Investigating the Spatial Variation of Renewable Energy Potential in the UK

Michael Blom – mike_geomatics@yahoo.com

September 2008

http://www.student.city.ac.uk/~abbj500/dissertation/UKWindSolarResourceTool.kmz

Submitted in partial fulfilment of the requirements for the degree of MSc in
Information Systems and Technology, City University, London

Abstract

When determining and reporting estimated energy production values, variables such as solar panel and turbine rotor size can be standardized, but the location variable - so important in determining accurate energy production values - is difficult to standardize. The location variable typifies the subjectivity inherent in calculating renewable-energy system outputs. With a variety of GIS tools, location is the variable analysed within this paper. This paper analyzes the entire UK and identifies what levels of renewable energy might be expected from both solar and wind renewable energy systems with respect to location. Met Office weather station data, taken from 2006, are analysed for wind speed and solar irradiation values, calculating probabilistic total yearly potential energy available from the respective renewable energies. Weather station data is interpolated into a raster surface to render full UK coverage with raster cell centroids and their potential energy attributes entered into a database. A visualisation tool is provided by an interactive virtual globe interface allowing users to query wind and solar energy values within defined regions. Confidence ratings are returned for each location to indicate the certainty with which estimated energy values are reported. The certainty of energy values is affected by factors such as the flux in wind and solar energy values from year to year and irregular energy values caused by micro-environments.

1 Introduction

Scientific research has worked hard to convince the general public that the use of alternative energies not only should, but indeed must need to increase in the future. There are very few who will still argue against this point. The debate as to which means of harnessing alternative energies is the most efficient (and just how efficient it is) has emerged as the most popular theme in scientific circles and consequently the press. Within this debate there are two alternative energy systems that receive the majority of attention: solar panels and wind turbines. Unfortunately, for the individual trying to gain a better understanding of these two systems, the statistics on their expected energy outputs are often unclear or misleading due in large part to each system's strong dependence on the location variable.

Gordon Brown installed solar panels in late 2005, and David Cameron had a small scale wind turbine installed in March of 2007 (The Times, 13/12/07). These two men have been portrayed by the press as a microcosm of the renewable energy debate, having chosen opposing systems. To interpret the opposite ends of the wide ranging media-opinion-spectrum, they've either taken steps towards contributing to a greener future while leading the way in advocating the use of renewable energy sources, or they've blindly jumped on the green bandwagon, and become caught up in increasingly popular "eco-cons" (Rowlatt, 2006).

Figure 1: Wind Turbine atop David Cameron's house in West London

Depending on which numbers are used, both sides of the argument can be successfully proven. The front page of *The Times* newspaper from December 13, 2007 quoted numbers stating that Brown will have recouped the investment in his solar panels in 100 years, while Cameron is looking at 60 years before he's recouped the cost of his wind turbine[1]. The next day, the political bloggers had picked apart these numbers which had ignored wind gusts, inflation of energy prices and life spans of the systems (Political Penguin 14/12/07, R Kyriakides's Weblog 14/12/07). The varying statistics in small scale renewable energy production are a direct result of the wide range of variables inherent in determining available wind and solar energy, the most important of those variables being location.

Lost somewhere in the renewable energy debate is the scientific truth, and in finding that truth lays the investigation of the spatial distribution of renewable energy. With such varying landscape and climate conditions in the United Kingdom, the location variable cannot be ignored when considering renewable energy production from the weather. The estimated energy outputs for a wind turbine on the shores of western Scotland are obviously not going to be applicable to a turbine on a roof in suburban London. If potential energies can be qualified into regions, and probabilistic statistics can be applied against annual weather records for that region, then the public can begin to understand a clearer idea of the effectiveness a wind turbine or solar panel might have at their home.

---

[1] No author was attributed to this caption which accompanied a photo of the systems atop the Brown's and Cameron's respective roofs.

Currently, the companies manufacturing renewable energy systems tend to err on the side of ambition. Solar manufacturer power ratings are based on ideal sunshine conditions, which rarely occur in the real world (Kemp, 2005), while manufacturer's wind energy projections very rarely see full projected energy values put to use (Gipe, 1999). Unfortunately, in order to simplify things for public consumption, small scale renewable energy harnessing installations are generally described in very broad terms, giving the assumption that one number can be applied across all of the UK, which simply isn't the case. With so many variables in flux, attributing a benchmark value to a renewable energy system will always be fraught with the highest levels of uncertainty. Manufacturers will assume a perfect setup, a well maintained system, and very generous efficiency in energy conversion in an effort to sell a few more systems. Applying manufacturer's projected energy to a specific location will be left up to the prospective customer, who will need to research and investigate expected wind or solar conditions. The values which are likely to be found are expected yearly average wind speeds and expected monthly solar values for a region, both completely inefficient for estimating expected energy from wind and solar systems. Once these values are found, the know-how of applying them to a wind or solar system is another hurdle.

This project explains why the current manufacturer's estimated energy values are often so far off what is actually achieved. A more accurate (specifically location based) and relevant (expected savings based on the current energy price) result is given for locations across the UK at a resolution that considers the network of weather station's available.

The relevance of determining spatial distribution of renewable energy in the UK can be seen in new laws which are being passed by most countries within the European Union. The UK has set a renewables target of 10% of our electricity being supplied by renewables by 2010, and also set out aspirations to double that share to 20% by 2020 (Wilson, 2003). Ever increasing amounts of money are being spent on determining ideal locations for large scale wind and solar factories, but the results of these studies aren't readily available to potential consumers of small scale wind turbines and solar panels. The power produced from optimally placed small scale renewable energy systems can be incorporated into the wider grid network and help the government to achieve its set goals.

This paper investigates weather patterns in the UK and provides a means of advanced geovisualization and user interaction in giving the public a service to determine the potential renewable energy that exists in their area, taking into consideration uncertainty resulting from station density. Users will become better equipped with accurate and relevant renewable energy information, allowing them to see beyond the vague and imprecise benchmarked values of solar panels and wind turbines.

2 Data

Data for this project has been obtained through the British Atmospheric Data Centre (BADC). The data is property of the Met Office, while the BADC is entrusted with licensing and distribution of the datasets. Since 1959 data has been stored digitally, and since 1998 the data has been stored in the Met Office Data Archive System (MIDAS) database which contains a relational structure (Met Office, 2007). The two data tables downloaded and used for this study were:

midas_radtob_200601-200612.txt

midas_wind_200601-200612.txt

At the time of downloading these datasets from the BADC website, the 2007 datasets were incomplete, so research has been conducted based on 2006 weather statistics. 2006 hourly wind dataset contains 1.9 million records, while the solar dataset contains 1.3 million records.

MIDAS datasets can only be downloaded by analysts once BADC staff have granted access to that analyst's user name and password. This access will be granted upon the signing of a release form by the analyst and their dissertation supervisor, stating the academic reasoning behind the need for the data.

A spreadsheet detailing station information was also downloaded from the BADC site: src_id_list.xls[2] This dataset which is freely available to the public contains a list of all the Met Office weather stations, past and present.

2.1 Met Office Weather Stations

The UK Met Office has a system of stations based at sites that are representative of an area up to several tens of km from the station (Met Office, 2007). The Met Office provided spreadsheet of these stations contains 50,051 records. A station location has a distinct src_id value, but as can be seen in **Figure 2**, there can be many records attributed to one src_id. Station records under the same src_id represent a different style of report being fed back to MIDAS, but originating from the same location. These records vary in the types of messages which they record and report, and also in the temporal aspects of their existence - many of the records represent a set of collections messages that are no longer being recorded at that particular site.

---

[2] <*http://badc.nerc.ac.uk/data/ukmo-midas/Docs/src_id_list.xls*>

	A	B	C	D	E	F	G	H	I	
1	SRC_ID	SRC_NAME	ID_T	ID	MET_DOMA	SRC_CAP_B	SRC_CAP_E	LOC	POST_COD	HIGH
396	150	ABOYNE NO 2	RAIN	845987	CLMSN	01/01/1961	31/12/1990	ABDN	AB34 5	
397	150	ABOYNE NO 2	RAIN	845988	CLMSN	01/01/1961	31/12/1990	ABDN	AB34 5	
398	150	ABOYNE NO 2	DCNN	1241	DRADR35	01/01/1995	31/12/1996	ABDN	AB34 5	
399	150	ABOYNE NO 2	DCNN	1241	NCM	01/01/1994	31/12/3999	ABDN	AB34 5	
400	150	ABOYNE NO 2	RAIN	845988	SREW	01/06/1990	31/12/3999	ABDN	AB34 5	
401	150	ABOYNE NO 2	RAIN	845987	WADRAIN	01/02/1992	31/12/3999	ABDN	AB34 5	
402	150	ABOYNE NO 2	WMO	3080	HCM	01/01/1992	31/12/3999	ABDN	AB34 5	
403	150	ABOYNE NO 2	WIND	124101	HCM	01/01/1992	31/12/3999	ABDN	AB34 5	
404	150	ABOYNE NO 2	DCNN	1241	CARLOS	01/05/1989	31/12/3999	ABDN	AB34 5	
405	150	ABOYNE NO 2	DCNN	1241	HCM	01/01/1992	31/12/3999	ABDN	AB34 5	
406	150	ABOYNE NO 2	RAIN	845987	WADRAIN	01/05/1989	31/01/1992	ABDN	AB34 5	
407	150	ABOYNE NO 2	WMO	3080	SYNOP	01/05/1989	31/12/3999	ABDN	AB34 5	
408	150	ABOYNE NO 2	WMO	3080	SREW	01/06/1990	31/12/3999	ABDN	AB34 5	
409	150	ABOYNE NO 2	WMO	3080	NCM	01/01/1992	31/12/3999	ABDN	AB34 5	
410	150	ABOYNE NO 2	RAIN	845988	NCM	01/02/1992	31/12/3999	ABDN	AB34 5	
60045										

Figure 2: MET_DOMA column describes the system of reporting while SRC_CAP_B and SRC_CAP_E represent beginning and ending collection dates.

The Met Office divides its stations in various networks which are designed to meet particular user requirements. The synoptic network meets the requirements for real time observations taken at intervals between 1 and 3 hours (Met Office, 2007). Stations used within this project are part of the synoptic network, reporting wind and solar values at hourly intervals is essential for the accuracy of this research.

According to the Met Office's numbers, the current UK synoptic network has an average station spacing of less than 50km (Met Office, 2007). This is the spacing which has been deemed as efficient to determine an accurate picture of the UK's weather characteristics. Weather elements which are more likely to vary over smaller areas, such as wind, have better coverage and are more densely reported than solar irradiance which tends to vary more slowly in relation to distance.

The Met Office's stated 50km UK wide coverage considers all domains of stations found within the synoptic network. In order to maintain consistency of weather data used throughout the project, only stations reporting under the HCM (hourly climate messages[3]) domain were chosen. This reduces the number of both solar and wind recording stations, and will negatively affect the resolution with which estimated energy values can be reported to users; however, it also gives this study improved standardisation and consistency in reporting.

[3] Hourly Climate Messages: <http://badc.nerc.ac.uk/data/ukmo-midas/ukmo_guide.html#4.4>

As the HCM name acronym suggests, the data in these tables was recorded on an hourly basis ensuring that precise wind and solar energy potential could be calculated at the station locations. Hourly recordings are the most precise weather datasets in the UK collected on a large scale. Knowing the weather at a station for every hour of any given year, it is possible to deterministically calculate how much energy was available at that station for that yearly period. The precision of these results can only be reported back as being accurate to the hour. More precise data is collected in the UK, but this occurs at isolated locations where detailed investigation into weather patterns is necessary. The Cardington (in Bedfordshire) dataset [4] is an example of wind speed collections occurring by the minute.

Station automation has allowed for hourly weather reports to be created on a wide scale. Unfortunately, an unavoidable consequence of automation has been the increased numbers of missing observations in the climate record due to station system crashes. When a station fails, it often ends up missing multiple days of reporting while repairs are conducted. For this project, the missing data from these stations is investigated to determine whether or not this station has collected substantial data, and not missed a large continuous period of reporting.

Site exposure is also a very key factor in weather reporting. The standard exposure at a meteorological station recording wind is over level, open terrain at a height of 10m above the ground. Obstructions and nearby ground features are removed from around stations as these will cause large discrepancies in wind speeds values, and also affect solar readings by creating shade (Met Office, 2007).

[4] <http://badc.nerc.ac.uk/data/cardington/>

2.2 *Solar Station Recordings*

Global irradiance is the value that best represents the energy that a solar panel would take up from the surrounding environment. Global irradiance is the total irradiance received on a horizontal surface direct from the whole sky including the sun. Diffuse irradiance is the total amount of solar energy falling on a horizontal surface from all parts of the sky apart from the direct sun (Met Office, 2007), and direct irradiance only considers irradiance which comes directly from the sun, and ignores scattered irradiance.

The Met Office's solar recording data reveals that 95 Met Office stations collected solar data on an hourly basis (HCM) in 2006. For a station to have recorded every hour over the period of a year, it would require 8760 recordings. In **Figure 3**, we can see that not all of the stations achieved this value: many stations had less than 8700 records, while some stations with over 8800 recordings had apparently somehow managed to record at an even higher rate than hourly.

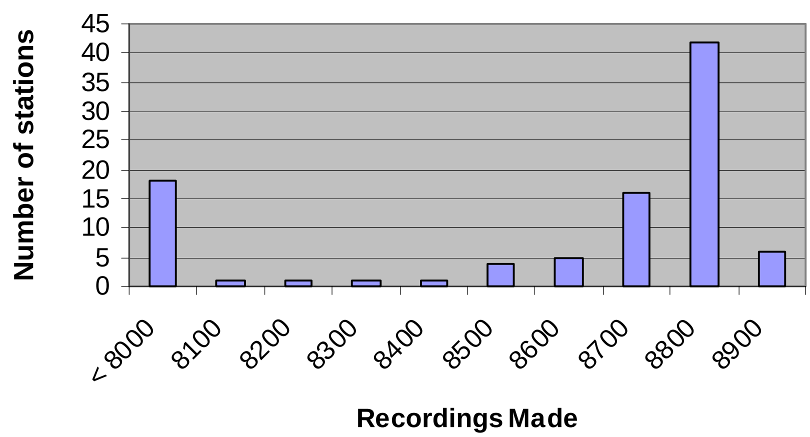


Figure 3: 42 of 95 solar stations made between 8700 and 8800 recordings in 2006 (8760 being ideal), 78 recorded greater than 8000.

If one were to calculate absolute (deterministic) irradiance values at each station over the period of a year, recordings for every hour throughout that year would be required. Where an hourly report was missing, assumptions for irradiance levels over that hour would be required. These assumptions could be based on the time of year while perhaps also considering the weather patterns surrounding the missing dates.

Allowing for inevitably missed data reports and considering all solar stations that managed to record 8700 reports (having then only missed 60 hours over the year) would leave 48 solar stations available for analysis. Unfortunately, the sparse distribution over the UK of these 48 stations negatively affects the confidence with which potential energy numbers can be reported back to users. Within this 48 station network the average distance any point in the UK is to a solar station is 42 km so the station to station average can be assumed to be approximately double that. This value of 84 km does not come close to the Met Office's stated station spacing of 50 km even distribution across the UK. Using only the 48 solar stations with 8700 or greater reports, there are many locations in the UK that would be over 150 km from the nearest solar station, with the maximum distance from a station being about 170 km.

In order to take advantage of a greater number of solar stations, the deterministic method of calculating a station's potential energy by summing the hour to hour records must be reconsidered. Instead of attempting to quantify exact potential energy values, the approach of qualifying station data – looking at distribution and classification of the station reports – allows the analysis of a location's solar characteristics to be performed in a probabilistic manner.

The immediate benefit is more flexibility where station reports went unrecorded due to station failure. Moving to a probabilistic approach in station analysis, stations which have missing gaps in weather reports can be considered. Station data collection patterns were examined to determine how much missing data was acceptable before a station should be excluded. Thirty missing days, or approximately 760 missing hours was chosen as the acceptable limit. The missing data reports would only negatively affect the precision of potential energy numbers if all the missing reports were from a consecutive 30 day period. If missing solar data was from 30 days in July and August, statistical analysis would report much lower expected energy values for a yearly period from the station in question. Station reporting patterns were analyzed to ensure that missing data did not fall into one lump period.

Choosing the arbitrary value of 8000 reports for a year rendered 78 solar stations usable. This created a much denser network of stations. In this network, the farthest any point is away from a solar station used in this study is 106 km; however, the average distance any point in the UK is from a station is 31 km, suggesting a denser network of stations spaced at an average distance of 62 km, close to the stated 50 km network which the Met Office uses to provide a good overall coverage of the UK.

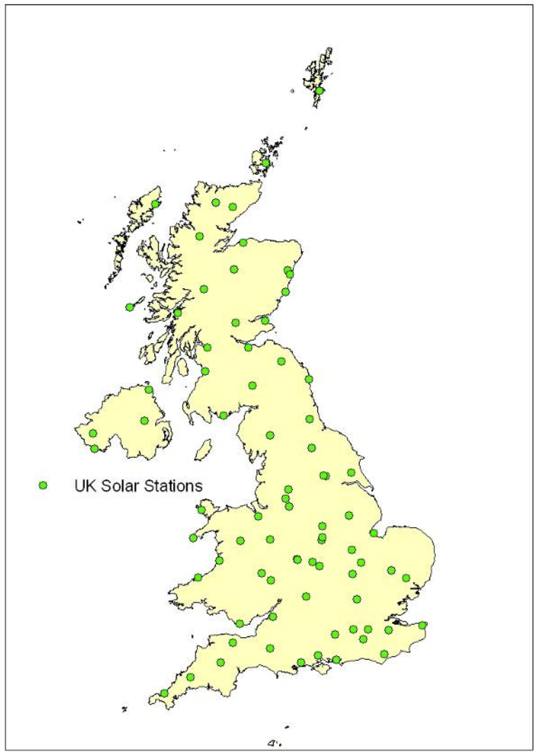


Figure 4: Complete set of 78 solar stations used in the research.

2.3 Wind Station Recordings

At Met Office weather stations, hourly wind recordings are represented by a wind speed mean taken from a ten minute period within an hour. Hourly data will also record a maximum gust value achieved over the one hour period. In the UK, these speeds are always recorded in knots (Met Office, 2007), but in the case of this project, for unit consistency, wind speed units have been converted to meters/second.

The hourly wind data recordings shows that 154 Met Office stations collected wind data (under the HCM station domain) on an hourly basis in 2006. As was the case with solar data, if what is needed is the absolute wind energy potential for 2006, recordings from every hour over the whole year are required. Limiting analysis to wind stations that reported back over 8700 records (losing only approximately two days of data recordings) renders 90 usable stations with an average distance of any point in the UK to a wind station being about 23 km. This suggests a 46 km average spacing station-to-station in our wind network. The farthest any point in the UK is from a wind station is approximately 78 kilometres.

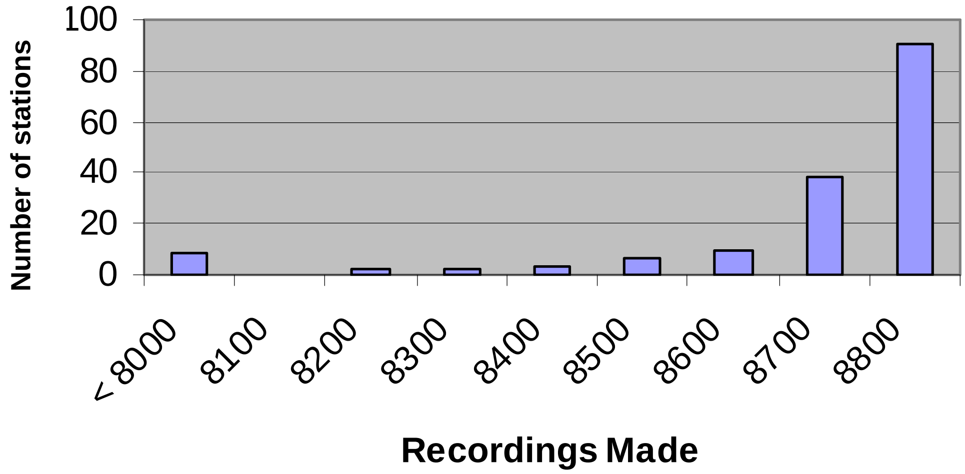


Figure 5: 90 of 154 stations reporting HCM hourly wind speed reported at least 8700 records, while 146 stations recorded greater than 8000.

As with solar data, the decision was taken to distribute wind data from each station into histograms, and utilise probability statistics to predict wind speeds and avoid deterministic calculations of potential energy. The cut-off point for missing data was again set at 8000 recordings, meaning that almost 30 days worth of data is missing in the worst case. However, this allows the opportunity to take advantage of a tighter network of stations, with 146 stations reporting at least 8000 records. Eight stations were thrown out in this process, each of those stations having missed more than 30 days worth of reports throughout the year.

Gust recordings are important for determining potential wind energy, as the power available from the wind is equal to the cube of the wind speed. In theory, doubling of wind speed will increase the available energy by a factor of eight. Consequently periods of gusts will contribute greatest to the potential energy available. Jack Park explained it best by stating the average of the cubes is greater than the cube of the averages (Gipe, 1999)[5]. To further clarify this statement, the average of the cube of different speeds (i.e. gusts) over time is greater than the cube of the average speed. While the Met Office does collect hourly maximum gust speeds, this single value is insufficient for the needs of accurately calculating energy available from wind gusts through out an hourly period.

High elevation wind stations also need to be considered, as they often report consistently high wind speeds that are not representative of lower elevation residential housing that would be attributed to the station. There are no clearly defined rules for wind station elevation cut-off levels, the acceptable elevations would need to be considered on a project by project basis. The town with the highest elevation in the UK is Shap, located at 412 metres. Considering that individual houses could be located at higher elevations, for this project the cut-off elevation for usable wind stations was set at 600 metres. One station, Cairnwell in Scotland at 933 metres was subsequently removed due to its high elevation. This station did incidentally report very anomalous readings throughout the year, attributable to its very high elevation.

