https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

Geoscience Communication
Discussions
Open Access

Geo-locate Project: A novel approach to resolving meteorological station location issues with the assistance of undergraduate students

Simon Noone[1], Alison Brody[1], Sasha Brown[2], Niamh Cantwell[2], Martha Coleman[1], Louise Sarsfield Collins[2], Caoilfhionn Darcy[2], Dick Dee[3], Seán Donegan[1], Rowan Fealy[1], Padraig Flattery[1], Rhonda Mc Govern[4], Caspar Menkman[2], Michael Murphy[2], Christopher Phillips[1], Martina Roche[2] and Peter Thorne[1]

[1]Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography, Maynooth University, Ireland
[2]Department of Geography, Maynooth University, Ireland
[3]European Centre for Medium-Range Weather Forecast (ECMWF), Shinfield Rd, Reading RG2 9AX, UK
[4]Trinity Centre for Environmental Humanities, Department of History, Trinity College, Dublin, Ireland

Correspondence to: Simon Noone (simon.noone@mu.ie)

Abstract.

The Global Land and Marine Observations Database aims to produce a comprehensive land based meteorological data archive and inventory. This requires the compilation of available land-based station meteorological data information from all known available in-situ meteorological data repositories/sources at multiple timescales (e.g. sub-daily, daily and monthly). During this process the service team members have identified that many of the data sources contained stations with incorrect location coordinates. These stations cannot be included in the processing to be served via the Copernicus Climate Change Service until the issues are satisfactorily resolved. Many of these stations are located in regions of the world where a sparsity of climate data currently exists, such as Southeast Asia and South America. As such, resolving these issues would provide important additional climate data, but this is a very labour-intensive task. Therefore, we have developed the Geo-locate project enrolling the help of undergraduate Geography students at Maynooth University, Ireland, to resolve some of the land-based station geolocation issues. We have successfully run two Geo-locate projects, the first in the second semester of the 2017/18 academic year and the second in the 2018/19 academic year. Both iterations to date have been very successful with 1926 out of 2168 total candidate stations ostensibly resolved, which equates to an 88% success rate. At the same time, students gained critical skills helping to meet the expected pedagogical outcomes of the second-year curriculum, while producing a lasting scientific legacy. We asked the class of 2018/19 to reflect critically upon the outcomes and present the results herein which provide important feedback on what students felt that they gained from their participation and how we may improve the experience and learning outcomes in future. We will be continuing to run Geo-locate projects over the next few years. Due to the success of the Geo-locate project we encourage other organisations to investigate the potential for engaging university students to help resolve similar data issues while enriching the student experience and aiding the delivery of learning outcomes. This paper provides details of the project, and all supporting information such as project guidelines and templates to enable this.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

1 Introduction

A comprehensive Global Land and Marine Observations Database is being developed containing surface meteorological data records spanning the entire history of instrumental observation. This work is being carried out under contract with the Copernicus Climate Change Service (C3S) by the Irish Climate Analysis and Research Units (ICARUS), Maynooth University, Ireland (lead Institute). We have sub-contracted partners in the United Kingdom (UK) at the Met Office, the National Oceanographic Centre and the Science and Technology Facilitates Council. We are also working closely with the National Oceanic and Atmospheric Administration's National Centers for Environmental Information (NOAA/NCEI) who are based in the United States. The Copernicus Climate Change Service (C3S) is implemented by ECMWF on behalf of the European Commission in an effort to make climate data and information more easily available to society. Data sources for the database include major collections of historic weather data archived at NOAA/NCEI, as well as substantial holdings of meteorological *in situ* observations used for numerical weather prediction and climate reanalysis at the European Centre for Medium-Range Weather Forecasts (ECMWF). Many additional sources of data will be included, from national weather service providers, atmospheric research institutions and the multitude of historical data rescue activities taking place around the world.

Building the database requires development of data source inventories, retrieving data from all available sources, converting the data to a common representation, merging them into harmonised records, and applying quality checks at all levels. Over time, the database will be continually updated with additional observations as they become available. Access to the database will be provided by C3S via an internet-based Climate Data Store (CDS; https://cds.climate.copernicus.eu), which will also offer many other datasets and tools needed to enable development of applications of the data for a variety of purposes.

