Citizen Science

1. Introduction

Citizen Science, also referred to as crowd-sourced science or crowd science, is where scientific research is conducted with the assistance of volunteer, amateur or non-professional scientists. It covers the many different ways in which citizens can participate in scientific research. Some good examples, include Agent Exoplanet (https://lco.global/agentexoplanet/) in astronomy, where interested citizen scientists are examining imagery from space to determine the presence of exoplants; ARTigo (http://www.artigo.org/) where people are helping to search vast data repositories for reproductions of historical art and the International Environmental Data Rescue Organization (IERDO) (http://iedro.org/) which is trying to rescue millions of historic climate data records globally through the mass digitisation of decaying paper based records. Globally, hundreds of thousands of people have participated in citizen science projects that have led to a new and improved understanding of our world and the people and places that occupy it.

As part of this module, you will be contributing to new citizen science project. The outcome of your involvement will lead to significant improvements in our existing climate data holdings and ultimately lead to new and interesting scientific discoveries. While the topic of this citizen science project looks at improving climate data holdings, the skills you will develop as part of this citizen science project are directly transferable to any research area. Essentially, you are being trained as a researcher, rather than being taught how to be a researcher. You will play a central and active role in achieving this.

Being involved in the process, in addition to partaking in actual, meaningful research, you will develop new research skills and enhance the ones you already have. To successfully complete the research, you will need to use a variety of tools, some of which will be new; you will develop skills in comparing different sources of information which may be in conflict with each another; you will learn how to access different sources of information; you will be working with, and undertaking an analysis of, data. You will also develop your reasoning skills and lateral thinking ability. These are the qualities and skills that you will need when you go on to Third Year and are strongly sought after by employers, in an increasingly competitive job market.

2. Project Context

Copernicus Climate Change Service

The Copernicus Climate Change Service (C3S) is operated by the European Centre for Medium Range Weather Forecasting (ECMWF) on behalf of the European Union. The service will provide access to information for monitoring and predicting climate change, required to support climate change adaptation and mitigation policy development across Europe and globally. In order to deliver the service, C3S has commissioned a number of research projects to evaluate, quality control and consolidate all available, existing climate data information which currently exists in numerous different formats and locations. One of these commissioned projects, entitled '*Copernicus Climate Change Service Global Land and Marine Observations Database*', is being led by the Irish Climate Analysis and Research UnitS (ICARUS) at Maynooth University. The project team is comprised of Prof. Peter Thorne (Project Lead), Dr. Simon Noone, Dr. Anthony Kettle and Corinne Voce.

Copernicus Climate Change Service Global Land and Marine Observations Database

We cannot predict what is not observed, and we cannot analyse what is not archived. It is vital that the management and curation of land-based meteorological data holdings is improved. To address this issue the team at *ICARUS* are working together with international partners with the aim to provide an easily accessible comprehensive and consolidated global set of data holdings. The collated meteorological data will be made available in one place and one format, which will realise the true global value of the historically collected meteorological data.

The Challenge

There are several challenges within these climate data holdings. We have data from in excess of 250 different sources. These challenges exist not just in the data itself (eg. missing values, incomplete data etc) but also in the metadata (eg. information about the data) – such as the geographic location of the stations. Many of the station locations are simply incorrect. For example, over 10,000 land stations have coordinates that place them in the ocean! It may be just that the station coordinates have slightly misplaced the land station over the water or, in extreme cases, the sign of the station's geographic coordinates place it in the wrong hemisphere!

It is this aspect which shall be the focus of your work. There is no substitute for human eyes and investigatory powers to disentangle these issues. Specifically, your eyes and investigatory powers will help us resolve issues for land stations with coordinates that are known to be incorrect. All stations resolved will then be used in the final data holdings so you shall be integral in building this new set of data holdings which will be made available to global research community.

Good luck, we have every faith in your ability, collectively working as citizen scientists, to solve this puzzle.

Project: GEO Locate land **O**bserving **C**lim**A**te s**T**ation attribut**E**s (GEO-LOCATE)

3. Project Details

Each of you has been allocated 15 stations on an excel spreadsheet. Each of the 15 stations you have been supplied with has associated geographic coordinate information, which is known to be incorrect. All the stations that you will be working with are known **LAND BASED** stations, but due to either the incorrect recording of their geographic location information (ie. latitude and longitude) or the use of imprecise geographic coordinate information, they have coordinates that place them over water bodies.