[5] Comment attributed to Jack Park, early American pioneer in wind power

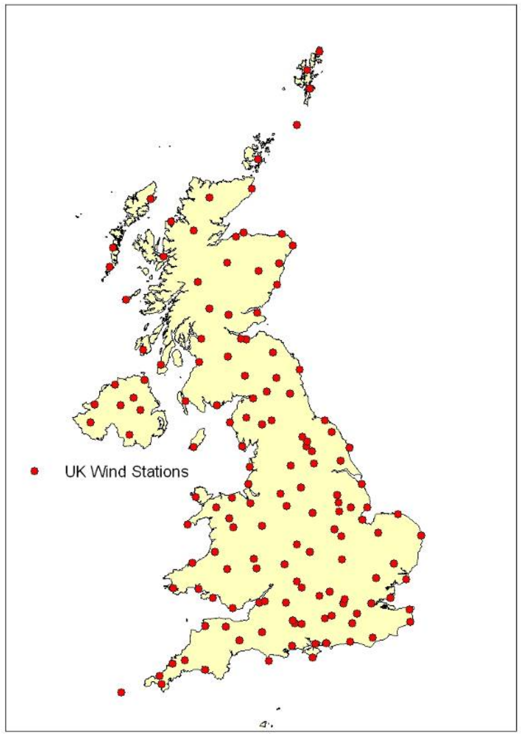


Figure 6: complete set of145 stations used in the research and analysis.

2.4 Data Summary

The problems inherent with automated weather stations periodically failing and missing record collections can be overcome through looking at probability distribution patterns of existing stations. A high density network of stations has been achieved, one that comes close to the Met Office's overall network which is describe as sufficient to give an overall impression of the climate in the UK. Stations which have not reported accurately or consistently have been removed, allowing calculations to be made more confidently, and expected energy values to be reported back to the client with improved precision.

3 Research Methodology

3.1 Calculating Renewable Energy at Station Locations

Estimation of both solar and wind potential renewable energy is done through probabilistic statistics, involving placing hourly station data and into discreet bins. Sampling of data into bins is a common practice in statistics, but bin sizing can be very arbitrary. Generally, the idea is to bin data into segments small enough to give good spatial resolution but large enough to give reliable sampling (Gastner & Newman, 2004). Bin sizes for this research were chosen with this principle in mind.

One algorithm is heavily relied upon for calculations within the wind analysis section of this project; the well known and accepted wind power from wind speed equation. An important modification is made on this algorithm by applying a weighting value to account for the extra energy that is available within wind gusts.

Solar energy values are collected by the Met Office directly in $W^{\cdot}H\ /m^2$ energy units; hence no radiation to energy algorithm is needed. For both solar and wind energy calculations, it is required to apply an energy conversion efficiency factor.

3.1.1 Calculating Potential Wind Energy at Stations

Each of the 145 wind stations used in this study has recorded somewhere between 8000 to 8800 weather reports over a yearly period, in line with a report being produced every hour of every day for a year, taking into consideration periods of downtime where stations are not operational. Looking at all mean hourly wind speeds for these 145 wind stations, values for 2006 were found to range from 0 to 54 knots.

An SQL query was built in order to group stations by their uniquely identifying src_id and classify the data into histograms with 27 bins representing 2 knots each. Running **Equation 1** below created 3915 (145 stations x 27 bins) separate insert queries.

```
select distinct
'insert histogram_table
select '+cast(src_id AS varchar(10))+', 2, count(*) from
dbo.windSpeeds_2006
where mean_wind_speed <=2 and mean_wind_speed >0 and
src_id ='+cast(src_id AS varchar(10))+'; insert histogram_table...
...........
...select '+cast(src_id AS varchar(10))+', 52, count(*) from
dbo.windSpeeds_2006
where mean_wind_speed <=54 and mean_wind_speed >52 and
src_id ='+cast(src_id AS varchar(10))
from dbo.windSpeeds_2006
group by src_id
```

Equation 1: portion of SQL query to create 27 separate insert statements for each of the 145 wind stations used in this research.

Table 1 below shows the classified data for the Fair Isle station in the Shetlands which resulted from running the SQL insert statement for station src_id = 3 created in **Equation 1**.

src_id	Windspeed_kts	frequency
3	0-2	293
3	2-4	611
3	4-6	908
3	6-8	959
3	8-10	951
3	10-12	850
…	…	…
3	44-46	2
3	46-48	2
3	48-50	0
3	50-52	0
3	52-54	0

Table 1: Wind speed data grouped into bins, with relative frequency calculated in last column.

With the data distributed into bins, it becomes possible to create histograms for each wind station. From here, comparisons and assumptions can be made as to how wind turbines at specific locations might perform.

The wind station at Hereford, Credenhill, consistently recorded some of the lowest mean hourly wind values in the UK.

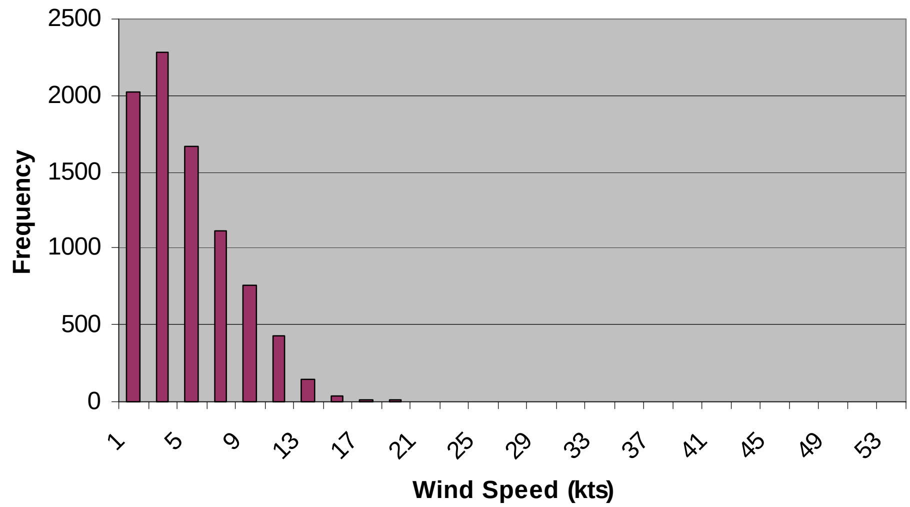


Figure 7: Frequency distribution of wind speeds at Hereford, England.

As shown in **Figure 8**, the wind station in Aberdaron, North West Wales, consistently recorded some of the highest mean wind speeds in the UK.

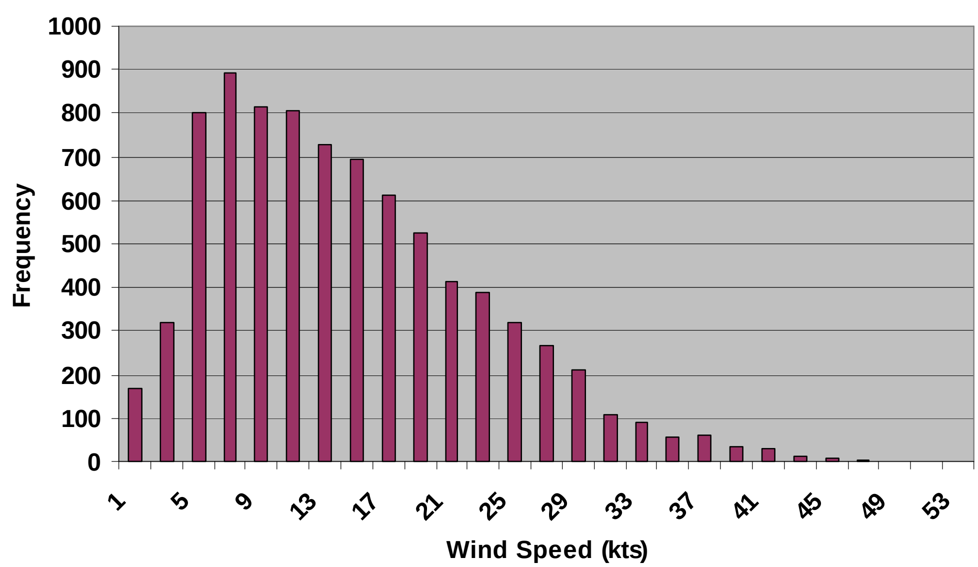


Figure 8: Frequency distribution of wind speeds at Aberdaron, North West Wales

The relative frequency values of wind speeds at these stations allow us to make assumptions about yearly wind speed values. The curves which can be derived from the histograms suggest how a station might perform. Curves with fatter tails and a mean positioned further to the right of the graph represent locations which are more likely to record high wind speeds on a regular basis, when compared to curves with a thin tail and a mean that sits more to the left of the graph.

Figure 9 looks at the Hereford, Credenhill station again, but this time considers relative frequency distribution. Through analysis of this graph, it can be concluded that this location will experience wind speeds of 3 kts (±1 kt) approximately 27% of the time.

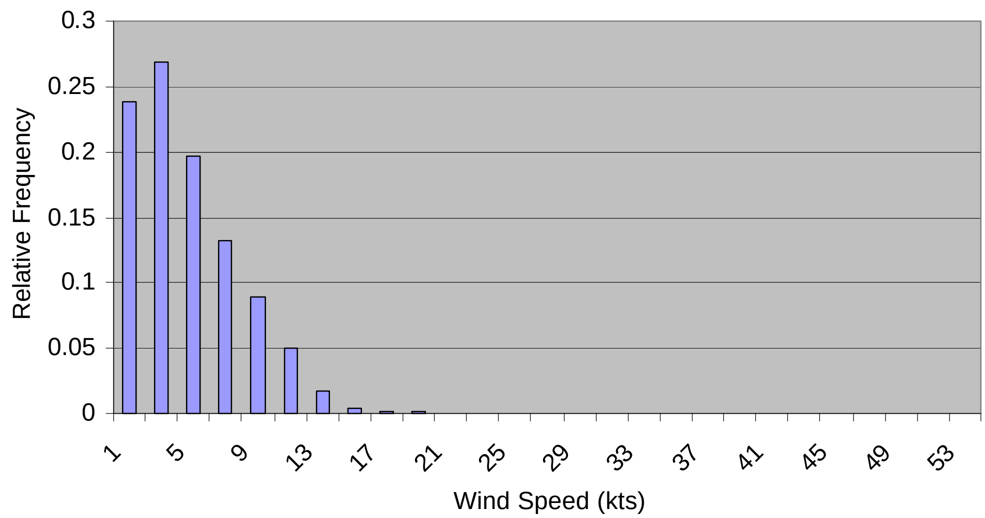


Figure 9: Relative frequency distribution at Hereford, England.

Figure 10 looks at Aberdaron station in the very windy North-West Wales again, and considers relative frequency distribution. During 2006, Aberdaron would have experienced wind speeds of 7 kts (±1 kt) approximately 10-11 % of the time.

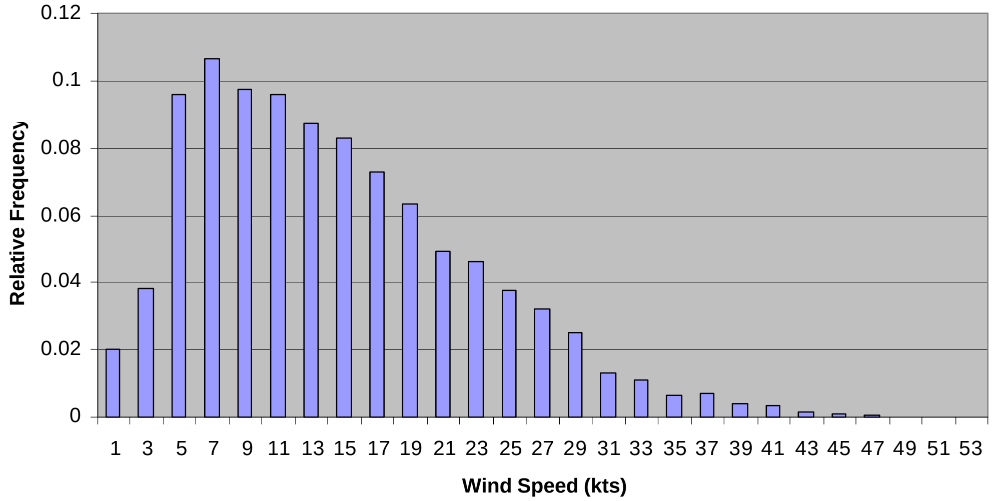


Figure 10: Relative frequency distribution at Aberdaron, North West Wales.

Summing all of the relative frequency values equals 1, or put in another way, gives 100% temporal coverage of wind characteristics for a given location. If the time period chosen is one month, it can be stated that at Aberdaron, wind speeds of 7 knots will be experienced over approximately 3.3 days.

$$\text{Number of Days at 7 kts} = (10.5\%) * (31 \text{ days}) \qquad \textbf{Equation 2}$$

Once the wind speed distribution at a location has been established, the next step is to estimate the power that this produces. To begin to determine a station's expected potential wind energy, the equation that calculates potential wind energy must be understood and applied correctly.

The usable power ***P*** available in the wind is given by:

$$P = \frac{1}{2}\alpha\rho\pi r^2 v^3$$ **Equation 3**

where P = power in watts, • = an efficiency factor determined by the design of the turbine, • = mass density of air in kilograms per cubic meter, r = radius of the wind turbine in meters, and v = velocity of the air in meters per second (Twidell and Wier 2006).

At its core, **Equation 3** is really just an extension of the conservation of Kinetic Energy Law, with an efficiency rating applied to determine what percent of kinetic energy from wind is actually converted to mechanical energy of a rotating shaft. Turbine efficiency has been reported as low as 7%, but is generally assumed to convert anywhere from 10-30% of potential wind energy into usable energy. The theoretical maximum power efficiency of any design of wind turbine is 16/27 or 59%, according to the Betz Limit law (Betz, 1926). This means that the most potential energy that could be converted to actual energy is 59% assuming that a wind turbine was able to operate at 100% efficiency. Considering Betz law, and the efficiency of the turbine rotor system, the average conversion of kinetic energy in the wind to mechanical energy is around 20% (Gipe 1999, Kemp 2005). This is the value used within this research for the efficiency factor (**•**) in **Equation 3**.

The mass density was calculated at each wind station, using the known equation for mass density at a known elevation.

• $= 1.225 - (1.194 * 10^{-4}) * (\text{Elevation})$ **Equation 4**

Equation 4 above was applied to each station using the elevation values provided in the Met Office station data.

The radius (**r**) for **Equation 3** was chosen to be one metre. Two metre diameters are a common size for domestic roof mounted turbines.

Using **Equation 3**, calculations were made to produce 27 different energy output values for each station. These energy output values are based on the 27 different wind speed bins, shown in the Power_watts column in **Table 2**. Each calculated value represents the energy in Watts (or Joules/second) that a turbine would be expected to produce at the wind speed used in the formula. From the Joules/second value Joules/day is determined, then KJ/year, and eventually the KW·H/year.

Wind_kts	Relative_frequency	Power_watts	Joules_day	Energy_KJ_yr
0-2	0.034349355	0.416841886	1237.0968	451.5403326
2-4	0.071629543	3.33473509	20637.9835	7532.863979
4-6	0.106447831	11.25473093	103510.803	37781.44299
6-8	0.112426729	26.67788072	259140.114	94586.14149
8-10	0.111488863	52.10523578	501910.861	183197.4643
10-12	0.0996483	90.03784743	775191.033	282944.7272
…	…	…	…	…
44-46	0.000234467	5071.71523	102742.367	37500.96401
46-48	0.000234467	5762.422236	116734.65	42608.14716
48-50	0	6513.154473	0	0
50-52	0	7326.412993	0	0
52-54	0	8204.698847	0	0
	sum=1		Sum=	8015546.725

Table 2: Potential Energy Faire Isle station in the Shetland Islands

The wind speeds recorded by the Met Office are of mean wind speed over an hourly period. While the Met office does record max gust speed occurring over the hour, it is impossible to absolutely determine wind energy potential unless the precise minute by minute, or even more desirable, second by second details of wind gust speeds are known at the concerned location.

Equation 3 used for energy production from wind can only calculate energy produced from wind at an instant in time. The instant in time being used represents mean wind speeds, but fails to consider gusts, which are vital to a turbine's energy output.

As was first mentioned in section 2.3, and in line with **Equation 3**, doubling the wind speed increase the available kinetic energy by a factor of 8:

Wind Speed = 2 kts $\quad v^3 = 2*2*2 = 8$ **Equation 5**

Wind Speed = 4 kts $\quad v^3 = 4*4*4 = 64$

Ignoring wind gusts implies that less energy can be extracted than it is available in the wind. This extra energy generated from unknown wind gusts outside of mean wind speeds is represented by a factor known as the "energy pattern factor" (Gholam, 2007), and sometimes called the cube factor.

$$= \frac{\Sigma V_i^3/N_a}{V_a^3} \qquad \textbf{Equation 6}$$

The energy pattern factor calculated from **Equation 6** is also often referred to as the cube factor due to the fact that it relates to the effect that cubing of wind speed has on energy output. The cube factor value gives an idea of the degree of deviation in wind speeds from the mean wind speed over a period of time - in the case of this study, one year.

By multiplying this cube factor against the expected energy output values calculated from mean wind speeds, the effect of energy dense wind gusts can be considered in determining a more accurate total potential energy value for a location.

3.1.2 Calculating Potential Solar Energy at Stations

The Met Offices solar stations record the global and diffused solar values. In order to consider the full amount of irradiation a solar panel would receive, global irradiation values which include direct and reflected irradiance were considered.

The same method of developing histograms that was applied to wind station data was also applied to solar irradiance data. However, hourly global irradiation values were much more widely ranging than recorded wind speeds. Global hourly irradiation values recorded in the UK in 2006 ranged from 0 to 3645 W·hr/m^2. A large majority of recorded values were at the lower end of this scale, and in order to show distinction between values at this lower end, a value of 30 W hr/m^2 was chosen as the discrete bin size. This bin size rendered 122 bins in total.

The solar station Thorney Island in West Sussex exhibited the highest solar levels over 2006. The absolute frequency histogram from this station is seen below in **Figure 11**.

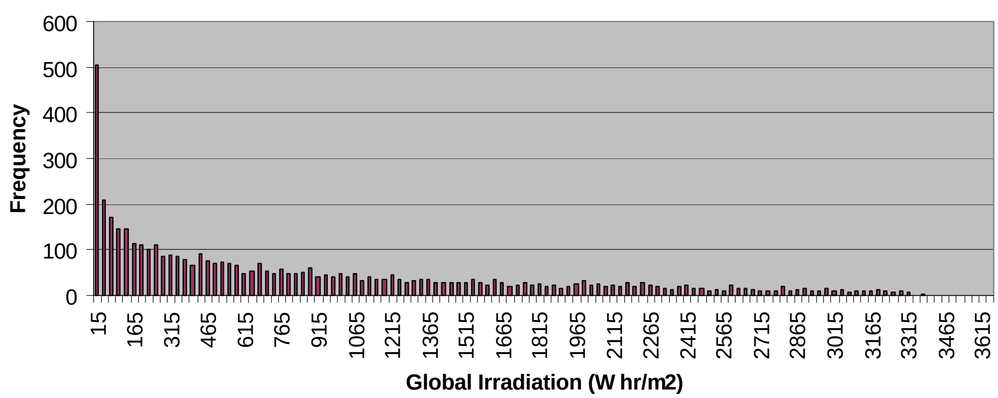


Figure 11: Absolute frequency distribution at Thorney Island, Wet Sussex.

The relative frequency histogram for the same area is below in **Figure 12**. Summing all the relative frequency values will equal one, representing 100% coverage for the time period of a full year.

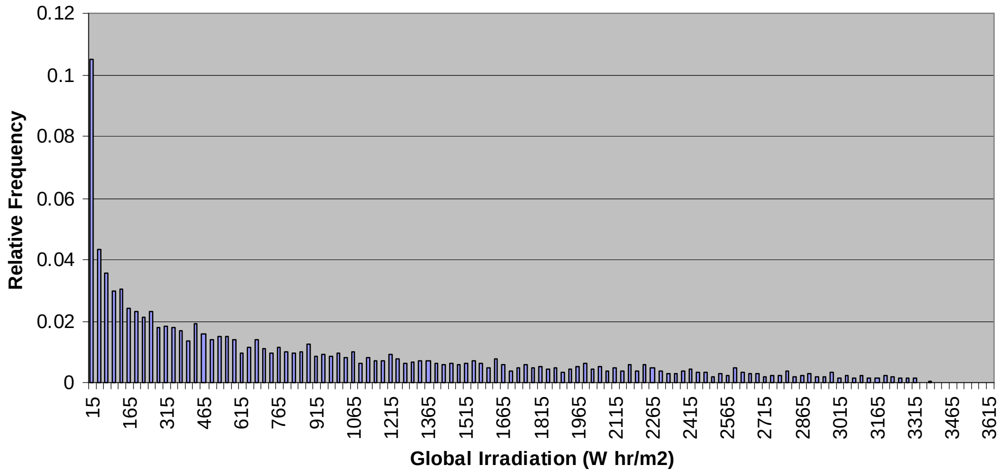


Figure 12: Absolute frequency distribution at Thorney Island, West Sussex.

Creating a relative frequency histogram of each station's solar reports throughout 2006 allows the extrapolation of yearly potential solar energy, in the same way that hourly wind values were extrapolated to create yearly expected wind values.

The solar irradiance values represent global irradiance in Watts received over a one hour period on a one square metre area. Once relative frequency values were calculated, the extrapolation to total KW·H at each individual station is a simple equation:

$$\frac{\sum_{i=0}^{i=121}((15+(15*i))*relFreq(i)*(8760hr/year))}{1000}$$

Equation 7

where i represents the bin value, of which there are 122 solar bins that increment by a value of 15. The relFreq(i) represents the probability of a solar reading occurring within the given bin.

Commercial solar panels generally convert 10-20% of solar energy into actual energy (Kemp 2005). Single cells that can convert up to 40% of irradiation into energy have been created in lab environments, but these are single cells rather than panels, and remain in development stages until lower cost production methods can be devised. For the solar panels in this project, a 15% efficiency rate has been assigned and a two square metre area has been assumed.

3.2 Point Interpolation

To determine complete and continuous coverage of the UK's potential energy, values taken from stations' point locations were interpolated to produce a smooth raster surface. The interpolation method used in this project is known as inverse distance weighted. In order to determine the most accurate representation of the overall energy potential in the UK, multiple inverse distance weighted (IDW) rasters need to be created, analyzed and compared.

Inverse distance weighting interpolation requires input of parameters from the user. There is no definitive right or wrong value for these parameters; creating surfaces through inverse distance weighting is an iterative process. It is left to the discretion of the analyst to create a surface that most accurately represents what the data is attempting to display. In the case of wind patterns, which are very sensitive to environmental changes, the variance of wind speeds in between stations can very difficult to accurately model and predict through simple interpolation.

Figure 13 displays the options that are available when creating a raster surface through IDW.