The Global Land and Marine Observations Database aims to produce a comprehensive land based meteorological data archive and inventory. This requires the compilation of available land-based station meteorological data and information (metadata) from all known available in-situ meteorological data repositories/sources and at multiple timescales (e.g. sub-daily, daily and monthly) (Thorne et al, 2017). Observations form the foundational basis for understanding how our climate has changed and continues to change. By collecting, documenting, and curating these sources in partnership with the National Oceanic and Atmospheric Administration's National Centers for Environmental Information (NOAA/NCEI) the long-term and fail-safe availability of these meteorological data sources can be assured for future generations.

However, while compiling these data inventories it has become very clear that many data sources contain stations with demonstrably incorrect location coordinates. This is most obvious for those land-based stations which, when mapped, have coordinates that situate them over a water body. Until such issues are resolved, these stations cannot be included in the process to be served via the Copernicus Climate Change Service (CDS). Critically, many of these issues are related to stations located in regions of the world where a sparsity of climate data currently exists, such as Southeast Asia and South America. Therefore, a goal of this classroom-based exercise was to resolve these station location issues, so these important stations can be included in the CDS.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

1.1 Resolving data issues using a crowdsourced student approach

Once the station location issues have been identified it can take a considerable amount time and resources to resolve the correct geolocation. In most cases a lack of station metadata (historical station information) can hinder this task. Nevertheless, the process is relatively simple and repetitive in nature with the same sequential steps required to try to remedy the situation each time. The methods are inherently geospatially based. Taking advantage of the nature of the problem and noting the concurrent need to refresh and revamp our undergraduate program to meet new stated educational curriculum expectations (Section 5.1), we implemented a pilot project, which was rolled out to undergraduate Geography students. These students are in their 2nd year of a 3-year degree at Maynooth University in Ireland, and this project forms part of the geographical research methods class which is a mandatory module. Previously, this class had considered a range of methods-based problems, but these were based upon existing data and had no broader benefit beyond the educational outcomes for the students. The revamped citizen science-based approach far better met the stated target educational outcomes for the year as discussed further in Section 5.1.

The concept of crowd sourcing or citizen science is not new, with many global projects recruiting between them millions of citizens to help with specific labour-intensive tasks. Many of these volunteers are non-scientists, yet with appropriate guidance and instruction can help with such tasks as data transcribing, data verification or categorization as well as conducting analysis of all types of scientific data (Bonney et al., 2018). There have also been substantial efforts in climate data rescue which involves the digitizing and transcribing of recorded instrumental observations and climate data which is at risk of being damaged or lost (World Meteorological Organization, 2016). For example, in the climate research sphere the OldWeather.org (www.oldweather.org/), IEDRO.org (www.iedro.org), DataRescue@Home (www.data-rescue-at-home.org) and Weather Rescue (www.zooniverse.org/projects/edh/weather-rescue) all have ongoing projects that successfully recruit the help from citizens to rescue environmental and climate data. There has also been important work that utilized citizen science and crowd sourcing to rescue historical climate data in regions such as Africa, Europe, and Australia (Ashcroft et al., 2016; Brönnimann et al., 2018; Jacobson et al., 2018; Kaspar et al., 2015). Other projects such as the Cyclone Centre (https://www.cyclonecenter.org/) and the Climate CoLab (www.climatecolab.org/page/about) engage the help of thousands of individuals to analyse and / or verify climate data. In addition, projects like the climateprediction.net (www.climateprediction.net/getting-started/) are successfully running climate modelling experiments using the combined power of home computers of thousands of volunteers.

Crowd sourcing can also have explicit educational aims. For example, the Global Learning and Observations to Benefit the Environment (GLOBE) (https://www.globe.gov/about/overview) platform is an international science and education programme that works with citizens, students, teachers and scientists across the globe in helping them partake in data collection and the scientific process. This initiative allows the contributors to help us better understand the Earth system and global environment, while providing them with important insights into real-world research (Allen et al., 2011; Mitchell et al., 2017; Vitone et al., 2016).