You will be responsible for carrying out the research and to gather evidence to support the relocation of each of the 15 stations to their correct geographic coordinates on land. You will need to find any and all available evidence to support your conclusions as to why a particular station's coordinates should be as you indicate.

You will need to employ a variety of research tools at your disposal, including Google Earth, Google Maps, Google searches, Google Translate, dedicated climate data information sources (eg. national weather services from different countries) and any other available research tools and information in order to determine the correct locations for your chosen list of stations.

While we are there to help and assist you, you will be carrying out the basic research to provide results that support this international research effort on improving our meteorological data holdings.

The project deliverables will be based on

1) A spreadsheet with the station list, original coordinates and new, corrected coordinates. The spreadsheet also requires you to detail how you obtained the updated coordinates and to add comments briefly outlining your sources for the new coordinate information. An example spreadsheet is provided in the Session6 folder. Marks for the completed station .xls file will be based on the correct number of stations completed with full details/comments/supporting information. (**CA2: 35%**).
2) A presentation detailing the research methodology and methods you undertook to identify and correct each station's geographic coordinates. The presentation should also contain an overview of the arguments to support the relocation of a station to its new location. (**CA3: 15%**).

4. Getting a sense of the issues with the existing station locations

This section will outline the methods needed to import the station data into Google Earth and to visualise the current stations locations. For many of the existing station locations, you should be able to determine/narrow down your research focus on the basis of viewing the station locations relative to the surrounding land.

Open your allotted station data in excel, by browsing to the file location (eg. F:\GY201 Data) and double clicking the station file. Explore the data. You will note that some of the coordinate or other information has an NA (not available) associated with particular stations. This project is seeking to find out the missing or incorrect coordinate and other information for each of your selected stations.

As Google Earth can import comma delimited or CSV file types, but not .xls or xlsx (Excel file types), you need to change the file format of the station list. To save the file as a comma delimited or CSV file type, in MS Excel select **Save As**. Under *Save As Type*, select the CSV (Comma delimited) (*.csv) option. Give the file a suitable name and click **Save**. This will save the station file as a comma delimited or CSV file type which can be imported into Google Earth.

Google Earth Pro can be launched from the Windows Start menu, scroll down to the G programs block (all programs starting with G are listed here) and select the Google Earth Pro icon.

Check the latitude longitude coordinate system displayed in the bottom of the viewing window to make sure the coordinate information is being displayed in decimal degrees. If the latitude longitude coordinates are displayed in Degrees, Minutes, Seconds (with °,', '') you will need to change the Google Earth settings. This can be done under **Tools** > **Options**. Under the *Show Lat/Long*, change **Degrees, Minutes, Seconds** to **Decimal Degrees**. Click **Ok**.

To add the CSV file to Google Earth, go to the **File** > **Open** or **File** > **Import** on the menu bar. You will then be presented with an Open dialogue box, where you can browse to the location of your file. If you know the extension of your file (eg. .kml, .shp, .txt, .csv), select that file type from the file type drop down menu box. Select *Generic Text (*.txt *.csv)*. Once you have selected the file name and file type, click **Open**.

Google Earth will then launch the Data Import Wizard. You have previously imported .csv data into Google Earth. However, you will need to identify the columns in the data file that contain the **Original Latitude** and **Original Longitude** so that Google Earth can use this information to locate the stations on the map.

Select **Yes** to *Do you want to apply a style template to the features you ingested?* and *Create new template*.

Under the *Name* tab, select STATION_NAME in *Set name field*.

Under *Color*, accept the default of *Use single color*.

Under the *Icon* tab, choose an icon.

Under the *Height* tab, use the default option *Clamp features to ground*.

Click Ok and save the Style template to a location you remember (F:\GY201 data).

Saving the imported data as a .Kmz file

Once the file has been imported successfully, save it as a .Kmz file which can be directly opened in Google Earth without having to import the file every time you want to view it.