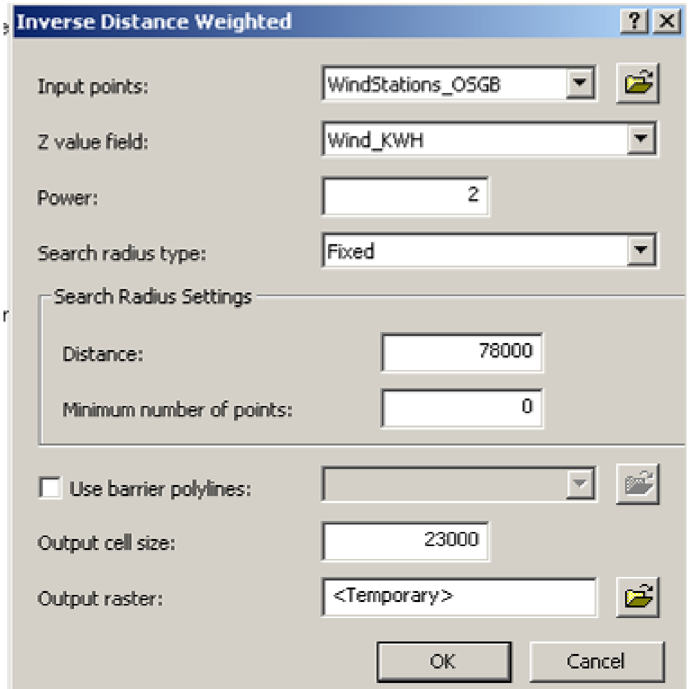


Figure 13: The dialog set up for inverse distance weighted wind speed raster surface.

The Z value field represents the field on which the raster will be based. This field, Wind_KWH, represents a specific station location's total wind energy available to a one metre radius wind turbine attributed with 20% conversion efficiency.

The Power field represents the degree that a station with a known value will influence a raster cell with respect to distance. Distance and influence are inversely proportional: a station's influence on a grid cell decreases as distance increases from the station. This follows what is regarded by many as the first law of geography: Everything is related to everything else, but near things are more related than distant things (Tobler, 1970). Higher power values cause exponentially greater influence of stations on the interpolated cell. A cell at **d** distance from a station will be weighted by multiplying the original stations value against $\mathbf{1/d^x}$ where **x** is the Power value, and **d** represents the distance from the centre of a raster cell to an influencing station.

$$w_i = 1/d_i^x$$ **Equation 8**

The search radius type is set to *fixed* rather than variable, in order to ensure that all raster cell locations consistently consider stations within and up to the same distance. Choosing *variable* search type results in a "whichever comes first" style search. A raster cell will be affected by either the first x number of stations surrounding it, or all the stations that fall within y distance, whichever occurs first. Maintaining consistency in inversed distance weighting, and using only the distance parameter to determine stations which influence a cell, it must then be ensured that every location within the UK will be influenced by at least one station. For the solar energy raster, a search radius size of 106 kilometres was used, to ensure that all points in the UK would consider at least one solar station, remembering that the farthest any point in the UK is from a solar station is 106 km. For the wind energy raster, a search radius size of 78 kilometres was used to also ensure that all points in the UK would consider at least one wind station. Raster cells values for both wind and solar were determined only on stations within these buffer sizes, and the variable search method was avoided to maintain consistency in the parameters used to determine a raster cell's value.

The output cell size parameter specifies the resolution of the raster cells which describe energy levels. Output cell size for wind and solar rasters was also based on the respective station network densities. Considering that the average distance any location in the UK is from a solar station for this project, the output raster cell size was set to 31 kilometres, while the output wind raster cell size was set to 23 kilometres. These two distances are significant because they represent the average distance that any location in the UK is from a wind station and the average distance any location is from a solar station.

The parameter which was experimented with in an iterative fashion was the power value. This value determines how distance from a station affects a particular cell's value. Two values were considered: one and two. A wind energy raster created using a power value of one, shown on the left in **Figure 14**, reduces the sphere of influence that anomaly stations have. A cell's value will still be affected by distances from stations, but to a lesser degree than if a higher value is used for power. The raster image on the right of **Figure 15** uses a power of two, and locations where there is a station with a anomalous value are more pronounced, as this station affects its neighbouring regions to a greater degree than when analysis is done with a power value equal to one.

Spatial interpolation, inverse distance weighting included, has been described as the GIS version of intelligent guesswork (Longley et al). This project has considered wind and solar properties and determined that a location should be highly influenced by weather stations which are closer to it, and consequently uses a power value of two to represent distance weighting in the creation of raster surfaces. This means that the weight given to a point (or weather station in this case) drops by a factor of four when the distance to the point doubles (**Equation 8**).

As is inherent with inversed distance weighting, no raster cell will have a value higher than the highest energy producing station, or lower than the lowest, all interpolated values will fall between the highest and lowest valued stations.

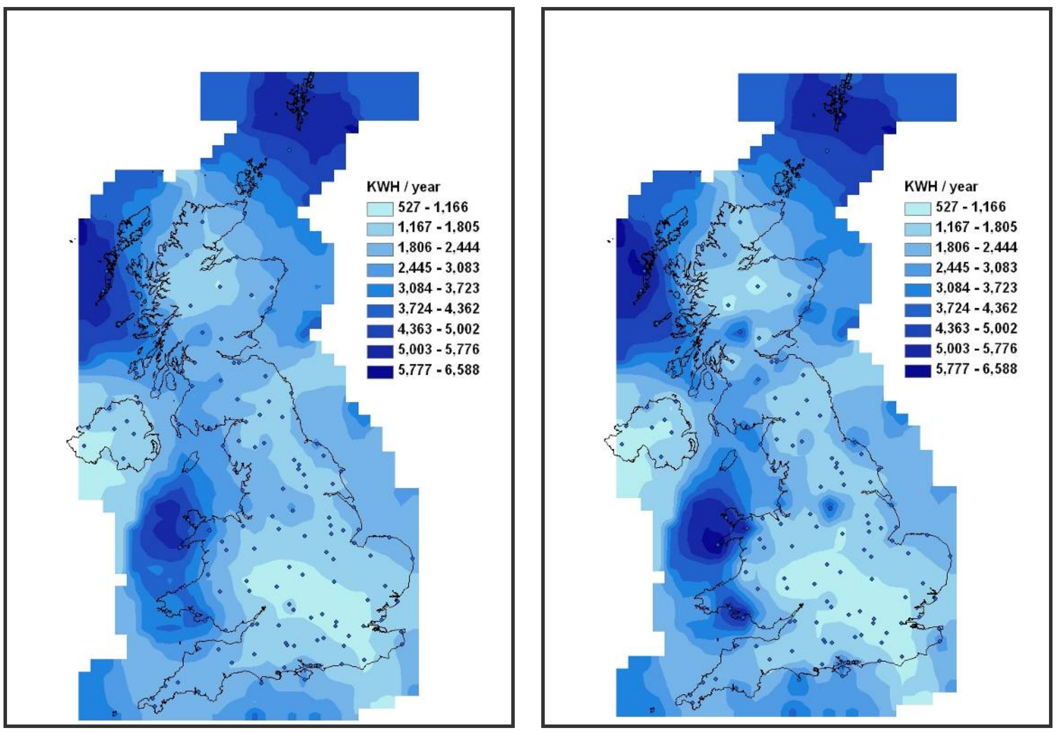


Figure 14 & 15: The two images shows Inverse Distance Weighted rasters of potential wind energy in KWH. The left raster uses a power factor of 1, while the right uses a factor of 2.

An inverse distance weighted raster acts as the data for the back-end system database for each of wind and solar energy values. Values taken from the wind and solar rasters will be reported back to the user. Inverse distance weighting interpolation for this project has been carried out while taking uncertainty into consideration, namely by varying the radius distance around a point which will influence that point, and by varying the resolution with which the raster is outputted.

3.3 Station Network's Influence on Uncertainty

The initial project scope did not include a means of reporting uncertainty of projected energy results back to the user. As the research was carried out, the importance of not only reporting back to the user expected energy savings, but also an indication of confidence in those values became evident. The goal of the project was to help users to break out from the idea of expecting a specific energy output determined by manufacturers without considering location. While this project improves vastly on general energy estimates which don't include localized wind speeds or solar radiation, there is the danger of users interpreting the results over confidently. Users need to consider that weather patterns can be very localized, especially in the case of wind. By reporting a confidence level back to the user, it's possible to understand degrees of deviations in wind speeds at given locations, and the sensitivity of wind speeds in relation to location changes. The likelihood of a specific level of energy being harnessed is important to avoid a continued misrepresentation of available renewable energies.

As has been discussed, the Met Office aims to achieve as a picture of UK weather as possible. However, the distribution of weather stations throughout the UK is inconsistent, and looking at overall station placement suggests that more weather stations are built around populated areas. Stations will also often record irregular values that are not in agreement with any neighbouring stations. This follows the thinking that some areas have more diverse weather conditions and do require more weather stations, but it is unlikely to be the case that more populated areas tend to have more varied weather conditions, it is more likely that populated areas are of greater interest, and hence contain more weather stations.

Regardless, the accuracy of estimated potential energy values across the UK will vary due to these inconsistencies. Where a user's requested location may be surrounded by 3 or 4 weather stations within a very short distance, all of them reporting consistent energy values, the estimated potential energy levels can be thought to be an accurate representation, whereas a location having two stations within that same radius size and reporting widely varying energy values, can not be assessed as accurately of the energy results being reported back.

As discussed in section 3.2, the potential energy values calculated for wind raster cells were determined by considering all stations within 78km of that location, while the solar cell values were determined by looking at all stations within 106km. These two distances will ensure that all locations in the UK will use at least one weather station in determining the expected energy levels from the respective renewable resource. Using these two distances within the interpolation process, detailed analysis can be carried out around the user's given location to determine the stations used in evaluating that location's potential energy cell.

Knowing the stations which are used to determine a raster cell's value, the two parameters most strongly influencing uncertainty are the total number of stations used, and the degree of deviation in expected energy values between the stations. The uncertainty rating system which has been specifically developed for this project, takes into consideration these two parameters. Confidence in reported potential energy values is based on a grading system that ranks the certainty of a returned energy value on a scale of one to ten, with one being not confident at all and ten being very confident.

Rating scales were kept the same between wind and solar station readings in order to accurately depict the differences in predictability between the two renewable energy systems. **Figure 16** shows the rating setup for the number of stations used.

Number of Stations	Rating Score
<=2	1
<=5	2
<=8	3
<=11	4
>=12	5

Figure 16: Confidence rating scheme relating to number of stations used.

A maximum value of five is attributed to a location if there are twelve or more stations within the search radius, while if two or less stations are used, a confidence rating of one is given.

Figure 17 shows the rating setup relating to the standard deviation amongst the stations being used.

Standard Deviation	Rating Score
<=300	5
<=800	4
<=1500	3
<=2200	2
>2200	1

Figure 17: Confidence rating scheme relating to number of stations used.

Standard deviation values were calculated using the following formula:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2}. \qquad \textbf{Equation 9}$$

Where N represents the total number of stations, $\overline{x}$ is the mean of the potential energy values for the stations used, and x_i represents a measurement taken at a station.

The standard deviation equation shown in **Equation 9** is a modification of the Root Mean Square Error equation where the mean error is assumed to be 0. The concept of the point RMSE has been widely applied to modelling of the accuracy of vector GIS (Shi, 1998), and in fact, many researchers have recommended the adoption of the mean error and the standard deviation as the basis of reporting error, superseding the RMSE (Fisher, 2005).

If only one station is used in determining a raster cell's value, a rating score of zero will be given for the standard deviation, and the total confidence rating score will thus be two.

3.4 System Architecture

Results for this research project are provided to the user through a 3-tiered tool based on the standard architectural structure of data - business logic - client interface. The overall architecture is shown in **Figure 18**, and encompasses all of the actors which play a role within the overall functionality of the system delivering location-based potential energy results to end-users.

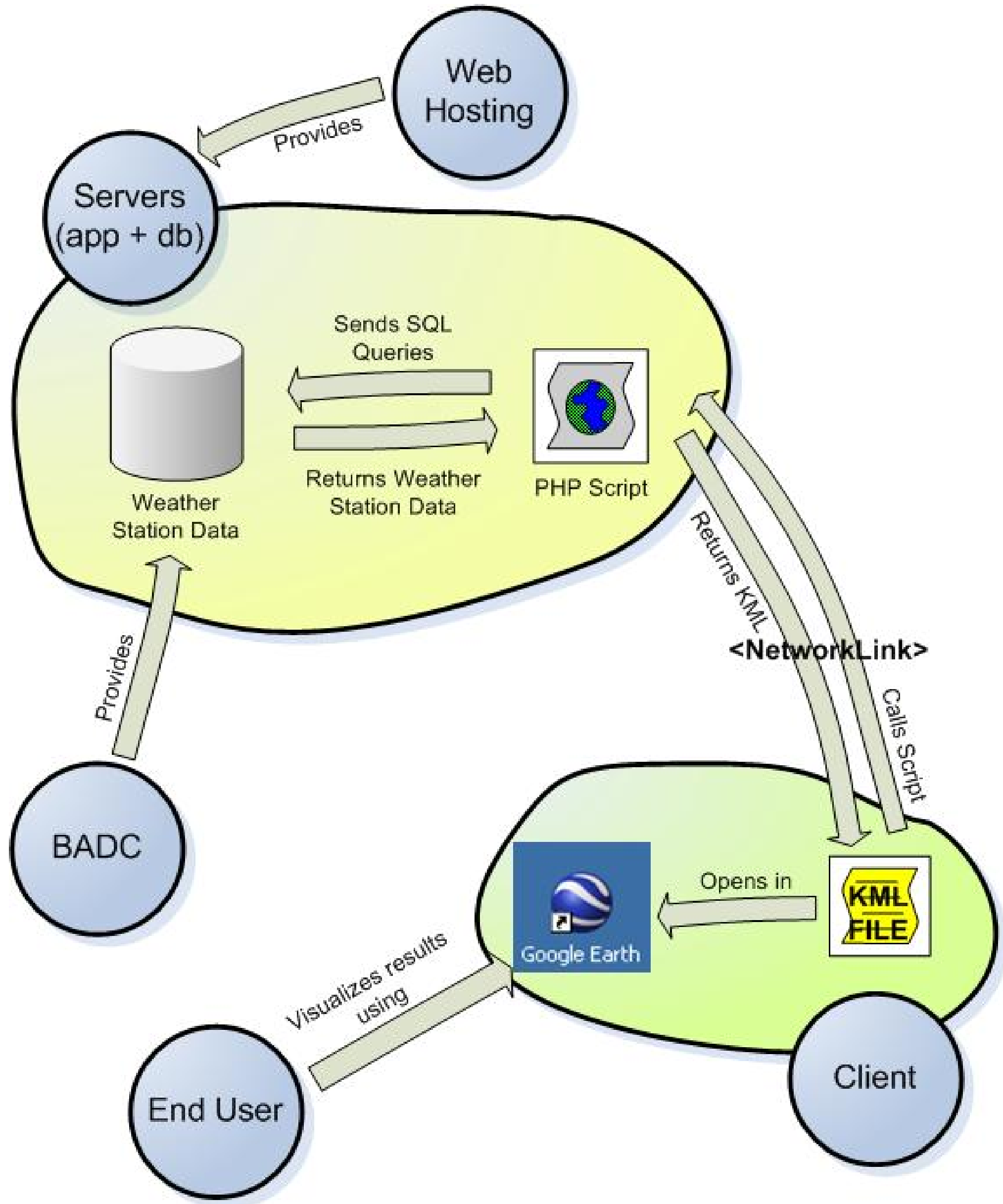


Figure 18: All the main actors within the project: the data, the PHP scripted business logic, and the client operating a digital globe.

3.5 *Database Setup*

Initially the project design called for the full 2006 hourly recordings data from the 78 solar and 145 wind stations to be stored in two separate tables. The calculations for potential energy at a location would be calculated on the server, synchronously, while the digital globe client waited for results to be returned. These deterministic calculations required intensive queries on the server's database. Trials of these queries on the MySQL tables to select the relevant 8760 or so records for each of the solar and wind stations revealed too much waiting time at the client end for the results to be returned. Pre-determining probabilistic energy values at each station helped to ease the work load required on the server side with each request from the client.

The freely available DBMS MySQL is used to store the necessary data. The proposal had initially discussed setting the system up on a privately hosted web server, but City University was able to provide access to MySQL databases on its servers. City University's "Vega" server currently hosts both the database and the application portions of this system. The MySQL database is protected by a username and password that is stored in a PHP script on Vega.

Five tables are necessary for the system to function:

1. 23km resolution wind energy point table (converted from raster)
2. 31km resolution solar energy point table (converted from raster)
3. DEM point grid of elevations (converted from raster)
4. Wind Stations table
5. Solar Stations table

3.5.1 DEM Grid Table

Elevation point grid data came from SRTM satellite which has recorded a full digital elevation model of almost the complete earth's surface. This data is recorded at 3" (80-90 metres) resolution. Using the ArcGIS raster to grid tool, the satellite's raster DEM data was converted to an evenly spaced point grid representing elevations above the ellipsoid. These elevation values are needed in the conversion process from Google Earth's WGS 84 latitude longitude to the OSGB Cartesian coordinates system used to calculate distances to weather stations used in the uncertainty calculations describe in section 3.2.

This elevation grid has a resolution of approximately 8 km, so each grid point is an average elevation of an 8 square kilometre area. This grid also plays a role in representing the UK as an evenly spaced grid dataset, a dataset which was used in analysis to determine the farthest any point in the UK is from a wind station. Throughout this paper, references are made to locations in the UK which are farthest from any wind or solar stations. This reference actually refers to the 8 km resolution grid point which is farthest away from any wind or solar station. Hawth's Tools[6] provides functionality to measure the distances between every set of points in two point datasets. This tool was run against the solar station dataset and the UK 8 km point grid, and then the wind station dataset and UK 8 km point grid.

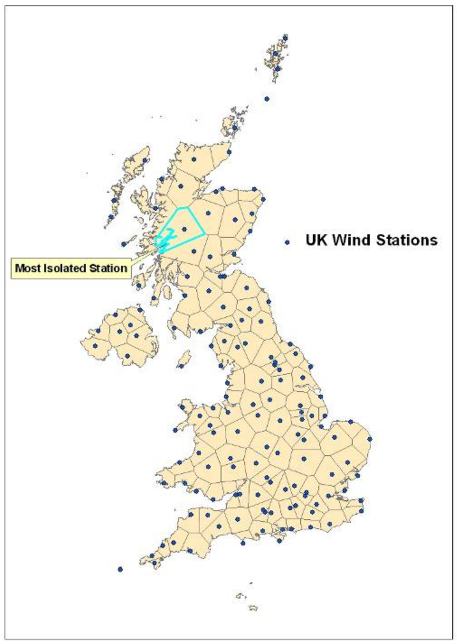
[6] http://www.spatialecology.com/htools/

3.5.2 Expected Energy Grid Tables

The model was first designed to consider the closest station to a user's location and apply the wind and solar energy attributes of that station to the location in question.

The closest wind and solar stations could be calculated through Pythagoras distance calculations. This method would result in discrete boundaries of potential energy regions. Stations having anomalous potential energy attributes could result in misleading data representation. If a station was isolated, the area over which the misleading data covered could be over 5000 km^2, the size of the largest region created from the Thyssien polygon algorithm in **Figure 19**.

Figure 20 shows the UK location (based on 8km grid) which is farthest from any wind stations. The buffer radius shown around this point is 78 km, within which only a single station is captured. Using the closest station premise would render this location influenced by only the single wind station, while ignoring the attributes of the four other stations which fall just outside the buffer.



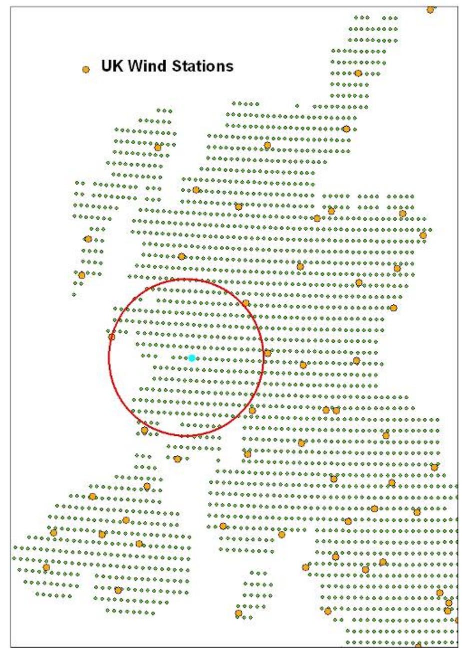
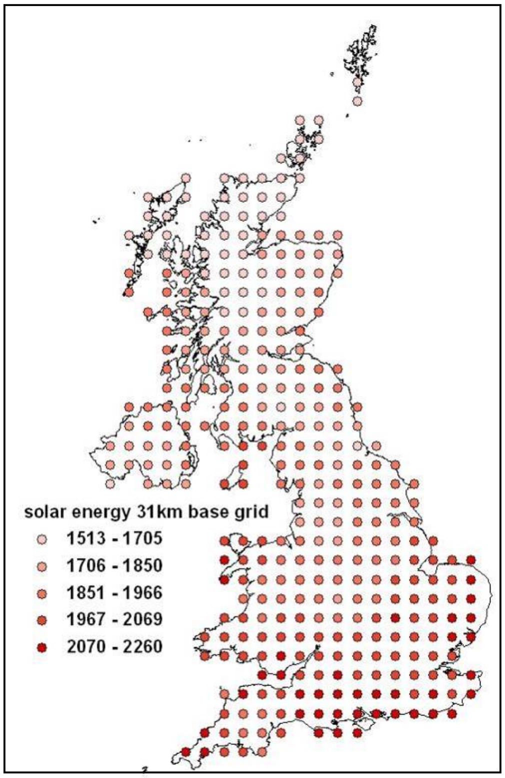


Figure 19: The largest region in the UK which is covered by a single wind station.

Figure 20: The location in the UK which is furthest away from a wind station, 78 kilometres.

Creating a continuous raster surface through point interpolation reduces the influence of irregular station data, and results in all locations throughout the UK being influenced by multiple surrounding stations rather than simply the one closest station.

The inverse distance weighted algorithm used ensures that any location's value will be more greatly influenced by stations which are closer to it. In the case of **Figure 20**, the location will be more or less equally influenced by five wind stations.