Until recently there had been little effort to integrate such approaches explicitly into the tertiary education classroom. Such approaches have potential co-benefits in terms of educational outcomes but also engaging a cohort of interested students undertaking activities with expert instruction and support. Maynooth University has undertaken two substantive efforts to integrate such approaches into the classroom and assess the results via its Geography programme in recent years, as follows:

- Ryan et al., (2018) showed that with careful guidance and planning a module for university students could be developed, where students can help with important data rescue tasks. The study developed an accredited assignment for final year Geography undergraduate students at Maynooth University. The students were given the tools to successfully transcribe 1300 years of Irish daily precipitation records from scanned hard copy sheets (Ryan et al., 2018). Students also provided feedback on the module with more than 90% of students providing a positive response to all aspects of the assignment. Since that publication a further two years of the assignment have been run and across the three cohorts of students in excess of 4000 years of early daily Irish rainfall records from across the island of Ireland have been digitised in collaboration with the Irish Meteorological service Met Eireann.

- Phillips et al., (2018) assessed whether citizen science projects can be used as coursework with real practical experiential-learning benefits, without affecting the citizen science project outcomes. Two groups of university students (from Maynooth University and University of North Carolina Asheville) and volunteers were assessed on their participation in the Cyclone Centre project using a skill score metric developed by Knapp et al. (2016). The results showed that there were no substantive differences in cyclone classification between credit awarded and volunteer participants (Phillips et al., 2018). Interestingly, the study noted that students overall had a positive opinion of participating in a citizen science project and of completing such a non-traditional assignment.

Both studies noted that their work demonstrates the potential for future projects to be developed that engage university students in meaningful real-world research (Phillips et al., 2018; Ryan et al., 2018). Therefore, we have undertaken to expand the curriculum activities in this area as described in the present paper. The remaining sections of this paper are structured as follows: Section 2 presents details of identifying the station location issues; Section 3 presents details on the Geo-locate project, which includes the workflow instructions given to students; Section 4 provides details on the results of the first two years of the Geo-locate project; Section 5 gives details of pedagogical aims of the Geography department and the expected learning outcomes of the module. Section 5 also presents details of what the students gained from doing the assignment. The same section presents the results of a project feedback survey that students were asked to complete; In section 6 we describe the ongoing work and future challenges of stations location issues; Finally, section 7 presents some final comments and concluding remarks.

2 Ascertaining the magnitude of the station geolocation issues

To quantify the number of currently inventoried land-based stations that were potentially situated over a water body, mapping software tools were used. All the station location points were first mapped according to the coordinates provided in each data source. Next, using the mapping tools, all station points were overlaid on a global country boundary shape file. Stations that did not lie within the shape file land boundaries were deemed to be situated in the ocean. In total 7975 daily stations and 9144 monthly stations have been identified in the sources inventoried to date as not being situated on land through this process. These spurious station location cases arose from a broad range of primary data sources. They were next checked to see if they could be identified as actual buoys, platforms or ships. If they were not, then they were extracted from the inventory for further consideration.

We present a map in Figure 1 that shows the daily stations (denoted by blue dots) and the monthly stations (denoted by red dots) which have been identified as having location issues. These stations are classified as land-based stations, are located over a water body, with most stations just lying off the coast, which indicates a geocoordinate precision issue. To give a sense of the scale of location problems, Figure 2 presents a histogram of the number of stations (daily and monthly) and the distance from land in meters (m). Figure 2 shows that 10,256 stations located the ocean are within 1000m of land, 1,259 are within 4000m, 876 are within 7000m, 423 stations within 10,000m, 255 stations within 13,000m, 146 stations within 16,000m and the balance of 66 are within 22,000m of land. There were 2,838 stations that could not be mapped due to missing latitude or longitude coordinates. The results in Figure 2 show that most stations lay within 1000m (one kilometre) of land which suggests that they are either lighthouses, platforms, buoys or that they are land stations with coordinate precision issues.