To save the imported data, right click on the layer name and select **Save Place As**. Under *Save as type*, select Kmz (*.kmz). Browse to a location that you want to save the file (ie. F:\GY201 Data) and give the file a name that is meaningful and which you will remember (eg. *Yourname*_stations). Click **Save**. This will save the file with given name as a kmz file in your chosen location (ie. F:\).

You can now delete the imported csv layer file (eg. right click in the layer file and select **Delete**) and add the saved Kmz file containing your station information (**File** > **Open**, browse to location that you saved the .Kmz file. Make sure either the Google Earth (*.kmz, *.kml etc) or All files (*.*) file type option is selected in the file type drop down menu. Select the Kmz file containing you station data and click **Open**). Once you have saved the file as a Kmz file type, you can directly import the file in to Google Earth in the future (**File** > **Open**).

5. Project Workflow

In the example below, the station layer file is called *My list of 15 stations*. If you followed the instructions above, the layer file should be visible as *Yourname*_stations.

Select the check box next to the layer file to make the stations visible.

As these stations are distributed globally you may not immediately see them displayed on the virtual earth; to quickly locate each station, click the ▶arrow next to the layer name to expand the contents of the layer (Figure 1.). You may need to expand subsequent layers (ie. Station List in Figure 1 below) to make the station names visible (assuming you set the STATION_NAME in Set name field in the Style Template).

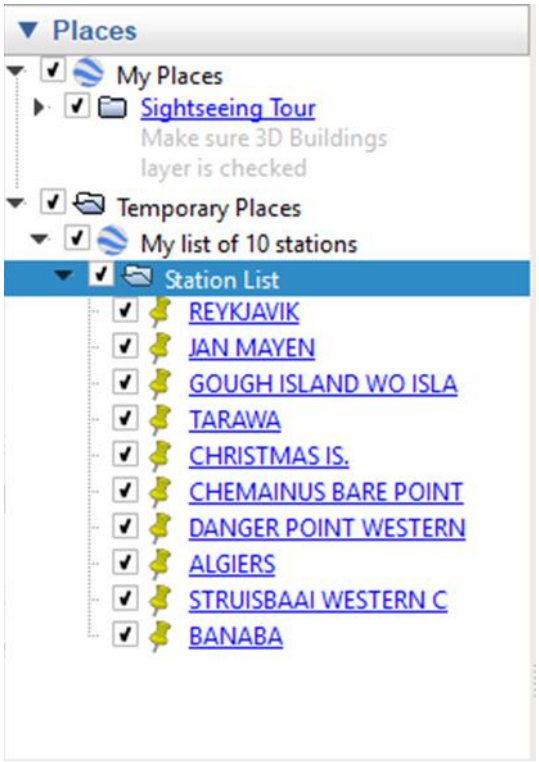

Figure 1. Expanded station list after clicking the ▶arrow next to the layer name and Station List. (Layer view, © Google Earth, earth.google.com/web/)

Double click on a station name and Google Earth will automatically zoom to that station's (incorrect) geographic coordinate location. You may have to zoom out to get a sense of where you are!

Ensure the **Borders and Labels** layer is turned on with a check mark, under the Layers module sidebar window. This will turn on all the available place names in Google Earth and make it easier to identify where you are on the virtual earth.

Based on the example of Reykjavik in Figure 2 below, it is clear that the station location has been misplaced over the sea, relative to the position of Reykjavik city and surrounding land surface. Therefore, the possible search area for the correct location of this station is likely to be well defined. Remember, this is just an example. It is possible that some of your stations have a '-' missing (ie. minus sign) from the coordinate information and be placed in the wrong hemisphere! For example, an incorrect latitude of 33.868160 (missing the – sign to place it in the southern hemisphere) would place Sydney, Australia, somewhere off the coast of Japan in the northern hemisphere!

Figure 2. Position of Reykjavik meteorological station, a land based station, based on the incorrect geographic coordinates provide in the station file. Also shown is the city of Reykjavik. (Map of Iceland. © Google Earth, 2019, earth.google.com/web/)

Please note, **DO NOT** just get the coordinates for Reykjavik city and use those for the new coordinates for the Reykjavik meteorological station. You are required to locate the station as close to its **ACTUAL** location through a process of research and be able to justify the selection of new coordinates.