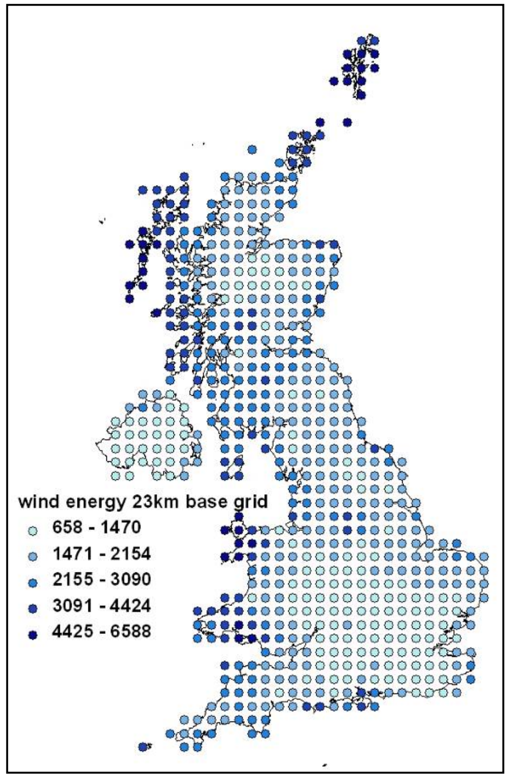


Figure 21: Solar Grid acting as the base point dataset for the database table “solarGrid”, units KW·H, 15% efficiency, $2m^2$ solar panels

Figure 22: Wind Grid acting as the base point dataset for the database table “windGrid”, units KW·H, 20% efficiency, 1m radius blade

ESRI’s raster to point conversion tool is utilised again here. This tool allows the wind and solar rasters to be converted to grid point layers. The two grid point layers contain attribute tables with columns that represent each point’s x and y coordinates, as well as the expected energy at the point. A SQL script is built around the tables to allow the data to be imported into MySQL tables. Once loaded into the database tables, the data is used as the base datasets for regional potential energy levels, to be queried by and reported back to the client.

3.5.3 Wind and Solar Station Tables

One table each represents the wind and solar stations. These tables contain a unique ID field, the x and y OSGB coordinates of the stations and the amount of potential energy available at the station after assumptions on parameters such as conversion efficiency, and solar panel/wind turbine size.

In order to determine uncertainty levels in expected energy outputs that are being reported back to the client, it is necessary to know the locations of stations which have played a role in evaluating that areas expected energy levels, and the energy levels at those locations. The two station tables provide the basis for queries relating to the evaluation of these uncertainty levels, as they allow calculations on number of stations used to determine energy levels and the standard deviation between those stations to be calculated.

3.6 *Setting up Business Logic*

3.6.1 IDE and Script Architecture

All code can be found in appendix A.

The Eclipse[7] IDE is used as the development environment for scripting and debugging the business logic code. Eclipse is a popular open source development environment. An extension for Eclipse, PHPEclipse[8], implements PHP development functionality into the Eclipse environment. Debugging in Eclipse was performed by outputting variables onto the console screen through the echo statement.

The business logic is written in PHP, a server side scripting language. The code was divided in four PHP files which could be considered as classes in more object oriented languages.

---

[7] http://www.eclipse.org/downloads/

[8] http://sourceforge.net/projects/phpeclipse/

phpsql_dbinfo.php: stores all the data necessary to connect to the database. This includes a username and password, and the database and server names.

constants.php: stores values that relate to conversions from the WGS 84 geographic coordinate system to the Cartesian coordinate system used, Ordnance Survey Great Britain.

projectionFunctions.php: contains methods to convert between WGS84 geographic coordinate system and the Cartesian coordinate systems upon which distance measurements can be made in determining the station which is closest to a user's location.

createkmlSimple.php: contains the code to initialize the database connection at the beginning of the business logic, and at the end of the business logic to construct the xml based kml, using the appropriate tags.

functions.php: contains a method to calculate distance through Pythagoras, and also the method to rate confidence levels through evaluation of total number of stations used and standard deviation in potential energy values amongst those stations.

windCalcs.php: contains the business logic behind querying the windGrid table to determine the closest grid point to the user's location. The logic then queries the windStations table to determine all wind stations which fall within 78km of the location queried by the client.

solarCalcs.php: contains the business logic behind querying the solarGrid table to determine the closest solar grid point the requested location. The solarStations table is then queried to determine all solar stations falling within106km of the location queried by the client.

3.6.2 Retrieving Users' Locations

The business logic is instantiated by a call to the createkmlSimple PHP script by the client side application, Google Earth. The PHP script receives from Google Earth the bounding box coordinates of the user's current view. These coordinates come in a single string, and so need to be split (or 'exploded') into their four individual elements: North, East, South and West.

Using these four bounding box coordinates, the centre of the client's view is calculated. The coordinates passed from Google Earth are in WGS 84 geographic coordinate system (latitude and longitude, in decimal degrees). WGS 84 is the most widely used and internationally recognized datum defining the shape of the earth, and as such has been adopted as the default datum for use in virtual globes.

3.6.3 Datum Transformation

It is necessary to transform Google Earth's latitude and longitude values to Ordnance Survey Grid coordinates (X,Y). Within the PHP scripting, distance calculations are used to determine wind and solar stations within specified distances of the user's location. Distance calculations are more easily calculated from a projected coordinate system rather than a geodetic system because measurements can be made in two dimensions, avoiding the need to consider the curvature and shape of the earth. It should be remembered that in order to portray meaningful relationships for a complex, three-dimensional world on a flat sheet of paper or a video screen, a map must distort reality, or tell "white" lies (Monmonier, 1996).

The datum transformation process is broken up into four main stages. The coding for these processes resides in the projectionFunctions.php script, the details of which can be found in appendix A. The steps are described below:

WGS84 Lat, Long, Ellipsoid Height to WGS84 Cartesian XYZ

In order to convert from a GCS based on latitude and longitude to a Cartesian coordinate system, we need the height above the ellipsoid, h.

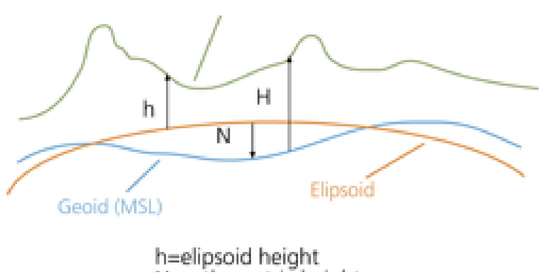


Figure 23: Height above ellipsoid is given by SRTM digital elevation model, but Google Earth reports height above Mean Sea Level.

The dempointsUK table (a table of 8km resolution grid points) is first queried to determine which grid the user is located in:

```
SELECT ELEV FROM demPointsuk
WHERE Latitude >= '.$userlat.' and
Longitude >='.$userlon.' order by
Latitude, Longitude
```

Equation 10

This query will look for all grid points to the North - East of the user's location, and by ordering the results by latitude and longitude, we return the point to the North – East which is closest to the user's location, a point which is a representation of the mean elevation above the ellipsoid in the surrounding 8 km^2 area.

With the known latitude, longitude and height above ellipsoid, the appropriate functions are applied to achieve the conversion of GCS coordinates to an ellipsoid based XYZ coordinate system.

Helmert Datum Transformation (WGS84 to OSGB36)

The XYZ coordinate based on the WGS84 ellipsoid is then converted to the XYZ coordinate system based on the OSGB36 datum. Both sets of XYZ coordinates describe the same spot on the earth, only their interpretations of the earth's parameters are slightly different, hence so are the relating XYZ values.

The Met Office had provided grid coordinates for Northern Ireland's stations in OSNI grid. During energy analysis, once a raster surface for wind and solar energy had been created for the UK and converted to a point grid, the OSGB coordinate system was adapted for both Great Britain and Northern Ireland.

OSGB36 Cartesian XYZ to OSGB Lat Long and Ellipsoid height

The description of the points' location is transferred back into latitude and longitude values, but now these values are based on the OSGB36 datum rather than WGS84.

OSGB36 Lat Long Ellipsoid height to OSGB36 Easting and Northing

The final British National Grid coordinates are calculated. The Northing and Easting results which are calculated at this final stage can now be compared to the Northing and Easting coordinates that are attributed to the stations in our Met Office stations table. Standardisation of coordinates systems has been achieved.

3.6.4 Querying Accurate Wind and Solar Energy Levels

This code exists in the windCalcs.php and solarCalcs.php files. Knowing the user's Cartesian XY location, the system needs to query for each of the closest grid points from the two energy raster grids. The script runs four queries, selecting the grid points to the North-East, North-West, South-East and South-West. This script, one of the four, will query for the closest grid point to the North-East.

```
SELECT * FROM (SELECT *
FROM windGrid where OSGBX
>= '.$OSGB_X.' and OSGBY >=
'.$OSGB_Y.' order by OSGBX
asc, OSGBY asc) s limit 1
```

Equation 11

where $OSGB_X represents the user's X coordinate location, and $OSGB_Y the user's Y coordinate location. Iterating through these four grid points surrounding the user's location, the point falling closest to the user is selected. This process is performed for wind and solar grids, and the resulting solar and wind grid points contain the details of how much potential energy is available from respective resources at that point. A solar grid point should never be more than 22km away from a user's location, and a wind grid point should never be more than 16km away from a user's location, these two distance values being the case of a user's location was in the exact centre of a set of four grid points.

Within the PHP scripts, a price of 12 pence per KW·H is set for the price of energy, and applied in the code against the final KW·H value so that potential energy levels can be converted to potential monetary savings.

3.6.5 Querying Uncertainty in Wind and Solar Energy Levels

Uncertainty rating is carried out in the **functions.php** script. Each grid point in the solar grid has been evaluated based on all solar stations within 106 km of the point while each grid point in the wind grid has been evaluated based on all wind stations within 78 km of that point. Knowing the radius distance value that each grid cell has used to consider stations used, a query is built to select all the stations falling within that distance, and the total count of those stations and their standard deviation is analyzed. The expected energy value of these stations can also be analysed, and the variance in their values will give an indication as to the consistency of wind speeds in the area, giving and indication of how confidently values can be reported back to users.

3.6.6 Formatting and Outputting KML

Included with PHP version 5, the DOM library assists in creating properly syntaxed extensible markup language (XML) documents. KML (Key-hole Markup Language) is a derivative of XML, hence the DOM API allows for creation of the open source KML.

The final echo statement outputs the formatted KML string, which is sent back to the virtual globe client. The KML contains the user's location and a callout balloon detailing the closest solar and wind stations, and the amount of money that could be saved based on a standardized solar panel size of $2m^2$ and 2m diameter wind turbine.

3.7 *Client-Side Structure*

In the last five years one of the biggest innovations to come out of the World Wide Web has been the virtual globe, with Google being the first to bring high resolution satellite imagery to the desktop through their Google Earth geo-browser.

In order to describe points, lines and polygons within their virtual globe, Google originally designed and created a derivative of XML called KML which contains a library of tags with semantics related to the study of geography. Probably the most significant KML tag, in terms of adding functionality, is `<networklink>`. This tag allows the front end Google Earth to query remote databases from remote servers. These databases can be static, as is the case in this project, or they can be regularly updated, providing live (or almost live) data to the Google Earth client.

The `<networklink>` tag of KML embeds a `<viewRefreshMode>` tag that further defines when the network link is to be called. By setting the value within the tag to 'onstop', the network link will be called once the view in Google Earth has come to rest.

More about Google Earth is discussed in section 5.3, Google Earth as a visualisation tool.

3.8 Operating the System

The Google Earth KMZ file is available for download from the following location:

http://www.student.city.ac.uk/~abbj500/dissertation/UK Wind and Solar ResourceTool.kmz

The tool requires version 4.2 of Google Earth, available for download from here:

earth.google.com/download-earth.html

The KMZ file should automatically open in Google Earth. Because the tool has been designed for the UK, it might be necessary to shift the view to be centred over the UK in order for the functionality to begin reporting expected energy values.

Panning the view around the UK will see the green location indicator circle update after a second or two, and be placed at the centre of the view. Stations used to determine the expected energy savings value at the green dot location will also be updated, but this might not be evident unless zoomed out to a larger scale.

4 Analysis and Discussion

What has been modelled in this project is potential energy from wind speed and solar irradiation in the UK; well known methods of inverse distance weighting have been used to accomplish this. The innovation exists in delivering the results to the user in an open source KML file to be viewed within a virtual earth browser. The challenge within this project exists in modelling to an acceptable degree of precision to allow anyone in the UK to understand the potential energy that could be harnessed at their location from these two resources.

In creating any model it must be remembered that all models are an abstraction of reality (Wainwright, 2004). Within this project assumptions are made based on how certain variables have behaved in the past in moving from reality to model. The affects that this transformation from reality to model has on precision have been quantified and provided as part of the final output to the client.

This research does not attempt to determine the definitive answer to which renewable energy source is more profitable than the other. There are too many variables which revolve around initial set-up cost that this project does not take into consideration. However, an attempt to put wind and solar systems on a level plane is made by using 2m^2 solar panels, and 2 metre diameter turbines as the standards to base comparisons on. These values are chosen as the standards because they were deemed to be a good average size for the systems that would be set up for a small scale home based renewable energy system.

4.1 *On Wind Energy Results*

When considering wind speed distribution, it is most accurate to look at a PDF, or probability distribution function (Barnsley, 2007). In studies of wind speed it is common to use either the Weibull distribution or the Rayleigh distribution as a PDF (Bowden et al, 1983).

By looking at data in terms of frequency distribution, probabilistic maths has been applied to consider what might happen in the future. Assuming that 2006 has been a good representation of an average year for wind speed across the UK, the histograms created in this research have been interpreted to produce expected yearly results. The Weibull curve generated through linear regression of empirical data will also take into account extreme events, leaving a small part of the curve's tail to consider events larger than the largest recorded for 2006.

The Weibull distribution curve is defined in terms of two parameters, c and k:

$$\Phi(u) = \frac{k}{c}\left(\frac{u}{c}\right)^{k-1}\exp\left[-\left(\frac{u}{c}\right)^{k}\right]$$

Equation 12

where u is the wind speed, • (u) is the probability density of wind speed u, c is a scale parameter, and k is a shape parameter. The Rayleigh distribution is a special case of the Weibull distribution where k = 2 (Barnsley, 2007).

For the wind speed data at Aberdaron in North West Wales, linear regression resulted in the converging of values of c and k after only 5 iterations:

```
After 5 iterations the fit converged.
final sum of squares of residuals : 0.00141416
rel. change during last iteration : -9.29844e-006
|
degrees of freedom (ndf) : 25
rms of residuals      (stdfit) = sqrt(WSSR/ndf)      : 0.00752106
variance of residuals (reduced chisquare) = WSSR/ndf : 5.65663e-005

Final set of parameters            Asymptotic Standard Error
=======================            ==========================

c               = 8.58268          +/- 0.1947       (2.268%)
k               = 1.9159           +/- 0.06288      (3.282%)
```

Figure 24: Values for c and k were determined after only 5 iterations

Convergence after only five iterations would indicate a wide distribution of wind speeds at Aberdaron, and looking at the distribution of data in **Figure 25**, and comparing the modeled results with **Figure 27** from Hereford, this can be seen to be the case. Using these determined values for c and k; 8.58268 and 1.9159 respectively, and fitting our stations wind speed and relative frequency data to the Weibull algorithm, we are able to plot a smooth probability distribution function in gnuplot[9]:

[9] http://www.gnuplot.info/download.html

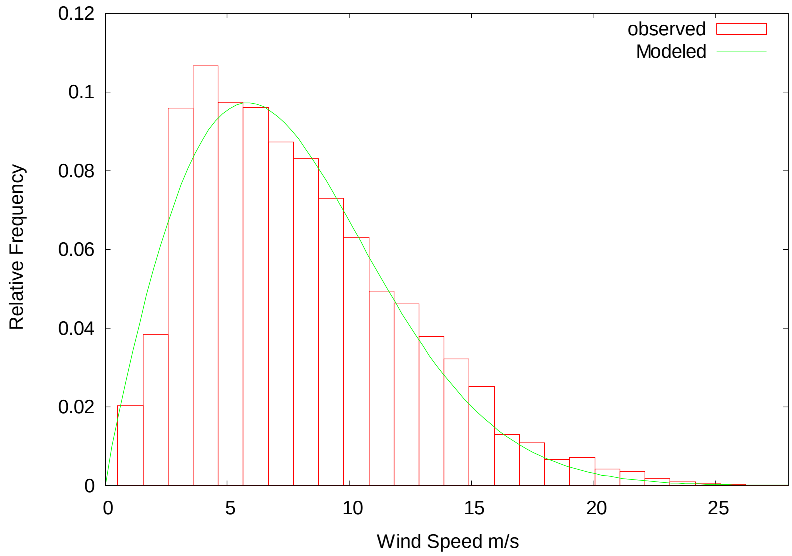


Figure 25: Modeled relative frequency distribution at Aberdaron.

The station that produced the lowest wind speeds, Hereford in South East England converged on values for c and k after 21 iterations:

```
After 21 iterations the fit converged.
final sum of squares of residuals : 0.00112826
rel. change during last iteration : -6.37455e-007

degrees of freedom (ndf) : 25
rms of residuals      (stdfit) = sqrt(WSSR/ndf)      : 0.00671793
variance of residuals (reduced chisquare) = WSSR/ndf : 4.51306e-005

Final set of parameters            Asymptotic Standard Error
=======================            ==========================

c               = 3.01616          +/- 0.04413      (1.463%)
k               = 1.66627          +/- 0.03582      (2.15%)
```

Figure 26: At Hereford, values for c and k converged after 21 iterations.

Convergence of c and k at Hereford was seen to take a relatively large amount of iterations when compared to Aberdaron. This is due to the fact that wind speed values are mainly grouped within the first three bins, in accordance with the low wind speeds that were experienced at Hereford. This has consequently created a modeled curve which is not as smooth as that at Hereford, and a curve that is harder to define through linear regression.

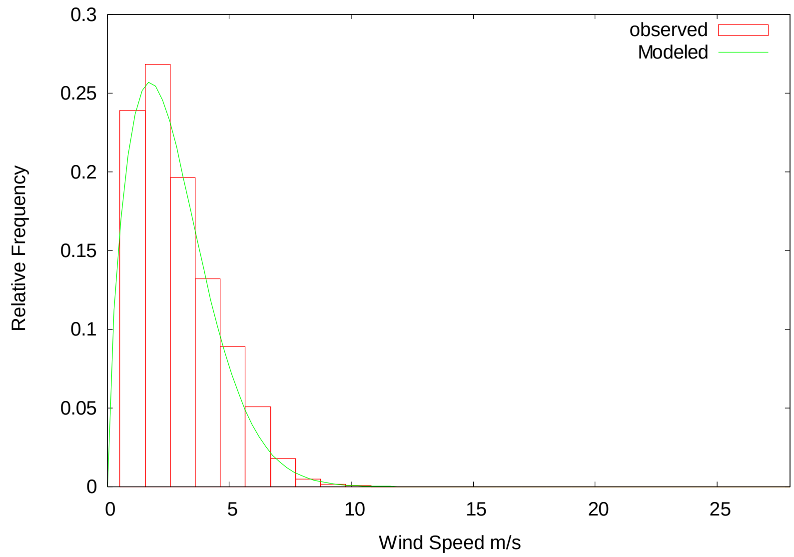


Figure 27: Relative frequency distribution at Hereford

The modeled curves in **Figures 25** and **27** represent the probability of a wind speed occurring at any given instant in time. In the case of the curve for Hereford, we would expect to experience wind speeds of 2 m/s about 27% of the time.

The area under each of these two curves for Aberdaron and Hereford is equal to 1, representing a dimensionless ratio value of 100%. By assuming that value of 1 to represent a year, yearly expected wind speeds were interpreted and expected energy results calculate.

The Cube Factor Formula attempts to quantify how positive deviations from mean wind speeds (wind gusts) affect increases in potential wind energy. Cube factor is very important when considering the energy available from wind at a location. It is important to understand however, that while the cube factor is often used to add a gust element to wind power calculations, it represents a degree of standard deviation amongst mean hourly wind speeds, and this might not always be a good indication of wind gusts. As the time period being analysed becomes larger, for example looking at a month's data versus a year's data, the possible explanations for varying mean wind speeds increase. Most notably, a change in the seasons will result in varying mean wind speeds being recorded.

The cube factor can be thought to represent the tail of a wind speed distribution curve. For a curve representing modeled wind speeds, a fatter tail signifies increased gusts, and consequently more energy dense winds. Wind speed curves from three different locations can potentially all have the same mean wind speed, but the difference in their tails gives an indication as to the potential energy that can ultimately be expected from a location whose wind patterns follow this curve. The fact that three locations with identical mean wind speeds can produce widely different energy values stresses the importance of including the cube factor into any wind energy calculations that are based on mean wind speeds.

In the UK, the highest wind cube (energy pattern) factor for 2006 was 2.99, found at Senny Bridge No 2 station, while the lowest wind cube factor of 1.59 was at found at High Wycombe station.

Inverse distance weighting interpolation was used in this research because of its status as the "workhorse of spatial interpolation, the method that is most often used by GIS analysts" (Longley et al). The IDW algorithm is not overly complicated, while providing a reasonable estimate of values where measured values don't exist. Using inverse distance weighting can itself be considered modelling. Values are being estimated, attempting to move from observed reality to an abstraction. The fundamental question is how accurate an interpretation of reality is the inverse distance weighting algorithm being used? Without going into too much detail about wind pockets, wind tunnels, and the effects of ground features on wind speeds, it can be said that wind is very localised element if nature. Thus, by using a power value of two in the interpolation algorithm, increased weighting is given to a location's closest stations, while still taking into lesser consideration all stations falling within 78 km.

While wind speeds can be seen to vary widely across small distances, the scale of potential wind energy variance across the UK is incredibly high. The range of expected wind energy outputs from a 2 metre diameter turbine ranges from 610 KW·H up to 6689 KW·H.

Surprisingly, Northern Ireland which is generally considered to be blustery reveals a low potential energy output based on 2006 wind data.

The wind energy surface shown in **Figure 28** represents potential energy with an efficiency conversion rate of 20% attributed to a 2 metre diameter turbine in order to calculate actual energy production levels.