The stations in the inventory that have been found to have location issues so far arise from several data sources. The daily stations with location issues are from 58 different data sources, the monthly station with location issues came from 41 different data sources. Table S1 in the supplementary-1 material presents a list of all data sources and the number of stations identified with location issues. The most daily station location issues were identified in NCDC's Global Summary of the Day (GSOD) Product (2506 stations) with the Global Historical Climate Network Daily (GHCND) dataset identified as having the second largest amount of station location issues (1649 stations). The balance of 56 daily data sources contain an average of 68 stations with location issues per source ranging from 1 to 881 stations per source. At the monthly timestep in Table S2 also in the supplementary-1 material shows that the UK Met Office data source has the most stations location issues with 2511 stations identified and the Monthly Climatic Data of the World (MCDW) data source has the second most with 844 stations location issues identified. The balance of the 39 monthly data sources contain an average 148 stations per source ranging from 1 to 784 stations per source.

Other issues which have come to light involve station clustering along the prime meridian which suggest station geolocation issues. These stations were also extracted from the inventory but are not considered further here. Undoubtedly there are additional station coordinate issues that lead to incorrect placement of sites, but over land. Future checks are planned using for example comparisons between target stations and apparent neighbouring stations or comparisons to reanalysis products but are outside the scope of the present analysis. Such comparisons should highlight stations that are grossly mis-located based upon both phase and amplitude of annual cycles and synoptic features.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

3 Working with Geography students at Maynooth University

We developed the Geo-locate project so that we could enrol the help of 2nd year undergraduate Geography students at Maynooth University, to resolve some of the land-based station geolocation issues identified. The pilot first round of the Geo-locate project was run in the second semester of the 2017/18 academic year. For this pilot we began with 880 daily-resolution stations identified with location issues. It was decided that three different students would attempt to resolve each station to try and attain triple verification of the revised location. We produced 88 excel sheets with 10 stations in each and divided the 264 students into three groups and each group of 88 was allocated the same 88 excel sheets one per student.

Each of the stations were supplied with associated geographic coordinate information, which was known to be incorrect as it placed the station over water. All the stations that students worked with are reporting as land-based stations. These stations will, in general, be over water for some combination of resolvable issues. These could include imprecise geographic coordinate information, incorrect conversion from degrees, minutes, seconds to decimal degrees, dropped minus signs placing in the incorrect hemisphere (N-S or E-W), or simply missing coordinates altogether.

Students were tasked with carrying out a sequential set of tasks to gather evidence to support the relocation of each of the stations they were assigned to their correct geographic coordinates on land. Students were required to find and provide all available evidence to support their conclusions as to why a station's coordinates should be as they indicated. They needed to employ a variety of research tools, including Google Earth, Google Maps, web searches, recourse to dedicated climate data information sources and a variety of additional research tools and information to determine the correct locations of their allocated stations.

Initially, students were tasked with mapping the stations by importing the station data into Google Earth and to visualise the current stations locations. For many of the existing station locations, students were be able to determine/narrow down the research focus on the basis of viewing the station locations relative to the surrounding land and the named features on the map which may correspond to the station name.

The students were provided with step by step instructions and guidance on how best to resolve the station location issues. A copy of the student handout sheet which describes the guidance and steps to complete the assignment is available in the supplementary-2 material of this paper. Figure 3 shows a summary workflow of the guidance and steps that students were asked to follow. Step 1 involved students firstly checking the World Meteorological Organization (WMO) Observing Systems Capability Analysis and Review Tool (OSCAR) (https://oscar.wmo.int/surface//index.html#/) to see if the station name exists in the database. The OSCAR database is the official metadata repository of the WMO Integrated Global Observing System (WIGOS) for all surface-based observing stations and platforms. For more information see (https://www.wmo.int/pages/prog/www/wigos/index_en.html).