To help achieve this, the World Meteorological Organisation (WMO) has developed a tool that contains details on many of the stations that contribute to international meteorological data collection efforts. The tool has a station search function and contains useful metadata, including actual geographic coordinates of many of the meteorological stations globally. While not all stations are available in this catalogue, it provides a very useful starting point to check if the correct station details are catalogued.

6. Suggested order of steps to finding the correct coordinates (latitude/longitude) for your stations

Step 1. World Meteorological Organization (WMO) Observing Systems Capability Analysis and Review Tool (OSCAR)

OSCAR is a resource developed by WMO in support of Earth Observation applications, studies and global coordination. It contains quantitative user-defined requirements for observation of physical variables in application areas of WMO (i.e. related to weather, water and climate). OSCAR is the World Meteorological Organization's official repository of WIGOS metadata (information about the data, including geographic location and station elevation data) for all surface-based observing stations and platforms.

a) Go to World Meteorological Organization Observing Systems Capability Analysis and Review Tool (OSCAR) webpage: https://oscar.wmo.int/surface//index.html#/.

b) Enter the station name from your station list in quick access drop down menu (Figure 3). Click on the *Station name* field under *Generate station report by* in **Quick access** and start to type the station name. As you start to type, suggested station names will start to auto populate the search field (Figure 4).

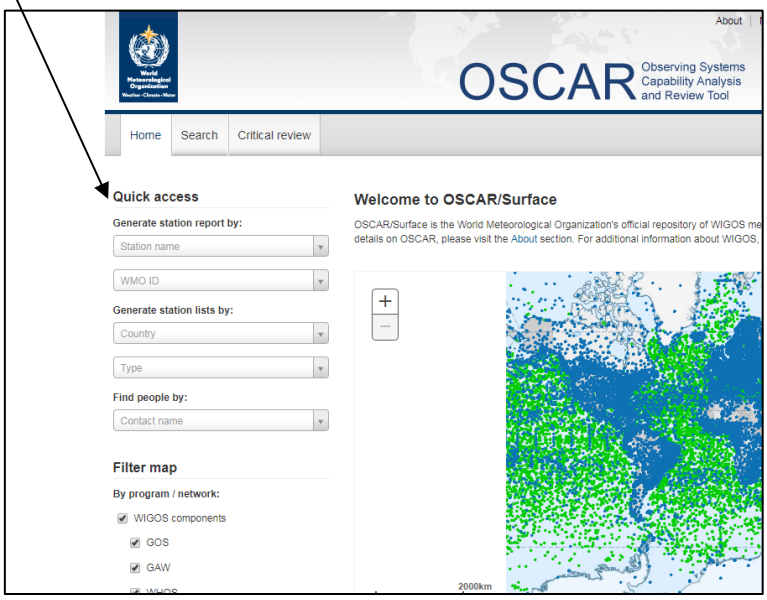

Figure 3. World Meteorological Organization Observing Systems Capability Analysis and Review Tool (OSCAR) (available at https://oscar.wmo.int/surface/#/).

In some instances, WMO OSCAR may present you a list with a number of stations with similar names (Figure 4). There may be multiple land based stations but if you check the latitude/longitude coordinates you will see that they are very similar and are located in close proximity to one another, so just choose the first **land based** (ie. Land (fixed)) station.

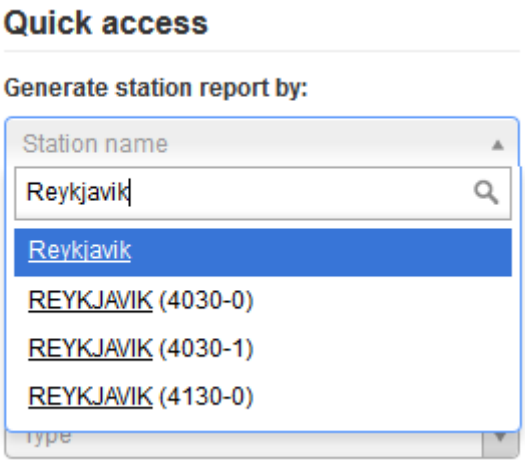

Figure 4. WMO OSCAR search function. (available at https://oscar.wmo.int/surface/#/).

c) **If the stations name exists in the OSCAR system**

Clicking the station name in the search box will open with the station information page catalogued in WMO OSCAR. *Station type* indicates whether it is a Land (fixed), sea (eg. data buoy) or underwater type station. Make sure the station you select is a **Land (fixed)** station type.