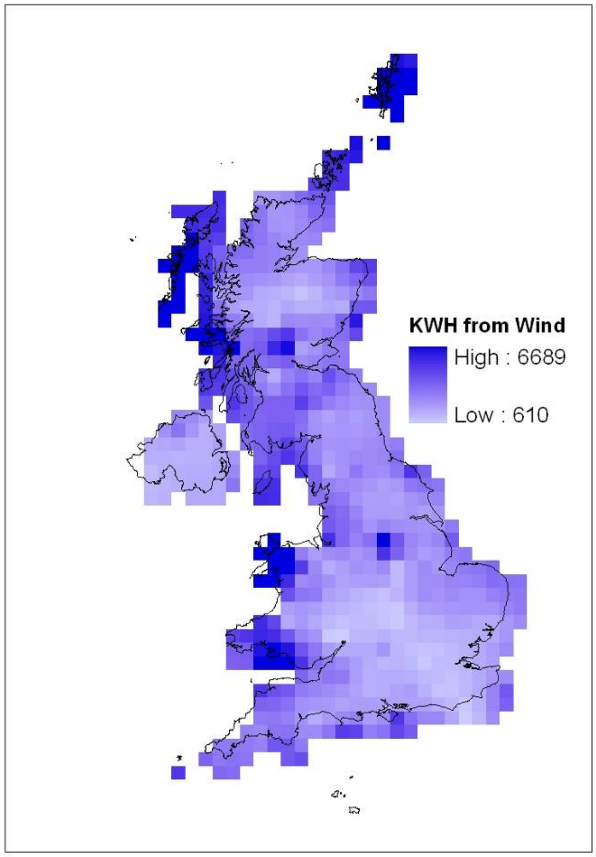


Figure 28: Potential wind energy fluctuates widely across the UK, and is punctuated by pockets of highs and lows.

Looking beyond the x,y dimensions of location, height (known as the z dimension) was found to be a strong variable in determining wind power at many locations. Stations at high elevations consistently produced greater amounts of energy due to higher wind speeds. Stations above 600 meters were removed in an attempt to reduce the influence that high elevation wind recordings would have on the final UK wide output.

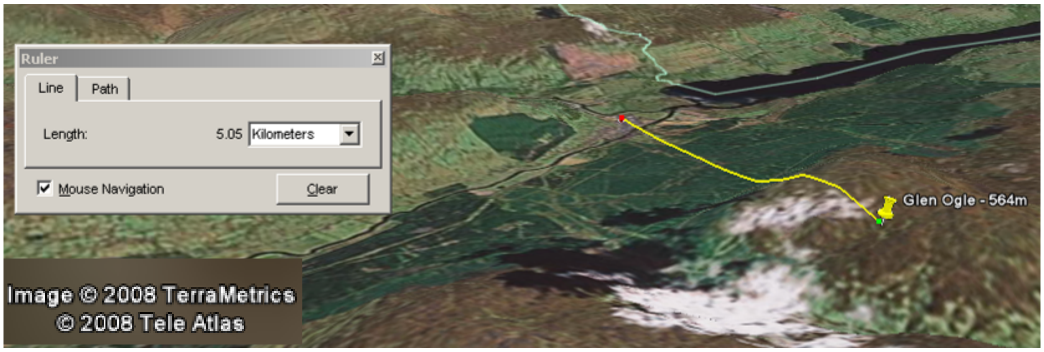
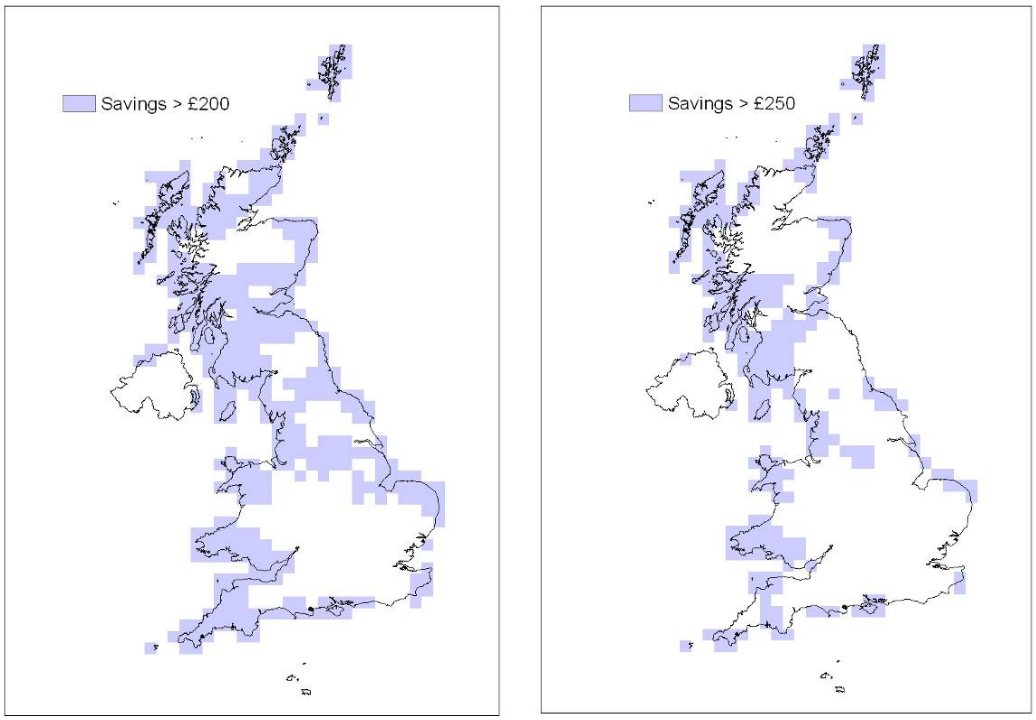


Figure 29: The influence of station elevations, and considering the highest station used Glen Ogle at 564 metres, 5 km from the town of Killin at 102 metres.

Glen Ogle weather station returned an expected wind energy value of 4777 KW·H for a yearly period. The expected energy levels at Killin, 5 km away, when using a Power value of 1 (see section 3.2) for the creation of the raster is 2777 KW·H, while if a Power value of 2 is used in creating the interpolated raster, Killin's expected wind energy is 3917 KW·H. As discussed in the section on point interpolation, a higher Power values will cause a location to be more strongly influenced by closer known locations, and using a Power value of 2 rather than 1 shows this clearly at Killin. Killin has 8 weather stations within 78 km radius, but the closest is Glen Ogle, hence this station plays the greatest influence on Killin's expected energy levels. Killin and Glen Ogle are at two elevations that differ by about 462 meters. This is an example of a weather station probably not being a good representation of the expected climate in the surrounding areas. In these situations it would be more desirable to reduce the influence from stations with irregular values, but overall, giving more weight to closer stations is the correct way to allocate interpolated values in the case of wind speeds.

Figures 30 and **31** show a characteristic of UK wind energy which becomes ever clearer when compared to figures XX and XX, the solar irradiance counter-parts. Approximately half of the locations in the UK could produce £200 worth of savings from wind energy in an average year, but many of those locations will still be capable of producing over £250 worth of savings. Most of these locations are along the UK coast, especially the West, suggesting a pattern where wind storms will tend to blow in from the West before dissipating somewhat overland while crossing East.

Areas not capable of £200 worth of saving tend to be inland, but not too mountainous regions, suggesting that rolling hills will dampen wind speeds. Inland regions that do manage higher energy production from wind are mainly located at regions of higher elevations. High elevations would appear to be the more governing factor in wind energy production than wind speed reducing effects of hills.



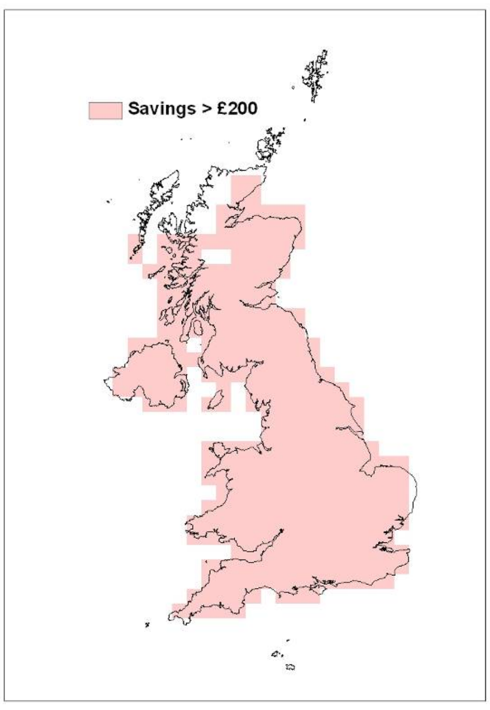
Figures 30 and **31**: Areas where projected energy savings from 2m diameter turbines are greater than £200 and £250.

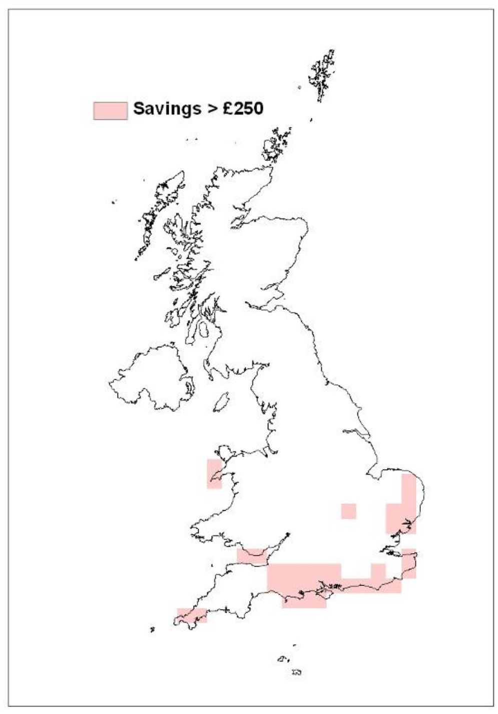
4.2 On Solar Energy Results

Solar radiation is collected by the Met Office in W·H /m^2, a simple conversion to the units which energy providers bill customers by, KW·H. This made the calculation of potential energy savings an easy one. Once relative frequency distribution for the year's worth of station reporting was determined, yearly potential energy was extracted.

Solar irradiation values across the UK are seen to be much more consistent than wind. The lowest expected potential energy for anywhere in the UK in 2006 is 1513 KW·H, while the highest expected value is around 2260 KW·H (based on 15% conversion efficiency, and 2 square metres of panels). This represents a very narrow range of 747 KH·W when compared to the 6079 KW·H range of energy values found from wind in the UK. Solar energy can consequently be viewed as a more predictable means of renewable energy production when compared to wind.

Figures 32 and **33** describe solar irradiance's defining characteristic of consistency across locations. Almost every location in the UK is capable of producing £200 worth of savings from solar energy on a yearly basis. When the UK is analysed for locations that would be expected to produce over £250 worth of savings from solar irradiance, the drop off is potential locations is sudden and dramatic. Only portions of the southern coast of England and one location in western Wales does the expected savings exceed £250.





Figures 32 and 33: Areas where projected energy savings from 2m^2 of solar panels are greater than £200 and £250. The drop off in viable locations is very significant.

As is the case for wind energy, high solar energy regions are generally found along the UK coast. This can be attributed to the fact that clouds tend not to form until weather systems have moved more inland. Higher wind speeds which were found along the coastal regions will also assist in sweeping irradiance blocking clouds inland preventing the clouds from hovering over the shorelines for extended periods of time. Not surprisingly, renewable energy from irradiation is low in Scotland, as mountainous regions tend to form close to the shorelines, and a northerly situation brings short days for most of the months outside of summer.

In **Figure 34**, the 31km resolution solar energy raster, pockets with lower potential energy can be seen running through central Great Britain, while a station in Cornwall reporting low irradiance values for 2006 can be seen to affect the raster's continuity. Greater analysis of regions of low and high anomalies, taking into account geographical land features could reveal mountains diverting clouds in certain areas, and valleys trapping clouds in other.

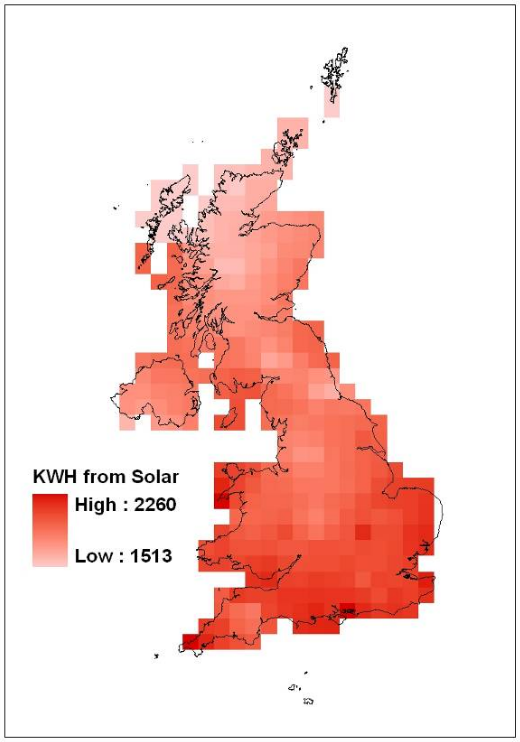


Figure 34: Potential solar energy across the UK is much more constant when compared to potential wind energy. Estimated energy values assume a 2 m^2 solar panel, at 15% conversion efficiency.

An important factor to consider at any location is the set-up of solar panels. A surface tilted towards the South at 35° will intercept around 11% more light energy than a horizontal surface, and angling panels 45° East or West of South will result in a drop of 25% in energy output (Kemp). While the Met Office aims to collect accurate and consistent irradiance levels at a location, it does not aim to collect the maximum levels of irradiance through panel manipulation. Privately owned panels could be configured to track the path of the sun throughout the day, and at the least should be tilted at a suitable angle for a given location. The Met Office's solar panels do not track the sun, and rest horizontally on the earth's surface.

4.3 *On Analysis of Uncertainty*

The abstraction of reality to model will always leave a degree of uncertainty. The relative degree of this uncertainty is important to communicate to users.

Empirical values reported at stations were interpolated to obtain modeled values, but with many stations reporting irregular values with respect to neighbouring stations, interpolation increasingly becomes guess work. By reporting on the irregularities of stations contributing to a location's expected energy output, confidence levels are determined and reported alongside potential energy values.

Throughout the UK, each location will report back unique expected energy levels at confidence levels relating to the attributes of nearby stations. Where a location's expected energy levels is contributed to by 10 stations with a relatively small degree of standard deviation between the stations, the expected energy values can be reported back to the client with a high degree of confidence. While every wind and solar grid cell value will have been determined by at least one weather station, in some cases there are up to 17 stations influencing a grid cell value.

Uncertainty varied greatly between potential wind and solar energy values reported back to users. Uncertainty within reported potential energy from wind resulted from the high degree of variance between wind stations, which makes a strong statement about the unpredictability of wind speeds in any location throughout the UK. Generalizations on expected energies can be made in referring to the coast, or in referring to higher elevation locations, but reporting expected winds across the UK using the network of stations that the Met Office provides is difficult to do with confidence.

In all cases, solar energy received the maximum score of five for variance. Variance of solar stations in a region was almost always under 100 KW·H per year, with a few locations reporting values up to 135. Variance in reported potential wind energy is rarely under 300, and is often over 2000, especially along the UK coastline. Because there are more wind stations than solar, wind energy values will usually be reported back using more stations than solar. This results in wind energy confidence ratings being higher than solar in relation to number of stations used.

Overall, when "station variance" scoring and "number of stations used" scoring are totalled, potential solar energy is consistently reported back with more confidence than wind. This is in line with the overall feeling of values being reported from Met Office stations in 2006, where wind is seen to represent a very localized weather element which varies widely over smaller distances, while solar irradiation has fewer pockets of high and low potential energy values.

Calculations for uncertainty are performed synchronously on the server as the user navigates the geobrowser. This option was chosen rather than the option of building a raster surface of uncertainty values which could have been designed into another database table to be queried in the same manner as the potential energy grid point tables.

4.4 On Google Earth as a Visualisation Tool

Google Earth created the KML formatting, but since August 2007, KML's advancement has been in the hands of the OGC (Open GIS Consortium). This has helped to ease the geocommunity's concerns of a proprietary GIS format taking too large a share of the market. This has also encouraged other digital globe manufacturers to implement support for the KML format. While sometimes being overly verbose, KML's format can also be considered 'easy to use', written in ASCII text format to allow text editor based manipulation. KML essentially allows anyone to augment Google Earth's rich datasets with 'volunteered' geographic information (Goodchild, 2007).

The system designed in this project supports Shneiderman's (1996) oft-quoted "visual information-seeking mantra: overview first, zoom and filter, then details-on-demand". At the overview stage, the user is first required to enter their postcode on the left hand side of the Google Earth interface.

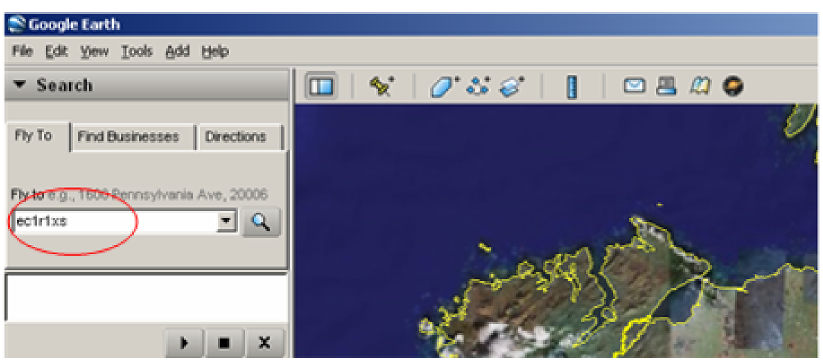


Figure 35: The user enters their postcode into the "Fly To" text box, and hits "return" key.

Google Earth will then navigate to the entered location completing the zoom stage. Filtering takes place once the query has been received from the network link, executed against the database, and returned to the Google Earth interface in the form of a KML string. In the case of this project, the filtering is of weather station data, where points which are not of significance to the user will not be considered for querying.

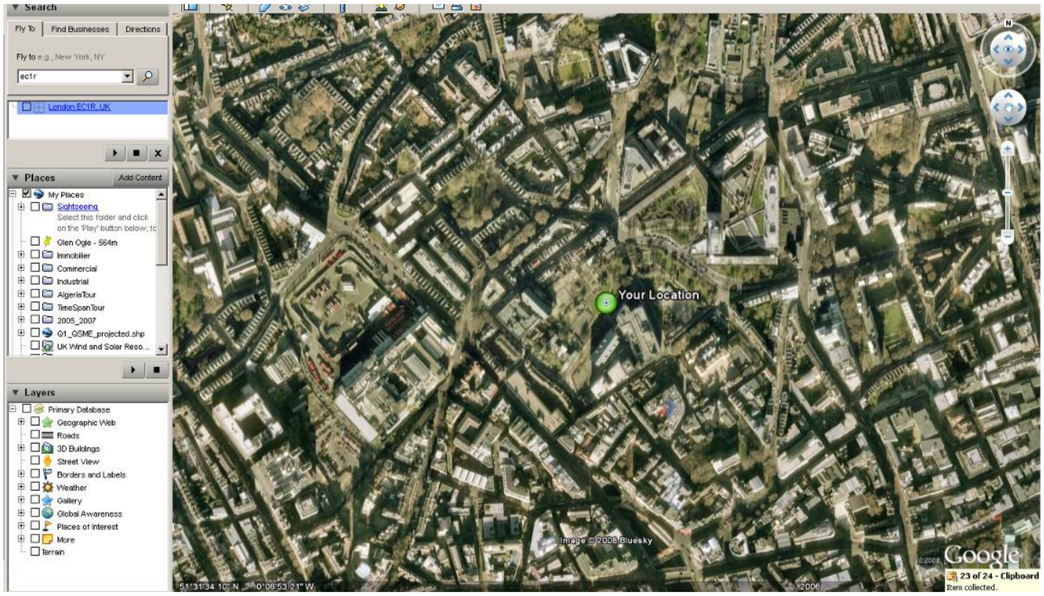


Figure 36: The green dot icon which is displayed at the centre of the user's view, representing the user's location.

Once Google Earth has parsed the KML, a placemark is pinned to the centre of the map's view. This placemark holds the details, which are available on-demand should the user choose to view them through clicking on the placemark.

Figure 37: Balloon displaying estimated savings in Islington, East London.

In order to give the end user a visualisation of the stations used in interpolating a value for their location, the resulting Google Earth KML also displays wind stations and solar stations that contributed to the location's potential energy.

Figure 38: All the wind and solar stations used to determine the potential energy at the location in Islington, the green spot at the centre of the map.

Google Earth's virtual globe is freely available, and can be download and used by anyone with a computer and an internet connection. This accessibility makes it a solid choice as an interface to provide the details of the research undertaken in this project.

The goal was to provide localised potential energy levels to a group of people covering a broad, non-local geographic region. Google Earth's smooth navigability allows users to quickly find their local information, rendering a global dataset quickly relevant at a very localised scale. Slingsby describes how Google Earth's intuitive user interface and the streaming technology it uses to collect data from remote servers allows spatially-referenced results to be disseminated and presented to wide groups of users (Slingsby, 2007). This describes the fundamental qualities of Google Earth which are harnessed through this project.

4.5 *On Possible Sources of Error*

4.5.1 Missed Energy Collection

With regards to solar energy, a factor not considered in this study is the summer months when energy being gathered and stored might have exceed that which is needed to accomplish a desired task (such as heat water to a maximum temp). In this case, the system shuts off, and extra solar energy will not be converted. Proper storage of energy when irradiance is in abundance is essential to ensure constant supply of energy during the dark winter months. Wind energy, while more abundant in the winter months, is not as dramatically season dependent as solar energy is. However, it is still important to have ample storage systems to ensure windy high-energy production days are taken advantage of.

4.5.2 Missing Reports from Stations

Missing data from the weather stations could lead to inaccurate yearly energy potential results for a location. Stations must have recorded 8000 reports throughout the year, so up to 760 hours of data is missing. When did these missing hours occur? Are they all in sequence? Are they spread across different periods in the year? Cawood weather station recorded 7968 readings for solar irradiance in 2006, so did not make the cut for acceptable weather stations. **Figure 39** shows the gaps in recording, most notably in June and November/December. The lack of results in these months would influence the frequency distribution of irradiance levels received at Cawood over the year.