If the station name was not found in the OSCAR database students were asked to enter details of their search in their allocated student sheet and move on to the next step. Step 2 involves students checking to see if the station is in any of the National Meteorological agency station information sheets which are provided. If the station is not in any of the information sheets the students were again asked to comment and then move on to step 3. In step 3 the guidance suggests that students conduct a web search combining the station name plus some key words such as "weather station" and "latitude longitude" to try and find any relevant information. If the station name cannot be found using any of the steps, then students were to comment that this station could not be found, and state details of each step taken. If the station revised location has been found, then students were requested to proceed to the evaluation step. As an extra step, students are required to try to verify the coordinates using alternative sources. For example, even if the WMO OSCAR data base contains geographic coordinates for a station, students are asked to 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 information files that have been provided). A snapshot of an example completed student sheet is presented in Figure 4 and shows the details of each step undertaken and the outcome.

During the project, teaching staff (consisting of faculty, post-docs and postgraduate students) provided ongoing support to the second-year students including regular scheduled workshops and question and answer sessions via an online forum and developed short video tutorials for each of the steps outlined above that students could access via an online e-learning environment. The overall aims and goals of the Global Land and Marine Observations Database activity were also delivered via an introductory lecture by the lead author of the present study. In addition, Dick Dee, the deputy head of the Copernicus Climate Change Service contributed an introductory video piece outlining the importance of the students work. The video was shown to students during the introductory lecture to help motivate the students and make them aware of the wider importance of the project to the scientific community. A copy of the Dick Dee introductory video is available at (https://doi.org/10.5446/41783).

The assignment deliverables and subsequent project marks (worth 50% of the overall module) were based on the students completing the following:

1) A spreadsheet with the station list, original coordinates and new, corrected coordinates. The spreadsheet also requires the student to detail how they obtained the updated coordinates and to add comments briefly outlining the sources for the new coordinate information and their justification. An example spreadsheet template is provided in the supplementary-3 material of this paper. Marks for the completed station .xls file were based on the number of stations completed with full details/comments/supporting information (**35%**). However, students were not penalised marks if they were unable to find the correct location, once they provide full details of all the steps conducted and include a full traceable account.
2) A group presentation detailing the research methodology that students undertook to identify and correct each station's geographic coordinates. The presentation should have contained an overview of the arguments to support the relocation of each station to its new location (**15%**).

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

4 Results of the pilot Geo-locate project

The Geo-locate project has now been run over academic years 2017/18 and 2018/19. The results are discussed in the following two sub-sections. In particular, lessons learnt from the pilot project in 2017/18 were applied in the following year, achieving both greater levels of output and an improved learning experience. In both years a substantial number of geolocation issues were resolved. The resolved locations and the metadata trail of the decisions made will be captured and used in the C3S Global Land and Marine Observations Database and at NOAA's National Centers for Environmental Information.

4.1 Pilot project

For the pilot project students attempted to resolve location issues at 811 stations. There were some initial problems with the distribution of the station sheets to the students, so not all stations were attempted by three different students (triple verification). The results show that 79 stations (10%) were attempted by one student, 310 stations (38%) were attempted by two students. There were 422 stations (52%) attempted by at least three students, and of these 38 stations were attempted by four students.

The updated geo-coordinates for all stations with single attempts required further checks by a service team member as a matter of course. Due to a lack of consistency between independent student assessments many of the other station updated locations were also checked by a service team member. These additional checks involved using mapping tools to map the station updated locations to visually check the validity of the revision. In addition, the distance between updated coordinates and original coordinates was assessed. Any updated station location greater than approximately 33 kilometres (.3°) from the original station location was also checked. Google earth was also used to zoom into the updated station location to verify the revision. The student comments were also read to make sure they made sense and that they had provided enough evidence to verify the updated location.

A service team member had to check and verify 249 stations locations for the pilot project, and as a result, only 77 station geo-coordinates had to be updated due to errors by students. In other words, less than 10% of the 811 stations attempted by students had to be updated to the correct geo-coordinates by a service team member, which builds confidence in the efficacy of students to undertake the project. It must be noted that these extra checks were much faster than trying to resolve the original station location issues from scratch. On completion of all the extra checks 794 station location issues were resolved (98%) of the 811 stations attempted. By a reasonable estimate getting these done from scratch by service team members would have taken of the order 1-2 hours per station equating to 4-5 person months of effort. The Chinese proverb 'many hands make light work' applies, in that by spreading the task across many individuals the workload on any one individual becomes much less onerous.