The Coordinates provided by OSCAR are the most accurate ones at the highest precision and should be entered into the revised latitude/longitude columns in the excel sheet for the selected station. Also enter the station elevation information (e.g. 52) into the revised elevation column but **omit** the m (Figure 5).

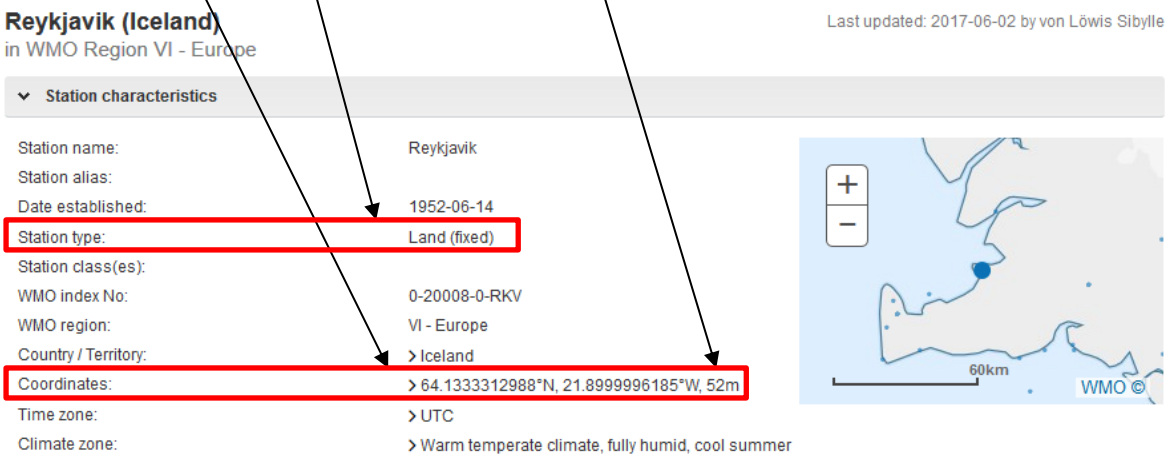

Figure 5. Details of the station entered in the OSCAR search box returned from the OSCAR database (available at https://oscar.wmo.int/surface/#/).

In the revised information source column on your excel spreadsheet containing your list of stations, enter the OSCAR derived details. Under the "**Step 1**" column header for that station, explain in detail how you got the coordinate information. In addition, enter any other relevant information on the station in the "**Comments**" column that may be useful. An example of a completed excel station file, with comments, is provided in the Session6 folder. **If the station name is not in OSCAR, put an NA under Step 1 for this station.**

IMPORTANT: You need to enter the decimal degrees (DD) latitude longitude coordinates in the excel sheet therefore:

If the OSCAR station latitude coordinates are ºS then you should place a minus sign before the numbers.

If the OSCAR station longitude coordinates are ºW then place a minus sign before the numbers.

In the excel sheet cell just type the – sign before the numbers and click **enter**.

For example, Reykjavik station in OSCAR is located at latitude 64.1333312988°N and longitude 21.8999996185°W. In excel, enter 64.1333312988 for the latitude coordinates. As the longitude is °W (west of Greenwich) you enter -21.8999996185 for the longitude coordinates.

N.B. **DO NOT** enter the symbols º S, º N, º E or º W in the excel spreadsheet.

If the station name is not in the OSCAR system, then try the following step next.