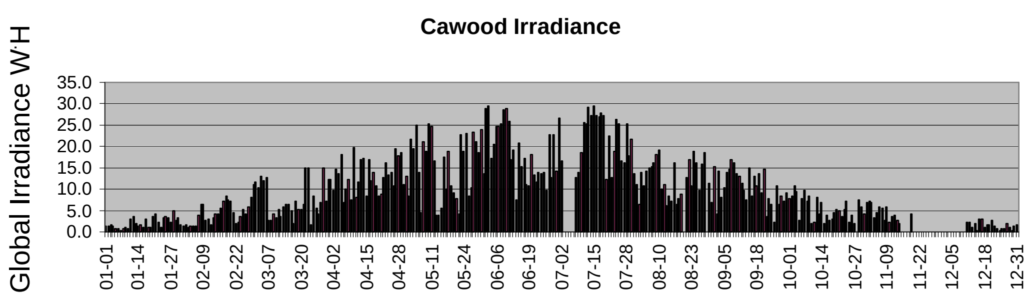


Figure 39: Global irradiance recorded in Cawood throughout 2006.

4.5.3 Affects of Abstraction in Modeling

The idea of localisation as a relative term is very relevant to this paper. Localisation covers a very broad spectrum, from completely non-localised, as wind and solar energy ratings tend to be when they pertain to no specific region, to the highest level of localisation, where a location's wind tunnelling, or shade producing characteristics are considered. The potential energy values produced in this study are taken directly from Met Office stations which are set up in ideal conditions, with minimal ground features to affect wind speed results and minimal shading to detract from solar irradiance levels. It is assumed that a user's location contains the same environmental conditions as its nearest wind and solar stations, but this of course will not always be the case.

In attempting to model accurately as possible, it is important to consider the key parameters affecting the model. Parameterization is a function of the scales at which a model operates (Harvey, 1985). Considering parameters which would increase the precision in wind and solar results to a house by house scale is a valid argument, but beyond the scope of this project. The project's resulting tool must be used with this limitation of parameters in mind.

4.5.4 Insufficient Time-Span for Base Data

Recordings have only been taken from 2006. Can the recorded wind and solar energy values for 2006 be considered a good representation of an areas mean weather patterns? The answer would undoubtedly vary by location. It might it have been more accurate to take results from numerous years and mean them, but large datasets made working with the data time consuming and difficult. In total, there are 1.37 million records for hourly irradiance reports in 2006, while in the same year there are 1.59 million records for hourly wind reports.

Random error will occur in these results throughout the UK, many locations will report back overestimated energy values when compared to the year to year average, while it is likely that equally as many locations will report lower than average expected energy results. Interpolation has served to minimize the presence of these randomly occurring errors.

4.5.5 Ever Changing Climate

A changing environment must also be considered when looking at future patterns of weather data. While a location which is currently sunny is unlikely to all of a sudden become sunny, there is the distinct possibility of parts of the UK which have experienced recording breaking sun and temperatures in recent years can continue to see this trend occur.

Much of the public's interest in wind and solar energy comes down to one value: pound sterling per KW·H. In other words, considering how much is spent on the system installation and setup, how many years are required before everything pays itself off. This again is a hard value to estimate, and depends strongly on the price fluctuations of gas and electricity in the future. This project chose not to speculate on the future prices of energy, and so chose not to speculate on a break-even point in the future for either of the two renewable energy systems. Price per KW·H is set at 12p, but this value could always be updated with the code to reflect changes.

4.5.6 Functional Errors

A functional error in the Google Earth system makes it necessary for the user to maintain a plan view with Google Earth, keeping the tilt at a 90° angle. The model relies on Google Earth's inherent capability to pass the bounding box coordinates of the user's view through to server side code via the Network Link KML tag. Once the view has become tilted, the North and South bounding coordinates (the Latitudes) become distorted. Entering a postcode value, and allowing Google Earth to fly into this location, the calculated bounding box will not render the centre of view and the postcode as being the same place if the view is tilted away from 90°.

5 Conclusions and Recommendations

An Observer newspaper article "Are wind turbines too feeble", reported on wind energy research from 15 sites in rural, suburban and urban locations. Only three of the sites generated more than 400 W·H per day (Jowit, 2008). Yearly energy savings at production rates of 400 W·H per day works out to about £18 per year.[10]

This project's research, done with calculations attributing turbines a 20% efficiency rate, estimates the lowest outputting 2m diameter turbine from the 145 wind stations to have a mean energy production of 1000 W·H per day over a year, with the average turbine at the 145 stations producing a mean of around 6000 W·H per day, to a savings of almost £200 per year. Discrepancies of values such as those reported by the Observer, and those calculated within this and other research projects make clear the necessity for further research and better overall understanding of the variables involved in considering renewable energy potential.

[10] Assuming a price of 12 pence/KW·H

While pessimistic newspaper articles report one way, major errors also occur in the other direction. Manufacturer's wind power estimations result from overestimating turbine efficiency, using mean wind speeds that will never be achieved in most areas, and also attempting to compensate for lack of wind gust data by using higher than normally achievable wind speeds in calculations.

The results of this research have proven that coastal areas, especially those in the south and the east stand to benefit the most from harnessing both solar and wind energy. The flat ground surface provided by the sea allows for wind speeds to gain momentum unimpeded, and also is not conducive to cloud build up. Undulating hills like those found inland in the UK, act to slow down winds in these areas. A rolling hills type of topology will see the sun shine more brightly than in seriously mountainous regions where cloud formation is dense and frequent, such as Scotland. These mountainous regions will also tend to experience high wind speeds (though not as high as the coastal regions) due to their higher elevations, and wind's tendency to increase with increasing elevation.

Most areas that benefit from one renewable energy system will also likely benefit from the other. In particular, Western Wales and Cornwall see high energy productions from both systems. Londoners would be wise to avoid wind turbines, as wind speeds would probably be even lower than reported from wind stations due to the dampening affect that houses and other undulating ground features would have on wind speeds.

Further research could investigate max hourly gusts, and determine how many gusts occur at this speed, or at speeds in between the max gust and mean wind speeds. More detailed wind gusts data needs collected by the Met Office would have benefited this research. At an increased level of precision, consideration on how quickly turbines react to wind gusts in order to take full potential of these sudden and short-lived burst of wind is also important.

With such varying landscape in the United Kingdom, the location variable cannot be ignored when considering natural energy production. The estimated energy outputs for a wind turbine on the shores of Western Scotland are obviously not going to be applicable to a turbine on a roof in suburban London. Indeed, solar panel customers are warned to de-rate manufacturer's power ratings by 20-40%, based on local atmospheric conditions (Kemp, p240).

In order to get the highest level of precision measurements for wind and solar localisation, scientific meteorological measuring instruments would need to be installed on individual rooftops throughout the UK. The finances and the logistics are obviously not there for the government to undertake an investigation of this magnitude; however, the idea of user contributed data does come into the scope of a future project here. Meteorological recording devices paid for and set up by amateur meteorologists at individual residences, with data being fed back to Met Office database via web services or RSS feeds, opens up the possibility of a much broader network of weather stations. Through this practice, the effects of neighbouring building structures, nearby trees, and other impediments to wind and solar energies can be modelled at a much larger scale. Osborn mentions modelling as a means to address environmental conditions where direct experimentation would be morally unacceptable (Osborn 1994). It is doubtful that the Met Office could morally justify the finances to directly experiment with wind and solar conditions at a resolution as precise as say, postcode unit. Instead, one works with what is available, and models to the best precision that can be measured.

Solar panels will also be most productive along the coastlines and not surprising, generally more productive in the South than the Northern regions. Confidence levels in these statements which are interpreted and reported back to the user, reveal that expected solar energy levels can be reported back to users with much greater certainty than wind energy values. This is despite the fact that there are far fewer solar stations that wind, and owes more to the fact that results reported back from solar weather stations are more consistent with much lower levels of standard deviation than those from wind stations.

In terms of difficulties in modeling, wind clearly presents a variable in much greater flux than solar irradiance. This makes wind energy a much more controversial topic than solar energy and also more complicated. Solar irradiation values change very slowly in relation to location, and station density would seem sufficient to cover regional solar variations. Estimating expected wind energy from a yearly mean wind speed is fraught with room for error, often leading to disappointment when returned energy levels are much lower than expected.

Research into wind and its vulnerability to influences from localized features requires more detailed data from Met Office weather stations. More realistic ground feature datasets which take into consideration detailed features such as foliage and buildings would also assist in more precise wind measurement taking. Investigations into three dimensional models which describe man made features would be a good place to start. Discounting the effects of geological features, such as valleys, mountains, buildings and foliage can be forgiven within a project of this resolution scale given the complexities that these features add to calculations. In order to increase model precision to the next level, these variables could not continue to be ignored.

6 Bibliography

Barnsley, Michael J, 2007. Environmental Modeling, a Practical Introduction. (London : CRC Press)

Bowden, G.J., Barker, P.R., Shestopal, V.O., and Twidell J.W., 1983. The Weibull Distribution and Wind Power Statistics. Wind Energy, 7, pp 85-98.

Betz, A., 1926. Windenergie und Ihre Ausnutzung durch Windmullen (Gottingen, Germany : Vandenhoeck and Ruprecht)

Fisher, P, 2005. Session 10: The representation of Uncertain Geographic Information. *Geographic Information Sciences Lecture Notes, City University*. pp 12

Gastner, Michael T., Newman, M.E.J., 2004. Diffusion-based method for producing density-equalizing maps. PNAS. May 18, 2004, vol. 101 no. 20 7499-7504.

Gholam, Riahy H., 2007. A New Method for Calculating the Power Output of a Wind Turbine based on Wind Speed Variations and the EPF. (Mashad, Iran)

Gipe, Paul. 1999. Wind Energy Basics: A Guide to Small and Micro Wind Systems. (United States : Chelsea Green Publishing Company)

Goodchild, M.F. 2007. “Citizens as sensors: the world of volunteered geography”. GeoJournal 69, pp 211-22.

Harvey, L.D.D. and Schneider, S.H., 1985. Transient climatic response to external forcing on 10^0 - 10^4 year time scales, 1, Experiments with globally averaged coupled atmosphere and ocean energy balance models, Journal of Geophysical Research 90, 2191-2205.

Jowit, Juliette, 2008. Home wind turbines 'too feeble'. The Weekend Observer, 06/01/08, page 21. [Accessed January 2008] <***http://www.guardian.co.uk/environment/2008/jan/06/windpower.alternativeenergy***>

Kemp, William H., 2005. The Renewable Energy Handbook. (Tamworth, Ontario : Aztext Press)

Longley, P., Goodchild, M., Maguire, D., Rhind, D., 2005. Geographic Information Systems and Science (Chichester, England : John Wiley & Sons Ltd)

Monmonier, M., 1996. How to Lie with Maps, Second Edition. (Chicago, USA : Chicago Press)

Osborn, T.J. and Wigley, T.M.L., 1994. A simple model for estimating methane concentrations and lifetime variations, Climate Dynamics 9, 181-193

Rowlatt, J., 2006. Are domestic turbines an eco-con? [Accessed August 2007]. <***http://www.bbc.co.uk/blogs/newsnight/2006/11/are_domestic_wind_turbines_an_ecocon.html***>

Shi, W. 1998. A generic statistical approach for modeling error of geometric features in GIS. *International Journal of Geographical Information Science*, vol 12 (2), pp. 131-144.

Shneiderman, B., 1996. "The eyes have it: A task by data type taxonomy for information visualizations", Visual languages, Boulder; CO, IEEE Computer Society Press, 1996.

Slingsby, A., Dykes, J., Wood, J., Foote, M., Blom, M., 2008. "The Visual Exploration of Insurance Data in Google Earth". Lambrick D (ed) Proceedings of GIS Research UK 2008, Manchester. pp 24-32.

Tobler, W. R., 1970. "A computer movie simulating urban growth in the Detroit region". Economic Geography, 46(2): 234-240.

Twidell, J.W. and Weir, A.D., 2006, Renewable Energy Resources, Second Edition (Cheshire, CT: Graphic Press)

UK Met Office. Abridged MIDAS Handbook, Maintained by Desktop and Databases branch. [Accessed August 2007] <***http://badc.nerc.ac.uk/data/surface/ukmo_guide.html***>

UK Wainwright, J., Mulligan M., 2004. Environmental Modelling, Finding Simplicity in Complexity. (London, England : John Wiley & Sons)

Wilson, B., 2003, UK Energy White Paper Speech. Tuesday, May 13, 2003. Power 2003 Conference, 7th European Power and Gas Conference, Madrid. [Accessed May 2008] <***http://www.berr.gov.uk/ministers/speeches/wilson130503.html***>

7 Acknowledgements

Special thanks to the Met Office for allowing the use of their data, and to the British Atmospheric Data Centre for acting as the main point of contact to access this data. The BADC was particularly helpful in providing quick responses to any queries.

8 Appendices

Appendix A – PHP code

createkmlSimple.php

```php
<?php

/*
 * createkmlSimple.php
 * Created on 28 Nov 2007
 * Author: Michael Blom
 * Code written with the assistance of OS data sheets
 */

//database information
require('phpsql_dbinfo.php');
require('projectionFunctions.php');
require('functions.php');
require('windCalcs.php');
require('solarCalcs.php');

//Opens a connection to a MySQL server.
$connection = mysql_connect ($server, $username, $password);
if (!$connection)
{
	die('Not connected : ' . mysql_error());
}

//Sets the active MySQL database.
$db_selected = mysql_select_db($database, $connection);
if (!$db_selected)
{
	die('Can\'t use db : ' . mysql_error());
}
```

```
$coords = explode(',', $_GET['BBOX']);

// calculate the approx center of the view
// note that this is innaccurate if the view is tilted

$userlon = (($coords[2] - $coords[0])/2) + $coords[0];
$userlat = (($coords[3] - $coords[1])/2) + $coords[1];

      //Determine the height above the ellipsoid at the user's entered location
      $queryDEM_ELEV = 'SELECT ELEV FROM (SELECT ELEV FROM dempointsuk WHERE
Latitude >= '.$userlat.' and Longitude >='.$userlon.' order by Latitude,
Longitude) s limit 1';
      $result_DEM = mysql_query($queryDEM_ELEV);
      $rowELEV = @mysql_fetch_assoc($result_DEM);
      $elevation = $rowELEV['ELEV'];

      //echo "elevation: ".$elevation."\n";

      //Perform the transformation of the coords from WGS84 to OSGB grid
      $arr = transform($userlat, $userlon, $elevation);

      $OSGB_X=$arr["X"];
      $OSGB_Y=$arr["Y"];

      $windResults=performWindCalcs($OSGB_X,$OSGB_Y);
      $confidenceRatingWind = rateConfidence($windResults);

      $solarResults=performSolarCalcs($OSGB_X,$OSGB_Y);
      $confidenceRatingSolar = rateConfidence($solarResults);

      //0.12 = the price of energy
      $windSavings = 0.12*(round($windResults["Energy"],-2));
      $solarSavings = 0.12*(round($solarResults["Energy"],-2));

      $totalSavings = round($windSavings,-1)+round($solarSavings,-1);
      // Creates the Document.
      $dom = new DOMDocument('1.0', 'UTF-8');

      // Creates the root KML element and appends it to the root document.
      $node = $dom->createElementNS('http://earth.google.com/kml/2.2', 'kml');
      $parNode = $dom->appendChild($node);

      // Creates a KML Document element and append it to the KML element.
      $dnode = $dom->createElement('Document');
      $docNode = $parNode->appendChild($dnode);

      //insert Icon style for wind station
      $windStyleNode = $dom->createElement('Style');
      $windStyleNode->setAttribute('id', 'windStyle');
      $windIconstyleNode = $dom->createElement('IconStyle');
      $windIconstyleNode->setAttribute('id', 'windIcon');
      $windIconNode = $dom->createElement('Icon');
      $windHref = $dom->createElement('href',
'http://vega.soi.city.ac.uk/~abbj500/renewables/windIcon.jpg');
      $windIconNode->appendChild($windHref);
      $windIconstyleNode->appendChild($windIconNode);
      $windStyleNode->appendChild($windIconstyleNode);
      $docNode->appendChild($windStyleNode);
```

```
	//insert Icon style for solar station
	$solarStyleNode = $dom->createElement('Style');
	$solarStyleNode->setAttribute('id', 'solarStyle');
	$solarIconstyleNode = $dom->createElement('IconStyle');
	$solarIconstyleNode->setAttribute('id', 'solarIcon');
	$solarIconNode = $dom->createElement('Icon');
	$solarHref = $dom->createElement('href',
'http://vega.soi.city.ac.uk/~abbj500/renewables/solarIcon.jpg');
	$solarIconNode->appendChild($solarHref);
	$solarIconstyleNode->appendChild($solarIconNode);
	$solarStyleNode->appendChild($solarIconstyleNode);
	$docNode->appendChild($solarStyleNode);

	//insert Icon style for user's location
	$locStyleNode = $dom->createElement('Style');
	$locStyleNode->setAttribute('id', 'userlocStyle');
	$locIconstyleNode = $dom->createElement('IconStyle');
	$locIconstyleNode->setAttribute('id', 'uerlocIcon');
	$locIconNode = $dom->createElement('Icon');
	$locHref = $dom->createElement('href',
'http://vega.soi.city.ac.uk/~abbj500/renewables/icon17.png');
	$locIconNode->appendChild($locHref);
	$locIconstyleNode->appendChild($locIconNode);
	$locStyleNode->appendChild($locIconstyleNode);
	$docNode->appendChild($locStyleNode);

	$windArray = $windResults["StationObjArray"];
	foreach($windArray as $value) {
		$node = $dom->createElement('Placemark');
		$placeNode = $docNode->appendChild($node);

		// Creates an id attribute and assign it the value of id column.
		$placeNode->setAttribute('id', $value["src_id"]);

		// Create name, and description elements and assigns them the values
of the name and address columns from the results.
		$nameNode = $dom->createElement('name','Wind:
'.$value["stationname"]);
		$placeNode->appendChild($nameNode);
		$descNode = $dom->createElement('description','Potential Energy /yr:
'.$value["energyatstation"].' KWH');
		$placeNode->appendChild($descNode);
		//styling
		$styleUrl = $dom->createElement('styleUrl', '#windStyle');
		$placeNode->appendChild($styleUrl);
		// Creates a Point element.
		$pointNode = $dom->createElement('Point');
		$placeNode->appendChild($pointNode);
		// Creates a coordinates element and gives it the value of the lng
and lat columns from the results.
		$coorStr = $value["LON"].','.$value["LAT"];
		$coorNode = $dom->createElement('coordinates', $coorStr);
		$pointNode->appendChild($coorNode);
	}

	$solarArray = $solarResults["StationObjArray"];
	foreach($solarArray as $value) {
		$node = $dom->createElement('Placemark');
		$placeNode = $docNode->appendChild($node);
```

```
            // Creates an id attribute and assign it the value of id column.
            $placeNode->setAttribute('id', $value["src_id"]);

            // Create name, and description elements and assigns them the values
of the name and address columns from the results.
            $nameNode = $dom->createElement('name','Solar:
'.$value["stationname"]);
            $placeNode->appendChild($nameNode);
            $descNode = $dom->createElement('description','Potential Energy /yr
: '.$value["energyatstation"].' KWH');
            $placeNode->appendChild($descNode);
            //styling
            $styleUrl = $dom->createElement('styleUrl', '#solarStyle');
            $placeNode->appendChild($styleUrl);
            // Creates a Point element.
            $pointNode = $dom->createElement('Point');
            $placeNode->appendChild($pointNode);
            // Creates a coordinates element and gives it the value of the lng
and lat columns from the results.
            $coorStr = $value["LON"].','.$value["LAT"];
            $coorNode = $dom->createElement('coordinates', $coorStr);
            $pointNode->appendChild($coorNode);
      }

      //Creates a Placemark and append it to the Document.
      $node = $dom->createElement('Placemark');
      $placeNode = $docNode->appendChild($node);

      // Creates an id attribute and assign it the value of id column.
      $placeNode->setAttribute('id', 'UserLocation');

      // Create name, and description elements and assigns them the values of
the name and address columns from the results.
      $nameNode = $dom->createElement('name','Your Location');
      $placeNode->appendChild($nameNode);
      $descNode = $dom->createElement('description', '
            <table width = "300">
          <tr><td width = "250"><h2><b>Possible yearly energy savings at your
location:</b></h2></td></tr>
          <tr><td width = "250"><p>Savings from a 2m diameter turbine:
£'.round($windSavings,-1).' / year</p></td></tr>
          <tr><td width = "250"><p>Wind reading confidence level:
'.$confidenceRatingWind["Confidence"].'  (1-10)</p></td></tr>
          <tr><td width = "250"><img width=115 height=120
src="http://vega.soi.city.ac.uk/~abbj500/renewables/turbine.jpg"></td></tr>
          <tr><td width = "250"><p>Savings from a 2sqm solar panel:
£'.round($solarSavings,-1).' / year</p></td></tr>
          <tr><td width = "250"><p>Solar reading confidence level:
'.$confidenceRatingSolar["Confidence"].' (1-10)</p></td></tr>
          <tr><td width = "250"><img width=115 height=120
src="http://vega.soi.city.ac.uk/~abbj500/renewables/solar.jpg"></td></tr>
          <tr><td width = "250"><p><b><font color="red"> Total <i>potential</i>
savings = £'.$totalSavings.' / year</font></b></p></td></tr>
              <tr><td width = "250"><p><b> For more information, read <a href
="http://www.student.city.ac.uk/~abbj500/dissertation/paper5.pdf">the paper</a>
behind the project.</b></p></td></tr>
        </table>');
      $placeNode->appendChild($descNode);
      //styling
      $styleUrl = $dom->createElement('styleUrl', '#userlocStyle');
```

```
        $placeNode->appendChild($styleUrl);
        // Creates a Point element.
        $pointNode = $dom->createElement('Point');
        $placeNode->appendChild($pointNode);
        // Creates a coordinates element and gives it the value of the lng and lat
columns from the results.
        $coorStr = $userlon.','.$userlat;
        $coorNode = $dom->createElement('coordinates', $coorStr);
        $pointNode->appendChild($coorNode);

        $kmlOutput = $dom->saveXML();
        header('Content-type: application/vnd.google-earth.kml+xml');

        echo $kmlOutput;

/*
 * end of createkmlSimple.php
 */

?>
```

constants.php

```
<?php
        /*
         * constants.php
         * Created on 28 Nov 2007
         * Author: Michael Blom
         */

         //constants for WGS84 ellipsoid
         $Aaxis = 6378137.000;
         $Baxis = 6356752.3142;

         //constants for Airy ellipsoid
         $AaxisAiry =6377563.396;
         $BaxisAiry =6356256.91;

         $PHI0_Airy=49;
         $LAM0_Airy=-2;

         $e0_Airy = 400000;
         $f0_Airy = 0.9996012717;
         $n0_Airy = -100000;
         $pi = 3.14159265358979;

         // translation constants
         $DX= -446.448;
         $DY = 125.157;
         $DZ = -542.06;
         $s = 20.4894;

         $X_Rot=-0.1502;
         $Y_Rot=-0.2470;
         $Z_Rot=-0.8421;

         $a_modAiry = 6377340.189;      // OSGBI semi-major
         $b_modAiry = 6356034.447;      // OSGBI semi-minor
         $e0_modAiry = 200000;          // easting of false origin
         $n0_modAiry = 250000;          // northing of false origin
```

```
        $f0_modAiry = 1.000035;       // OSGBI scale factor on central meridian
        $lam0_modAiry = -8.0;  // OSGBI false east
        $phi0_modAiry = 53.5;   // OSGBI false north
/*
        * end of constants.php
        *
        */
?>
```