The majority of the stations for the pilot project are situated in countries with sparse meteorological data coverage where resolution of individual issues has the greatest value to climate service users. Solving geolocation issues in data rich regions provides an incremental improvement, whereas in a data sparse or data void region this is a substantial advance. The 811 stations attempted in the pilot project derived from 11 data sources (original data provider) with a varying number of stations from each source and records at these stations spanning 1849-2017.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

Figure 5 shows a map of stations located in Java and parts of Indonesia that had geolocation issues, the blue dot represents the original locations and the red dot denotes the updated revised locations. The similar information is shown in Figure 6 for Malaysia, Sumatra, Kalimantan, Sulawesi and parts of the Philippines, Thailand, Vietnam and Cambodia; Figure 7 for Northern Australian stations; Figure 8 for stations located in Mexico.

The issue with many of these stations located in Australia and parts of Southeast Asia appears to have been the lack of precision in the original latitude/longitude coordinates, which meant many stations had been located off the coast. Other issues existed such as the coordinates were not converted from the original degrees, minutes and seconds to decimals correctly, or even not at all. Another common error was the latitude was entered as longitude and vice versa. The stations located in Mexico had no original station geo-coordinates, but when the station names were verified that they matched up with the city or town of the new location and made sense, they also were recoverable.

4.2 Results of round two of the Geo-locate project

Based on the what was learned from the pilot module it was decided that it would be acceptable for each station to be attempted by two different students, as there is a requirement for extra checks by a service team member regardless, as outlined in section 3.1. It was also decided to give the students 15 stations per sheet in round two, an increase of 5 stations per sheet from the pilot project. In the second round of the project we were also able to supply more current global National Meteorological Service station information sheets. In providing more station information sheets we would expect that some of the station location issues will be resolved much quicker as they contain correct land-based station location coordinates. It was also important to ensure that the station sheets were distributed correctly to the students so each of the stations were attempted by two students. The sheets were compiled and distributed to the student's groups using the same methods as the pilot scheme outlined in section 2. We divided the total number students into two groups and each group was allocated a duplicate set of excel sheets, one sheet per student.

For round two 100 sheets containing 15 stations (1500 stations) with location issues were produced for students. The 1500 stations derived from 33 original sources with a global spatial extent and spanned from 1797 to 2017. There were 198 students registered for the module and 181 completed station sheets were returned, which relates to 1357 stations. Of these there were 18 station sheets that could not be processed due to file corruption and / or not being correctly completed. The 163 completed station sheets were merged together, stations were sorted by name and checked by a service team member to verify that the revised coordinates were correctly entered. The revised station locations were checked using the same pilot project methods described in section 3.1. The results showed that there were 1222 stations attempted in round two of the project. There were 1170 stations (95%) attempted by two students, 30 stations (2%) had revised locations but had only been attempted by one student, and for 22 stations (2%) location issues could not be resolved. Results show that of the 1222 stations attempted 91 (7%) were found to be marine stations such as lighthouses, buoys or ships. A service team member had to conduct extra checks on 402 stations (33%) due to a lack of consistency in the student revised station location information. In total, round two of the Geo-locate project had resolved 1132 unique land-based station location issues and verified that a further 91 stations were in fact marine based stations.

Like the pilot project the issues with the coordinates in round two appear to be mainly due to poor coordinate precision with most stations incorrectly located just off the coast. The coordinate precision issue meant that many of the stations which should have been located on small islands across Canada, Alaska, Northern Europe, the United States and Japan were incorrectly located in the ocean. In addition, many times the station name was spelled incorrectly, which was also rectified when identified. Figure 9 shows a map of stations located in their original wrong location (denoted by blue dots) and the revised location (denoted by red dots) for Japan. Figure 10 shows the original and revised location of stations in Northern Europe and Figure 11 shows the stations located in the eastern region of Canada. It must be noted that some of the stations in Canada and Northern Europe were found to be actual buoys. Also, some stations in the Gulf of Mexico were found to be static marine platforms and others located around the coastline of different countries were found to be lighthouses.