Step 2. Check the meteorological agency of identified country of origin

1) Plot the station using the provided incorrect coordinates. In many cases, the incorrect coordinates are due to missing places of decimal. If this is the case, the incorrect station coordinates will place the station off shore from the country that it should be located in.
2) Once you determine the likely country that that station is from, check to see if station information from that meteorological agency has been provided for you on Moodle. A number of files have been placed on Moodle that contain station information for selected regions or countries. If the station does not exist in one of these files, please continue with the following steps.
3) Do a Google Search to look for the meteorological agency or service for that country. If you find a meteorological agency for the country, you can use Google translate to try to make sense of the webpages you find.
4) Many meteorological agencies provide station lists and station information, such as coordinates, on their websites. Look for somewhere on the site that might contain station information, often under station list or meta data.
5) If you manage to find a station list for a particular country, please download it and send it to the class teacher so that we can provide this for other students.
6) Coordinates from a meteorological agency or service should be treated as a 'trusted' source as they are typically government agencies.
7) Under "**Step 2**" in the spreadsheet, provide sufficient details on how you located the station coordinates using this method. Include a link to the meteorological agency website under the Revised Information Source column and provide any additional details under the Comments column.
8) If you are unable to locate the station coordinates using this method, put an NA under Step 2 in the spreadsheet.

For example, we can see from the plot of the station in Google Earth that the Reykjavik station is located just off the coast of Iceland. Using a google search with the terms 'Iceland meteorological agency' will locate the Iceland Meteorological Office home page. Fortunately, the web site is in English, therefore we do not need to use google translate to translate the page into English. It is likely that the weather stations are located under the Weather link at the top of the page. Click the link Weather. Under the map of Iceland, there is a link to Station List, from which you can select the station you are looking for ie. Reykjavik. Select the Reykjavik station link and then select *Details about the station*. Coordinate and height information about the station are listed. This is the information you add to your spreadsheet for this station. Make sure to add the relevant link to the station details and a comprehensive comment on how you found the station information, under the Comments section of the spreadsheet.

If the station name is not in the OSCAR system and you cannot locate the station via a national meteorological agency, then try the following step.

Step 3. Using Google Search

1) Type the station name into a Google search and add the words *weather station* and/or *latitude longitude* to the search. For example: *Reykjavik weather station*. If the information you have is not the full name of the station or location, then Google sometimes provides prompts of the full details. The search should show some webpages that provide information on the weather station. In this instance, the first hit is from the Icelandic Meteorological Office, the official meteorological agency of Iceland and thus, the information provided from this site should be considered of high/good quality. If you scroll down the list of sites returned, you will see results from sites such as accuweather.com, yr.no and weatherunderground.com. There are companies that sell or utilise weather information from personal weather stations for the home user, which these companies provide freely on the web (eg. https://www.wunderground.com/). As we are trying to locate the official weather station information, please ensure that you are not providing coordinate information for a personal weather station deployed in someone's back garden. For example, the official coordinates from the Iceland meteorological agency for Reykjavik weather station are 64.1275 lat and -21.9028 lon; the coordinates from the www.yr.no website for Reykjavik are 64.1355 lat and -21.8954 lon. You will need to make decisions about the validity, reliability and type of web page you are sourcing your information from to avoid this potential pit fall!
2) Check the original country information to see if it matches or makes sense. If the original country information is missing, then check the latitude longitude to see if they are similar or close in location. Often the coordinates are offset or may be placed in the wrong order (e.g latitude where longitude should be and vice versa).
3) You will then need to investigate these web pages and find the latitude and longitude coordinates for the station and enter them into the revised latitude/longitude columns. **Enter as you see them including the minus sign (-) if applicable**, **but without the º S, º N, º E or º W.**
4) Also, if possible find the station elevation/altitude and enter into the revised elevation column in the excel spreadsheet. **N.B** make sure elevation is in metres, if shown in feet then please divide by 3.28 to convert to meters.
5) Revise the station country location if needed.
6) In the "**Revised_Information_Source**" column on the Excel spreadsheet, enter the webpage link you used and in the "**Step 3**" column explain in detail how you obtained the coordinates. For example, you will need to include the search term you used to find the station on Google.
7) In addition, please enter any other information on the station in the "**Comments**" column that may be useful.
8) If you obtain station information from a site that is not from an official government or trusted entity, you should try to verify your information using additional sources.

Google Earth also has a very useful geocoding component. As you have a location description in the form of a station location name, enter the station name in the Google Earth search box to see if Google Earth can locate the station. However, you should try to verify your station coordinates if you use this method.

There are a number of websites that will allow you to convert between DMS (Degrees Minutes Seconds) and DD (Decimal Degrees). Make sure that your coordinate information is in Decimal Degree format for entry into the excel spreadsheet.