functions.php

```
<?php
/*
 * functions.php
 * Created on 10 May 2008
 * contains Pythagoris distance calcs and confidence rating structure
 */

 //calculate the distance between two points
 function distanceCalcs($originX, $originY, $destinationX, $destinationY)
 {
      $distance= sqrt(pow($destinationY-$originY,2)+pow($destinationX-
$originX,2));

      return $distance;
 }

 //get a confidence score on the results
 function rateConfidence($weatherResults)
 {
      $NumberStationsUsed = $weatherResults["NumberStationsUsed"];
      $averageEnergy = $weatherResults["SumValues"]/$NumberStationsUsed;
      $sumDifs = 0;

    for ($i = 0; $i < $NumberStationsUsed; ++$i)
    {
        $sumDifs = $sumDifs + pow($weatherResults["StationEnergyArray"][$i] -
$averageEnergy,2);
    }

      $standardDeviation = sqrt($sumDifs/$NumberStationsUsed);
      if ($NumberStationsUsed <=2)
      {
            $confidenceRating =1;
      }
      elseif ($NumberStationsUsed<=5)
      {
            $confidenceRating =2;
      }
      elseif ($NumberStationsUsed<=8)
      {
            $confidenceRating =3;
      }
      elseif ($NumberStationsUsed<=11)
      {
            $confidenceRating =4;
      }
      else
      {
            $confidenceRating =5;
```

```php
        }

        if($standardDeviation<1)
        {
              $confidenceRating += 1;
        }
        elseif ($standardDeviation <=300)
        {
              $confidenceRating += 5;
        }
        elseif ($standardDeviation <=800)
        {
              $confidenceRating += 4;
        }
        elseif ($standardDeviation <=1500)
        {
              $confidenceRating += 3;
        }
        elseif ($standardDeviation <=2200)
        {
              $confidenceRating += 2;
        }
        else
        {
              $confidenceRating += 1;
        }
        $arr = array("Confidence"=>$confidenceRating, "SD"=>$standardDeviation);
        return $arr;
 }

/*
 * end of functions.php
 */
?>
```

dbinfo.php

```php
<?php
$username="root";//root
$password="geomatics";//geomatics
$database="weatherdata";
$server="localhost";//localhost
?>
```

projectionFunctions.php

```php
<?php
/*
 * projectionFunctions.php
 * Created on 28 Nov 2007
 * Author: Michael Blom
 * Code written with the assistance of OS data sheets
 */

 require('constants.php');

 function transform($lat, $long, $height)
 {
      $arr = transformHelmerts($lat, $long, $height);
```

```
        return $arr;
    }

    function transformHelmerts($lat, $long, $height)
    {
        global $a_modAiry, $b_modAiry, $e0_modAiry,
        $n0_modAiry, $f0_modAiry, $e2_modAiry,
        $lam0_modAiry, $phi0_modAiry,$AaxisAiry,
        $BaxisAiry, $e0_Airy, $f0_Airy,
        $n0_Airy, $PHI0_Airy, $LAM0_Airy, $GRID;

        $Aaxis = $AaxisAiry;
        $Baxis = $BaxisAiry;
        $PHI0 = $PHI0_Airy;
        $LAM0 = $LAM0_Airy;
        $e0 = $e0_Airy;
        $f0 = $f0_Airy;
        $n0 = $n0_Airy;
        $GRID = 'OS';

        //x,y,z caresian values on the WGS84 ellipsoid
        $X_Value = latlongheight_to_x($lat, $long, $height);
        $Y_Value = latlongheight_to_y($lat, $long, $height);
        $Z_Value = latheight_to_z($lat, $height);

        $Helmert_X = helmertConversionX($X_Value, $Y_Value, $Z_Value);
        $Helmert_Y = helmertConversionY($X_Value, $Y_Value, $Z_Value);
        $Helmert_Z = helmertConversionZ($X_Value, $Y_Value, $Z_Value);

        $OSGB_Lat = XYZ_to_Lat($Helmert_X, $Helmert_Y, $Helmert_Z, $Aaxis,
$Baxis);
        $OSGB_Long = XYZ_to_Long($Helmert_X, $Helmert_Y);

        $OSGB_X = Lat_Long_to_East($OSGB_Lat, $OSGB_Long, $Aaxis, $Baxis, $PHI0,
$LAM0, $e0, $f0);
        $OSGB_Y = Lat_Long_to_North($OSGB_Lat, $OSGB_Long, $Aaxis, $Baxis, $PHI0,
$LAM0,$f0, $n0);

        $arr = array("X"=>$OSGB_X, "Y"=>$OSGB_Y);

        return $arr;
    }

    function transformLong($lat, $long, $height)
    {
        $X_Value = latlongheight_to_x($lat, $long, $height);
        $Y_Value = latlongheight_to_y($lat, $long, $height);
        $Z_Value = latheight_to_z($lat, $height);

        $Helmert_X = helmertConversionX($X_Value, $Y_Value, $Z_Value);
        $Helmert_Y = helmertConversionY($X_Value, $Y_Value, $Z_Value);
        $Helmert_Z = helmertConversionZ($X_Value, $Y_Value, $Z_Value);

        $OSGB_Lat = XYZ_to_Long($Helmert_X, $Helmert_Y, $Helmert_Z);
    }

    //convert from WGS84 lat/long to WGS84 cartesian X
    function latlongheight_to_x($lat, $long, $height)
    {
        global $pi,$Aaxis, $Baxis;
```

```
    //convert lat and long to radians
    $radLat = $lat*($pi/180);
    $radLong = $long*($pi/180);

    //calculate ecentricity (ellipsoid flattening)
    $e2 = ((pow($Aaxis,2)) - (pow($Baxis,2))) / (pow($Aaxis,2));

    //calculate the angle between the normal at the point
    //on the earth surface and the x (major) axis
    $v = $Aaxis/(sqrt(1-($e2*(pow((sin($radLat)),2)))));

    $X_Value = ($v + $height)*(cos($radLat))*(cos($radLong));
    return $X_Value;
}

//convert from WGS84 lat/long to WGS84 cartesian Y
function latlongheight_to_y($PHI, $LAM, $height)
{

    global $pi, $Aaxis, $Baxis;

    //Convert angle measures to radians
   $RadPHI = $PHI * ($pi / 180);
   $RadLAM = $LAM * ($pi / 180);

    //Compute eccentricity squared and nu
   $e2 = (pow($Aaxis,2) - pow($Baxis,2)) / pow($Aaxis,2);
   $V = $Aaxis / (sqrt(1 - ($e2 * pow((sin($RadPHI)),2))));

    //Compute Y
   $Lat_Long_H_to_Y = ($V + $height) * (cos($RadPHI)) * (sin($RadLAM));
   return $Lat_Long_H_to_Y;
}

//caculate z value
function latheight_to_z($lat, $height)
{
    global $pi,$Aaxis, $Baxis;

    //convert lat and long to radians
    $radLat = $lat*($pi/180);

    //calculate ecentricity (ellipsoid flattening)
    $e2 = ((pow($Aaxis,2))-(pow($Baxis,2)))/(pow($Aaxis,2));

    //calculate the angle between the normal at the point
    //on the earth surface and the x (major) axis
    $v = $Aaxis / sqrt(1 - ($e2 * pow((sin($radLat)),2)));

    $Z_Value = (($v*(1-$e2)) + $height)*(sin($radLat));
    return $Z_Value;
}

//returns the X cartesian value on the OSGB36 datum ellipsoid (Airy)
function helmertConversionX($X, $Y, $Z)
{
    //Input - XYZ cartesian values from WGS84 ellipsoid
    //Convert rotations to radians and ppm scale to a factor
   global $pi, $Y_Rot, $Z_Rot, $DX, $s;
```

```
    $sfactor2 = $s * 0.000001;
    $RadY_Rot = ($Y_Rot / 3600) * ($pi / 180);
    $RadZ_Rot = ($Z_Rot / 3600) * ($pi / 180);

      //Compute transformed Z coord
    $Helmert_X = $X + ($X * $sfactor2) - ($Y * $RadZ_Rot) + ($Z * $RadY_Rot) +
$DX;
    return $Helmert_X ;
 }

 //returns the Y cartesian value on the OSGB36 datum ellipsoid (Airy)
 function helmertConversionY($X, $Y, $Z)
 {
      //Input - XYZ cartesian values from WGS84 ellipsoid
      global $pi, $Z_Rot, $X_Rot, $DY, $s;

      $sfactor = $s * 0.000001;
    $RadX_Rot = ($X_Rot / 3600) * ($pi / 180);
    $RadZ_Rot = ($Z_Rot / 3600) * ($pi / 180);

      //Compute transformed Y coord
    $Helmert_Y = ($X * $RadZ_Rot) + $Y + ($Y * $sfactor) - ($Z * $RadX_Rot) +
$DY;
    return $Helmert_Y;
 }

 //returns the Z cartesian value on the OSGB36 datum ellipsoid (Airy)
 function helmertConversionZ($X, $Y, $Z)
 {
      //Input - XYZ cartesian values from WGS84 ellipsoid
      global $pi, $Y_Rot, $X_Rot, $DZ, $s;
      //Convert rotations to radians and ppm scale to a factor

    $sfactor = $s * 0.000001;
    $RadX_Rot = ($X_Rot / 3600) * ($pi / 180);
    $RadY_Rot = ($Y_Rot / 3600) * ($pi / 180);

      //Compute transformed Z coord
    $Helmert_Z = (-1 * $X * $RadY_Rot) + ($Y * $RadX_Rot) + $Z + ($Z * $sfactor)
+ $DZ;
      return $Helmert_Z ;
 }

 function XYZ_to_Lat($Helmert_X, $Helmert_Y, $Helmert_Z, $Aaxis, $Baxis)
 {
      //Input: - XYZ cartesian coords (X,Y,Z)
      global $pi;
      $RootXYSqr = sqrt(pow($Helmert_X,2) + pow($Helmert_Y,2));
    $e2 = (pow($Aaxis,2) - pow($Baxis,2)) / pow($Aaxis,2);
    $lat1 = atan($Helmert_Z / ($RootXYSqr * (1 - $e2)));
    $iterated_Lat = Iterate_XYZ_to_Lat($e2, $lat1, $Helmert_Z, $RootXYSqr,
$Aaxis, $Baxis);
    $OSGB_Lat = $iterated_Lat * (180 / $pi);
    return $OSGB_Lat;
 }

 function Iterate_XYZ_to_Lat($e2, $lat1, $Helmert_Z, $RootXYSqr,$Aaxis, $Baxis)
 {
      //Iteratively computes Latitude
```

```
      global $pi;
    $V = $Aaxis / (sqrt(1 - ($e2 * (pow((sin($lat1)),2)))));
    $lat2 = atan(($Helmert_Z + ($e2 * $V * (sin($lat1)))) / $RootXYSqr);

    do
    {
        $lat1 = $lat2;
        $V = $Aaxis / (sqrt(1 - ($e2 * (pow((sin($lat1)),2)))));
        $lat2 = atan(($Helmert_Z + ($e2 * $V * (sin($lat1)))) / $RootXYSqr);
    }
    while (abs($lat1 - $lat2) > 0.000000001);

    $iterated_Lat = $lat2;
    return $iterated_Lat;

}

//
function XYZ_to_Long($Helmert_X, $Helmert_Y)
{
      //Convert XYZ to Longitude (LAM) in Dec Degrees.
      //Input: X and Y cartesian coords in meters.
    global $pi;
    $OSGB_Long = (atan($Helmert_Y / $Helmert_X)) * (180 / $pi);
    return $OSGB_Long;
}

//Lat Long to East Conversion
function Lat_Long_to_East($OSGB_Lat, $OSGB_Long, $Aaxis, $Baxis, $PHI0, $LAM0,
$e0, $f0)
{
      //Project Latitude and longitude to Transverse Mercator eastings.
      //Input: Latitude (PHI) and Longitude (LAM) in decimal degrees;

      global $pi;//, $AaxisAiry, $BaxisAiry, $PHI0, $LAM0, $e0, $f0;

      //Convert angle measures to radians
    $RadPHI = $OSGB_Lat * ($pi/180);
    $RadLAM = $OSGB_Long * ($pi/180);
    $RadPHI0 = $PHI0 * ($pi/180);
    $RadLAM0 = $LAM0 * ($pi/180);

    $af0 = $Aaxis * $f0;
    $bf0 = $Baxis * $f0;
    $e2 = (pow($af0,2) - pow($bf0,2)) / pow($af0,2);
    $n = ($af0 - $bf0) / ($af0 + $bf0);
    $nu = $af0 / (sqrt(1 - ($e2 * (pow((sin($RadPHI)),2)))));
    $rho = ($nu * (1 - $e2)) / (1 - ($e2 * pow((sin($RadPHI)), 2)));
    $eta2 = ($nu / $rho) - 1;
    $p = $RadLAM - $RadLAM0;

    $IV = $nu * (cos($RadPHI));
    $V = ($nu / 6) * pow((cos($RadPHI)),3) * (($nu / $rho) - (pow(tan($RadPHI),
2)));
    $VI = ($nu / 120) * pow((cos($RadPHI)),5) * (5 - (18 *
pow((tan($RadPHI)),2)) + pow((tan($RadPHI)),4) + (14 * $eta2) - (58 *
pow((tan($RadPHI)),2) * $eta2));

    $Lat_Long_to_East = $e0 + ($p * $IV) + (pow($p,3) * $V) + (pow($p,5) * $VI);
    return $Lat_Long_to_East;
```

```
 }

 function Lat_Long_to_North($OSGB_Lat, $OSGB_Long, $Aaxis, $Baxis, $PHI0,
$LAM0,$f0, $n0)
 {
      //Project Latitude and longitude to Transverse Mercator northings
      //Input: Latitude (PHI) and Longitude (LAM) in decimal degrees; _

      global $pi;

      //Convert angle measures to radians
    $pi = 3.14159265358979;
    $RadPHI = $OSGB_Lat * ($pi/ 180);
    $RadLAM = $OSGB_Long * ($pi / 180);
    $RadPHI0 = $PHI0 * ($pi/ 180);
    $RadLAM0 = $LAM0 * ($pi/ 180);

    $af0 = $Aaxis * $f0;
    $bf0 = $Baxis * $f0;
    $e2 = (pow($af0,2) - pow($bf0,2)) / pow($af0,2);
    $n = ($af0 - $bf0) / ($af0 + $bf0);
    $nu = $af0 / (sqrt(1 - ($e2 * pow((sin($RadPHI)),2))));
    $rho = ($nu * (1 - $e2)) / (1 - ($e2 * pow((sin($RadPHI)),2)));
    $eta2 = ($nu / $rho) - 1;
    $p = $RadLAM - $RadLAM0;
    $M = Marc($bf0, $n, $RadPHI0, $RadPHI);

    $I = $M + $n0;
    $II = ($nu / 2) * (sin($RadPHI)) * (cos($RadPHI));
    $III = (($nu / 24) * (sin($RadPHI)) * pow((cos($RadPHI)),3)) * (5 -
pow((tan($RadPHI)),2) + (9 * $eta2));
    $IIIA = (($nu / 720) * (sin($RadPHI)) * pow((cos($RadPHI)),5)) * (61 - (58 *
pow((tan($RadPHI)),2)) + pow((tan($RadPHI)),4));

    $Lat_Long_to_North = $I + (pow($p,2) * $II) + (pow($p,4) * $III) +
(pow($p,6) * $IIIA);

    return $Lat_Long_to_North;
}

function Marc($bf0, $n, $PHI0, $PHI)
{
      $var1 =((5 / 4) * pow($n,2));
      $var1_5 = (5 / 4) * pow($n,3);
      $var2 = ($PHI - $PHI0);
      $var3 = (3 * $n) + (3 * pow($n,2));
      $var4 = ((21 / 8) * pow($n,3));
      $var5 = sin($PHI - $PHI0);
      $var6 = cos($PHI + $PHI0);
      $var7 = ((15 / 8) * pow($n,2));
      $var7_5 = (15 / 8) * pow($n,3);
      $var8 = sin(2 * ($PHI - $PHI0));
      $var9 = cos(2 * ($PHI + $PHI0));
      $var10 = (35 / 24) * pow($n,3);
      $var11 = sin(3 * ($PHI - $PHI0));
      $var12 = cos(3 * ($PHI + $PHI0));

    $Marc = $bf0 * (((1 + $n + $var1 + $var1_5) * $var2) - (($var3 + $var4) *
($var5) * ($var6)) + (($var7 + $var7_5) * ($var8) * ($var9)) - (($var10) *
($var11) * ($var12)));
```

```
	return $Marc;
}
/*
 * end of projectionFunctions.php
 */
?>
```

solarCalcs.php

```
<?php
/*
 * solarCalcs.php
 * Created on 27 May 2008
 * Takes care of calculations related to solar irradiation
 */

function performSolarCalcs($OSGB_X, $OSGB_Y)
{
	$queryArray = array();

	//select the four wind raster cells surrounding the users location, and
then choose the closest
	//TopRight
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM solargrid where
OSGBX >= '.$OSGB_X.' and OSGBY >= '.$OSGB_Y.' order by OSGBX asc, OSGBY asc) s
limit 1');
	//BottomRight
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM solargrid where
OSGBX >= '.$OSGB_X.' and OSGBY <= '.$OSGB_Y.' order by OSGBX asc, OSGBY desc) s
limit 1');
	//BottomLeft
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM solargrid where
OSGBX <= '.$OSGB_X.' and OSGBY <= '.$OSGB_Y.' order by OSGBX desc, OSGBY desc) s
limit 1');
	//TopLeft
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM solargrid where
OSGBX <= '.$OSGB_X.' and OSGBY >= '.$OSGB_Y.' order by OSGBX desc, OSGBY asc) s
limit 1');

	$OSGB_X_solarGrid = '';
	$OSGB_Y_solarGrid='';
	$solarEnergy=0;
	$minDistance = 31000;//none of the four grid points will be more than 31km
away for wind
	$resultsArray = array();
	foreach($queryArray as $query)
	{
		$result = mysql_query($query);
		$solarRasterPoint = @mysql_fetch_assoc($result);
		$distance =
distanceCalcs($solarRasterPoint['OSGBX'],$solarRasterPoint['OSGBY'],$OSGB_X,
$OSGB_Y);
		//choose the closest station
		if($distance < $minDistance)
		{
			$minDistance = $distance;
			$solarEnergy = $solarRasterPoint['solar_kwh2'];
```

```
				$OSGB_X_solarGrid = $solarRasterPoint['OSGBX'];
				$OSGB_Y_solarGrid = $solarRasterPoint['OSGBY'];

			}
		}

		//106 km is the search radius for wind station
		$OSGB_BB_North = $OSGB_Y_solarGrid + 106000;
		$OSGB_BB_South = $OSGB_Y_solarGrid - 106000;
		$OSGB_BB_East = $OSGB_X_solarGrid + 106000;
		$OSGB_BB_West = $OSGB_X_solarGrid - 106000;

		//looking at the closest raster cell, what stations were used to determine
it (within 106km)
		$querySolarStations106km = 'SELECT * FROM (SELECT * FROM solarstations
where OSGBX >= '.$OSGB_BB_West.' and OSGBX <= '.$OSGB_BB_East.' and OSGBY
>='.$OSGB_BB_South.' and OSGBY <= '.$OSGB_BB_North.' order by OSGBX, OSGBY) s ';
		$querySolarStations106kmResults = mysql_query($querySolarStations106km);

		$stationEnergyArray = array();
		$stationObjArray = array();
		$sumValues = 0;
		$stationName= '';
		$distance2SolarStation = 0;

		//iterate through stations within 106km bounding box to find how many are
actually
		//within 106km buffer ring and what their sum is (for standard deviation
calcs)
		while($rowSolarStations =
@mysql_fetch_assoc($querySolarStations106kmResults))
		{
			$distance2SolarStation = sqrt(pow($rowSolarStations['OSGBY']-
$OSGB_Y,2)+pow($rowSolarStations['OSGBX']-$OSGB_X,2));

			if($distance2SolarStation < 106000)
			{
				$stationName = $rowSolarStations['src_name'];
				$energyAtStation = $rowSolarStations['solar_KWH2'];
				$LAT = $rowSolarStations['WGS_LAT'];
				$LON = $rowSolarStations['WGS_LON'];
				$SRC_ID = $rowSolarStations['src_id'];
				$obj = array('src_id'=>$SRC_ID, 'LAT' => $LAT, 'LON' => $LON,
'stationname' => $stationName, 'energyatstation' => $energyAtStation);
				//create an array of solar station objects
				array_push($stationObjArray, $obj);
				array_push($stationEnergyArray, $energyAtStation);
				$sumValues += $energyAtStation;
			}
		}
		$len = count($stationEnergyArray);
		$arr = array("StationObjArray"=>$stationObjArray,
"StationEnergyArray"=>$stationEnergyArray, "Station"=>$stationName,
"Energy"=>$solarEnergy, "Distance"=>$distance2SolarStation,
"NumberStationsUsed"=>$len, "SumValues"=>$sumValues);
		return $arr;
}
/*
 * end of solarCalcs.php
 */
```

```
?>
```

windCalcs.php

```
<?php
/*
 * windCalcs.php
 *
 * Created on 25 May 2008
 * Author: Mike Blom
 */
function performWindCalcs($OSGB_X, $OSGB_Y)
{
	$queryArray = array();

	//select the four wind raster cells surrounding the users location, and
then choose the closest
	//TopRight
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM windgrid where OSGBX
>= '.$OSGB_X.' and OSGBY >= '.$OSGB_Y.' order by OSGBX asc, OSGBY asc) s limit
1');
	//BottomRight
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM windgrid where OSGBX
>= '.$OSGB_X.' and OSGBY <= '.$OSGB_Y.' order by OSGBX asc, OSGBY desc) s limit
1');
	//BottomLeft
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM windgrid where OSGBX
<= '.$OSGB_X.' and OSGBY <= '.$OSGB_Y.' order by OSGBX desc, OSGBY desc) s limit
1');
	//TopLeft
	array_push($queryArray, 'SELECT * FROM (SELECT * FROM windgrid where OSGBX
<= '.$OSGB_X.' and OSGBY >= '.$OSGB_Y.' order by OSGBX desc, OSGBY asc) s limit
1');

	$OSGB_X_windgrid = '';
	$OSGB_Y_windgrid='';
	$windEnergy=0;
	$minDistance = 23000;//none of the four grid points will be more than 23km
away for wind
	$resultsArray = array();
	foreach($queryArray as $query)
	{
		$result = mysql_query($query);
		$windRasterPoint = @mysql_fetch_assoc($result);
		$distance =
distanceCalcs($windRasterPoint['OSGBX'],$windRasterPoint['OSGBY'],$OSGB_X,
$OSGB_Y);
		//choose the closest station
		if($distance < $minDistance)
		{
			$minDistance = $distance;
			$windEnergy = $windRasterPoint['wind_kwh2'];
			$OSGB_X_windgrid = $windRasterPoint['OSGBX'];
			$OSGB_Y_windgrid = $windRasterPoint['OSGBY'];

		}
	}

	//78 km is the search radius for wind station
```

```php
	$OSGB_BB_North = $OSGB_Y_windgrid + 78000;
	$OSGB_BB_South = $OSGB_Y_windgrid - 78000;
	$OSGB_BB_East = $OSGB_X_windgrid + 78000;
	$OSGB_BB_West = $OSGB_X_windgrid - 78000;

	//looking at the closest raster cell, what stations were used to determine
it (within 78km)
	$queryWindStations78km = 'SELECT * FROM (SELECT * FROM windstations where
OSGBX >= '.$OSGB_BB_West.' and OSGBX <= '.$OSGB_BB_East.' and OSGBY
>='.$OSGB_BB_South.' and OSGBY <= '.$OSGB_BB_North.' order by OSGBX, OSGBY) s ';
	$queryWindStation78kmResults = mysql_query($queryWindStations78km);

	$stationEnergyArray = array();
	$stationObjArray = array();
	$sumValues = 0;
	$stationName ='';
	$distance2WindStation =0;

	//iterate through stations within 78km bounding box to find how many are
actually
	//within 78km buffer ring and what their sum is (for standard deviation
calcs)
	while($rowWindStations = @mysql_fetch_assoc($queryWindStation78kmResults))
	{
		$distance2WindStation = sqrt(pow($rowWindStations['OSGBY']-
$OSGB_Y,2)+pow($rowWindStations['OSGBX']-$OSGB_X,2));

		if($distance2WindStation < 78000)
		{
			$stationName = $rowWindStations['src_name'];
			$energyAtStation = $rowWindStations['wind_KWH2'];
			$LAT = $rowWindStations['WGS_LAT'];
			$LON = $rowWindStations['WGS_LON'];
			$SRC_ID = $rowWindStations['src_id'];
			$obj = array('src_id'=>$SRC_ID, 'LAT' => $LAT, 'LON' => $LON,
'stationname' => $stationName, 'energyatstation' => $energyAtStation);
			//create an array of solar station objects
			array_push($stationObjArray, $obj);
			array_push($stationEnergyArray, $energyAtStation);
			$sumValues += $energyAtStation;
		}
	}
	$len = count($stationEnergyArray);
	$arr = array("StationObjArray"=>$stationObjArray,
"StationEnergyArray"=>$stationEnergyArray, "Station"=>$stationName,
"Energy"=>$windEnergy, "Distance"=>$distance2WindStation,
"NumberStationsUsed"=>$len, "SumValues"=>$sumValues);
	return $arr;
}
/*
 * end of windCalcs.php
 */
?>
```

Appendix B

Project Proposal

Investigating the Spatial Variation of Natural Energy Potential in the UK

Introduction

Through the increasingly popular Google Earth geo-browser, this project will allow users in the UK the opportunity to enter their residence postcode, and view information on the potential renewable energy that could have harnessed in the last year if they had been using a solar panel and/or wind turbine.

The system architecture will be based on a MySQL database being queried by a PHP script that is called from a Google Earth KML file. If City University's servers prove to be a viable option, they will be used to host the scripts and data; otherwise a professional web hosting company will be contracted.

A continuous raster surface will be produced as part of the final report, to allow the user to visualize the full scope of potential renewable energy across the UK.

The data to be used belongs to the Met Office, recorded at various land surface weather stations throughout the UK. Data is provided through the British Atmospheric Data Centre (BADC).

Aims and objectives

This project aims to educate the user on the available renewable energy resources in their area. This will be achieved through the creation of a server scripted application which is called from a Google Earth KML file.

Specific objectives of the project will act as tasks or iterations in the project:

- Evaluation and cleansing of weather station data
- Set up and load data into local MySQL tables
- Design PHP code to determine closest Met Office station
- Design PHP code to query table for data from relevant station

- Determine algorithms to calculate solar and wind energy
- Design PHP code to calculate natural energy produced
- Design PHP code to output results as KML
- Return information on potential renewable energy production to user
- Display information clearly in a well designed layout.
- Obtain web domain and move system (database, PHP script) there
- Produce continuous raster of potential renewable energy production in the UK for overlay in Google Earth
- Investigate real time updates of weather station data (RSS feeds) to be read by network links from KML to determine highest potential energy production locations in real time
- The final tangible deliverables will be a KML file, probably no bigger than 5kb, and a report on the project's discoveries

Scope and definition

This project will cover the entire UK; the data analyzed will be over the 2006 calendar year. The data used will be *UK Hourly Weather Observation data*, very important if one considers the sensitive characteristics of energy production from wind turbines - the energy produced is directly proportional to the cube of the wind speed.

Consequently, measuring wind energy production is difficult because gusts of wind produce large amounts of energy, while a slight breeze produces very small amounts. To truly record those gusts and the energy they produce, we would need to consider wind speeds minute by minute. Our data is not that detailed, and the algorithms devised will make the best possible use of hourly wind speed averages and hourly max gusts.

The accuracy of results returned to users will be limited by the degree of spacing in between Met Office weather stations. The project will aim for an early February finish, with the accompanying website to be sacrificed if time constraint becomes an issue.

If the project is success, the updating of the data to 2007 weather records will be considered, and there is room for a comparison analysis between successive years.

Research context

Throughout the European Union, there is a push to better utilise renewable energy resources, mainly wind and solar radiation. In fact "Government legislation requires that by 2010, 10% of electricity supply must come from renewable sources" www.embracewind.com

With regards to solar energy, it is estimated that "within five years, solar power will be cheap enough to compete with carbon-generated electricity, even in Britain, Scandinavia or upper Siberia. In a decade, the cost may have fallen so dramatically that solar cells could undercut oil, gas, coal and nuclear power by up to half." www.telegraph.co.uk

This has led to many people purchasing wind turbines and solar panels, often at great financial expense. While the hearts of these natural energy pioneers are in the right place, questions have surfaced recently as to the actual effectiveness of their eco-friendly energy producers.

Across the UK, on any given day, the variation in wind speeds and solar radiation varies greatly. As a generalization though, many places tend to experience high winds and mostly cloudy days, while other areas are more prone to low winds and sunny days.

These variations in weather patterns over space will influence the amount of renewable energy that can be harnessed through wind turbines and solar panels. Indeed it could prove debatable whether or not small scale natural energy production is a worthwhile effort in some parts of the UK, while in other regions, citizens might be able to vastly cut their energy bills.

Search the internet for wind farms, and the opinions on their benefits differ widely. This is because there are so many variables involved in renewable energy production, allowing for the bombardment of statistics that could prove cases on either side of the argument. The main variables that we will consider are wind speed, solar radiation, and weather station location.

The BBC dedicates an expansive web site to climate change (http://www.bbc.co.uk/climate/), and a section of it to wind power (http://www.bbc.co.uk/climate/adaptation/wind_power.shtml). A BBC experiment has involved a man trying to live ethically, which involved purchasing a wind turbine.(.../are_domestic_wind_turbines_an_ecocon.html)

The British Wind Energy Association (http://www.bwea.com/) web site provides (somewhat biased) details on wind production.

The department for Business Enterprise & Regulatory Reform, BERR, (http://www.dti.gov.uk/energy/sources/index.html) provides good information on energy resources.

The Energy Savings Trust (http://www.energysavingtrust.org.uk/) provides good information on solar panels in the UK.

Michael J Barnsley teaches a course at Swansea University and has written "*Environmental Modelling: a Practical Approach*", which provides a few chapters on wind and solar energy modelling.

Methodology

This project will aim to follow the work plan as closely as possible (see below). A blog will detail the project's progress, and a bi-product of the project will be a website explaining methodologies and outcomes.

As this project is based on new and advancing geographic "NeoGeo" technologies, project goals and the tasks or methods to achieve those goals may require adaptations.

The initial phase of the project will involve planning. Requirements will be gathered, a project architecture diagram will be created, along with a database diagram detailing table schema.

The second phase of the project will deal with data cleansing. The Met Office dataset will be stripped down to the necessary data in order to improve performance. It is envisioned that GAWK, a programming language, will be good for this process. GAWK provides a language for stripping down txt files to their desirable elements.

The third phase will involve setting up of the MySQL database. Tables will be created through the MySQL administrator. It is hoped that code generated from a Visio database diagram will run against MySQL to create the necessary tables. Data will then be loaded. Poor batch loading functionality MySQL could be an issue, as it has been found to be in the past.

The fourth phase will involve planning and scripting of PHP code to query the data and imbed the results within KML. The bulk of the work will be in this phase, as there is a degree of learning required on my part to be competent with PHP.

There will be two main queries to the code. The first will deal with selecting the weather station closest to the user's location. This will involve distance calculations from the user's location the stations. These calculations will need to be performed in Cartesian coordinates and the user's location will be in latitude and longitude, meaning a conversion will be required to the chosen coordinate system, probably OSGB. The issue of differing ellipsoid datum between OSGB (Airy 1830 ellipsoid) and Google Earth (WGS84 ellipsoid) will further complicate the conversion process. Due to the high number of Met Office weather stations, there is also a question of how long a query might take to calculate the distance from the user's location to *all* the stations, and then selecting the closest station. It might be beneficial to first do a point in polygon calculation, or use some other algorithm to automatically eliminate stations that are too far away. This would greatly reduce the number of distances that are calculated.

The user's location will originate from Google Earth. The user will enter their postcode, and Google Earth will zoom into the region. The KML <NetworkLink> tag embeds a links to a script that produces KML for display in Google Earth. The <NetworkLink> tag will be set to be called when the zoom has settled on the appropriate region. When the <NetworkLink> tag is used, the only parameter it passes to the script is the bounding box of the user's current view. The centre of this bounding box will be the user's location.

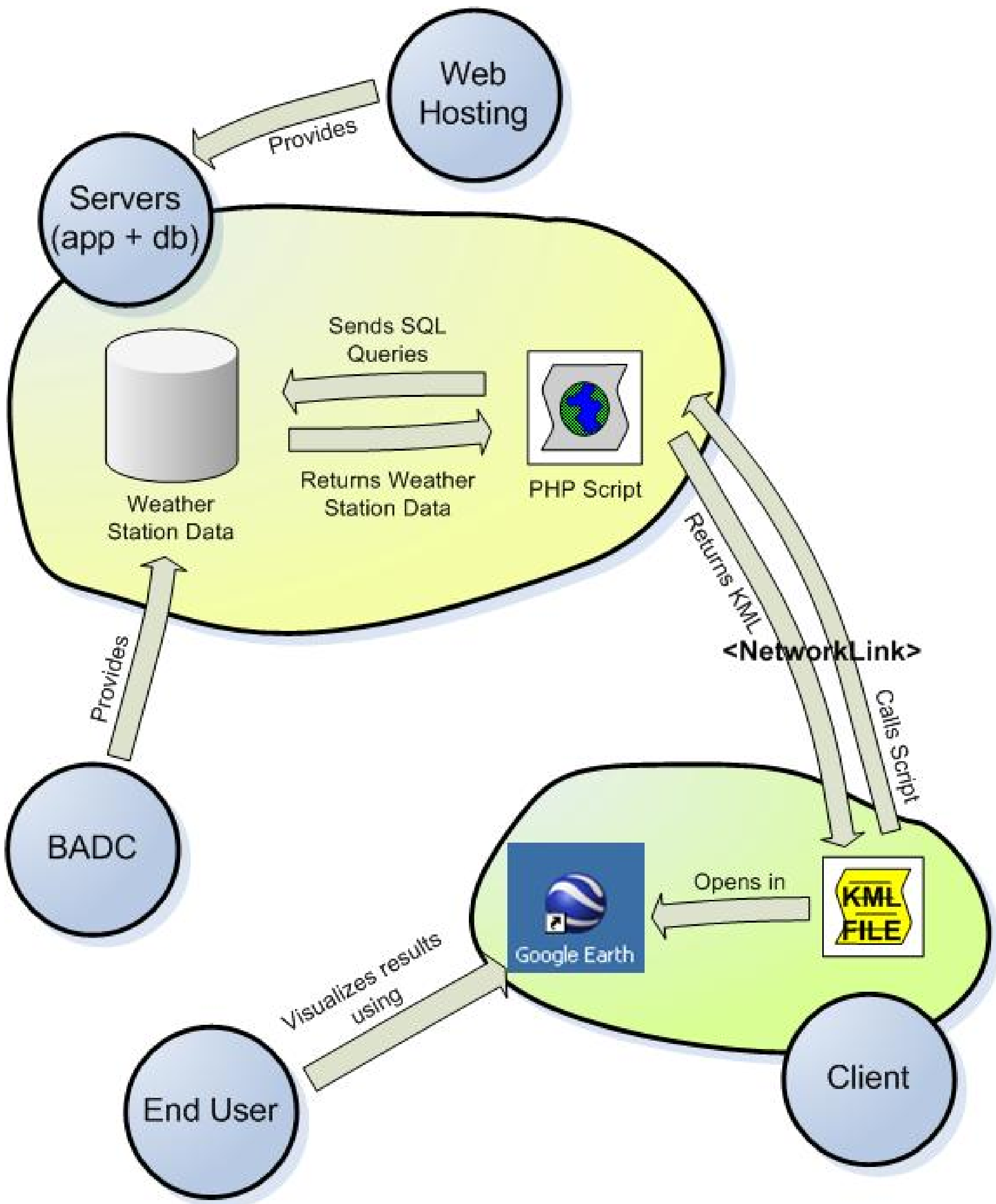
The second query will deal with selecting the data from the appropriate station; this should be straight-forward.

The fifth phase will consider algorithms for determining energy produced. A default size of solar panel and wind turbine will be chosen. The most accurate algorithms relating to the field will be scripted into PHP code to calculate the actual solar and wind energy produced throughout the entire last year (2006).

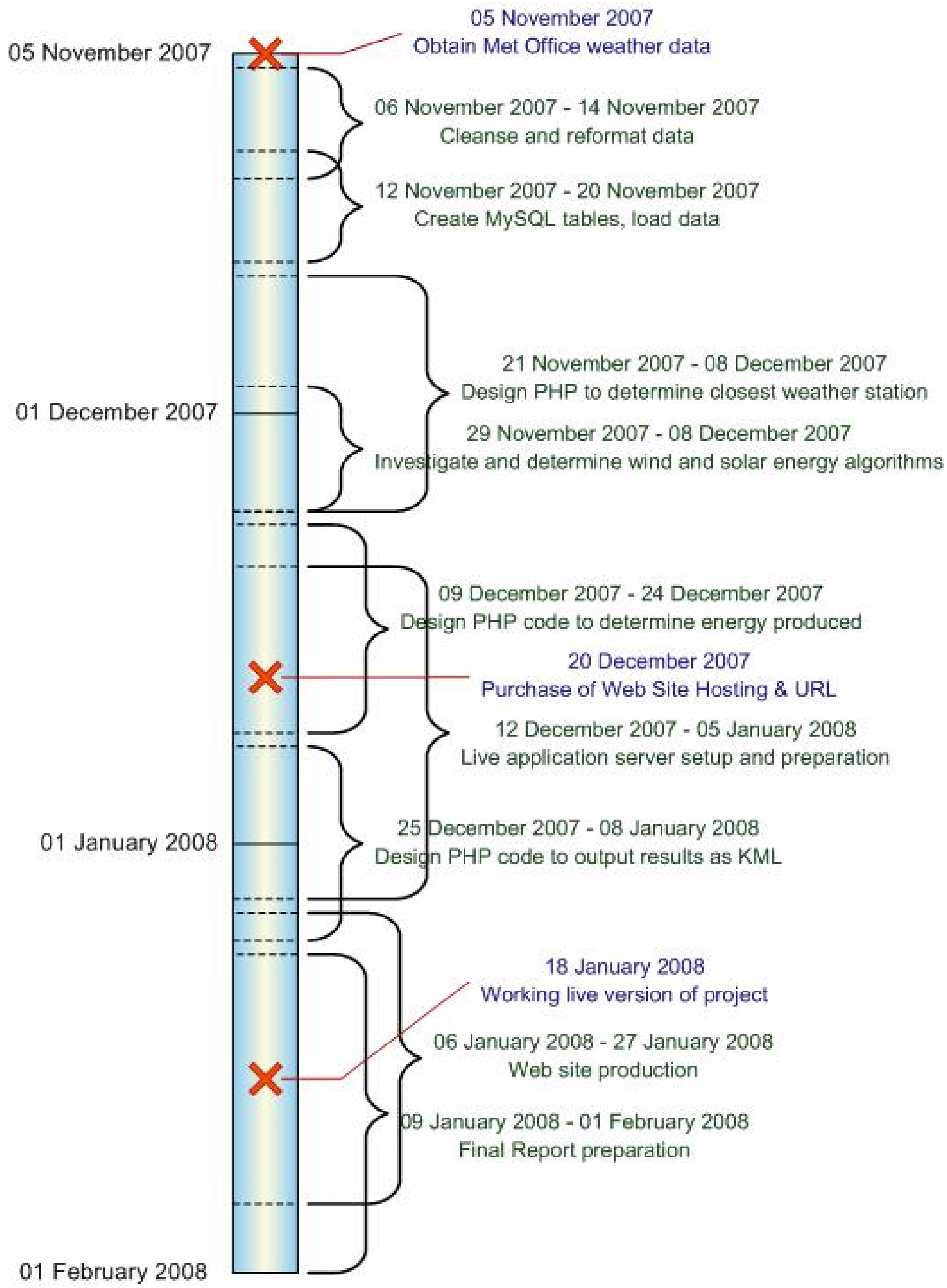
A sixth phase will deal with the creation of the KML tags that will display the output to the user. A pop up balloon style of display is envisioned; this will show energy produced by wind, by solar radiation, and by the two combined. Other interesting stats such as light bulbs lit, clothing washes done, or roast dinners cooked might also prove interesting.

The seventh and final phase will involve shifting the project over to a live hosted server where the data and PHP code can reside, and be accessed by the public viewing the KML file. The live server will be either the Apache HTTP server or Windows' IIS. The servers will need to be set up to run PHP code, and the MIME types will need to include KML and KMZ extensions to open automatically in Google Earth. Storage space issues will need to be considered as the datasets of hourly weather records is quite large (data as text file ~800mb).

System Architecture



Work Plan



05 November 2007
Obtain Met Office weather data
05 November 2007
06 November 2007 - 14 November 2007
Cleanse and reformat data
12 November 2007 - 20 November 2007
Create MySQL tables, load data
21 November 2007 - 08 December 2007
Design PHP to determine closest weather station
01 December 2007
29 November 2007 - 08 December 2007
Investigate and determine wind and solar energy algorithms
09 December 2007 - 24 December 2007
Design PHP code to determine energy produced
20 December 2007
Purchase of Web Site Hosting & URL
12 December 2007 - 05 January 2008
Live application server setup and preparation
01 January 2008
25 December 2007 - 08 January 2008
Design PHP code to output results as KML
18 January 2008
Working live version of project
06 January 2008 - 27 January 2008
Web site production
09 January 2008 - 01 February 2008
Final Report preparation
01 February 2008

Resources

The project will require the following:

- MySQL database
- GAWK programming language
- data of hourly wind and solar observations from weather stations across UK
- algorithms for conversion from WGS84 to Airy ellipsoid and then to OSGB projection
- algorithms for solar and wind energy production
- testing environment (local setup on laptop)
- live environment - a hosted server, approximately £60 /year
- Google Earth

Ethics

While there are no issues covered here deemed offensive, it is hoped that this project causes users to consider the ethics involved in their daily energy use.

Confidentiality

The Met Office weather station data provided through BADC is strictly confidential, and is only to be used for the purpose of this MSc dissertation. A confidentiality agreement was entered between myself and the BADC, and was also signed by Jo Wood.

Appendix C – Reflections on the Dissertation Process

The dissertation process has been a very good learning experience. Data procurement through the British Atmospheric Data Centre gave me experience with data licensing and confidentiality. The sheer volume of the weather station data sets provided me with valuable practice in handling large datasets.

Data manipulation is such an essential GIS skill, and this project provided vast amounts of experience. Spreadsheet skills were gained and structured query language (SQL) skills were learned in two different database management systems, Microsoft Sequel Server (MSSQL) and MySQL Needing to use a server scripting language within this project I have also learned a new programming language, PHP.

The importance of data uncertainty and its relationship with resolution, precision, and accuracy is another valuable skill learned within this project. While values reported back to the user are central to the project, of great importance is reporting back to the user an idea of the confidence in the accuracy of the values being given to the user.

The proposal mentions seven main phases. The initial phase of planning was conducted through the actual preparation of the proposal, while the second phase of data preparation and manipulation was conducted on schedule. The third phase of database creation and maintenance was conducted to schedule, and database loading problems did not materialise as steps were taken to keep database size down. The fourth phase involved coding in PHP to script database queries and KML semantics data structuring. Time taken to complete coding tasks was greater than originally thought due to the time required to learn coding languages. The fifth phase which involved calculating energy from weather station data reports, was actually performed much earlier in the project, because it was determined that raw data would not be imported in the project databases. The energy from stations was pre-calculated outside the main tool, and then entered into the project databases as base data. The sixth phase was formatting of the output, a sort of “prettying-up” to the Geo-browser’s output, and the seventh phase of moving the system to live servers hit snags with permissions and environment settings, all problems eventually solved.

Probably the largest amount of work actually involved the post-analysis of the results, and putting this analysis into a well structured paper, using strong arguments and existing known practices, while at the same time also attempting to involve my own methods of performing tasks and evaluating empirical data.