Evaluation Step: Even though the WMO provide the official station coordinates for all stations contained in the WMO OSCAR data base, these may be small accuracy errors in the official WMO station coordinates. Therefore, as an extra step, you are required to try to verify the coordinates using alternative sources. For example, even if the WMO OSCAR data base contains geographic coordinates for your station, you could verify the coordinates provided by OSCAR by following the instructions in Step 2 (ie. performing a Google search to locate a country's meteorological agency website and then looking for the station coordinates on the site or checking in one of the station meta data files that have been provided for you on Moodle). Please indicate your additional sources in the Comments for that station. If the national meteorological agency provide information that is different from OSCAR, use the coordinates from the meteorological agency – this is considered the most 'trusted' source of information, but clearly indicate in your comments that you have done this and also indicate that the station was found on OSCAR but that you provide the meteorological agency coordinates (include the web address).

Verifying your information is an important step in any research project you undertake.

Figure 6 shows the slight variation in the geographic coordinates derived from the different sources. The official coordinates provided by WMO OSCAR (Step 1) are very close (approx. 600 metres) to the actual location of the meteorological station found based on a Google search using the term "Iceland meteorological office" (Step 2) and locating a station list (under the "Weather" link, then "Weather Stations" and select "R" for Reykjavik and select "info" for the station).

Figure 6. Location of Reykjavik city centre, Reykjavik weather station location using WMO OSCAR coordinates (Step 1) (Yellow Placemarker) and coordinates of the Reykjavik weather station according to the Icelandic Meteorological Office (Red Placemarker) based on a Google search using the term "Iceland Meteorological Office" (Step 2). (Map of Reyjavik, Iceland. © Google Earth, 2019, earth.google.com/web/)

Downloading the station files for the project

Each of you has been allocated a number from 1 to 100 which is linked to specific file which contains your list of 15 stations. Each of the files are numbered from 1 to XX (eg. 1.xlsx to XX.xlsx). To check which set of station files you have been allocated, please check the **GY201Aallocations.xls** file, which is available on the GY201 moodle page (e.g. A, B , C etc). Check your number allocation.

Download the today's zip file (i.e. Session6.zip) from the GY201 Moodle page into the appropriate folder on your USB drive (i.e. F:\GY201 Data). When prompted select Save As .. and navigate to the correct drive, F:\GY201 Data, and hit Save. Then navigate to that drive and folder within My Computer or Windows Explorer and right-click on Session6.zip and select the option Extract all. You will see that the system is now writing the files into a new folder also called Session6.

To access your allocated station, unzip the Stations.zip in the Session6 folder. Select your allocated file (ie. 23.xlsx if you were allocated the number 23) and copy it to a new folder on your USB drive under the F:\GY201 Data folder. Name the new folder "Geolocate". You can use this folder to save any files associated with the GEO-LOCATE project (F:\GY201 Data\Geolocate).

Open the file in MS Excel and save the numbered .xlsx file as *Yourname YourGroupNumber Number of spreadsheet*.xlsx. For example, Joe Blogs in Group 4 was allocated the excel spreadsheet number 45 (ie. 45.xlsx), therefore the xlsx filename should be saved as JBlogs Group4 45.xlsx. Once you have completed the necessary research and filled in the spreadsheet – this is one of the files you will be submitting.

It is good practice to keep a master copy of the data, in case you need to go back and check anything in the original file. So **keep a copy** of the original xlsx file containing the station data that you copied into the Geolocate folder.

References

Strangeways, I. (2009), Using Google Earth to evaluate GCOS weather station sites. Weather, 64: 4–8. doi:10.1002/wea.334

Jones P.D., Lister D.H., Osborn T.J., Harpham C., Salmon M. and Morice C.P. (2012) Hemispheric and large-scale land surface air temperature variations: an extensive revision and an update to 2010. Journal of Geophysical Research 117, D05127. doi: 10.1029/2011JD017139.

Osborn T.J. and Jones P.D. (2014) The CRUTEM4 land-surface air temperature dataset: construction, previous versions and dissemination via Google Earth. Earth System Science Data 6, 61-68. doi: 10.5194/essd-6-61-2014