The second iteration was slightly more successful than the pilot project, with an increase in the number of stations location errors being resolved by fewer students. In the pilot project the students resolved 794 station location issues with 264 participating. However, project two resolved 1132 land-based stations and 91 marine stations with 181 students which is 83 students fewer than the pilot project. In addition, it took service team members only 3 days to collate and check the revised stations in round two, whereas the pilot project stations took over a week to sort and check. These results indicate that project two was more efficient as measured by scientific outputs.

5 Pedagogical aims and learning outcomes, student experience and feedback

The following sections presents the overall pedagogical aims of the Maynooth University Geography Department for 2nd undergraduate students and the specific expected student learning outcomes for the 2nd year Methods of Geographical Analysis (GY202) module. A formal student feedback survey was also implemented in round two of the Geo-locate project to gain some more quantitative insights into what students thought of the assignment and to hear some suggestions on how we could improve the assignment. The results of the survey are presented in section 5.2.

5.1 Geography department Pedagogical aims and module student learning outcomes

The following statement is taken from Department of Geography 2nd year student handbook and sets out the newly revised pedagogical aims of the department's teaching in that year of the program which this assignment partially aimed to fulfil:

> "The focus of this second year of the Geography undergraduate programme is on Methods and the Systematic Branches of the Discipline. Students are introduced to different systematic branches of Geography and learn that within both human and physical Geography there have emerged distinctive sub-areas with their own concerns and trajectories. In parallel, year 2 foregrounds the teaching of basic

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

> research methods. Students learn to work as individuals and as part of teams, and in the laboratory and in the field, to identify, source, collect and analyse primary and secondary data, and to evaluate and present research results and findings. In addition, students are provided with the opportunity of applying the research skills acquired in year 2 through field work in Ireland and overseas. All students will also learn the basics of GIS." (van Egeraat, 2018)

The specific expected student learning outcomes of the 2nd Year Methods of Geographical Analysis (GY202) module, of which the Geo-locate project was part, are as follows.

> On successful completion of the module, students should be able to:
> - Develop further data collection, processing, computer and presentation skills, based on work in First Year and in GY201
> - Learn the skills required for work in Second- and Third-Year Geography
> - Develop group working and co-operation skills
> - Gain basic experience of research methodology, useful in many areas of employment
> - Apply theoretical learning in practical situations
> - Relate theoretical learning to a local environment

The Geo-locate project was designed to meet the Geography department pedagogical aims and the module student learning outcomes. In particular, the project encouraged students to use several research methods new to them, working both as an individual and as part of a team. In addition, students had to explore varying online investigative skills to try and learn how to access, collect, compare and present different sources of information and data to resolve the station location issues. The project was designed to help students develop reasoning skills and allow students to gain computer and presentation experience. Students also used Geographical Information Systems (GIS) in the form of Google Earth mapping tools. Overall, the expectation was thus that the assignment should provide them with improved spatial awareness and a better understanding of potential issues with real-world data which they may well work with in their future careers. The Geo-locate project allowed the students to work with real data issues, be part of, and contribute to, a real-world climate data project. The Geo-locate project thus plays a substantive role in delivering the second year and module-specific pedagogical outcomes.

5.2 Results of student feedback survey

The survey was completely anonymous to ensure that the students could express their true opinions. A copy of the feedback survey sheet given to students is provided in the supplementary-4 information of this paper.
The student feedback survey was made available online and 152 students from the 2018-19 student cohort participated in reviewing the second iteration of the project. The questions 1a to 1k of the survey asked the students to indicate the extent to which they agree or disagree with specific statements about the project. Table 1 presents each of the questions and the subsequent results. Overall positive responses were received from students as follows:

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

- 74% agreed and 20% strongly agreed that they had gained important insights into data issues while only 5% disagreed and fewer than 1% of students strongly disagreed.
- 67% agreed and 14% strongly agreed that the supports in place were sufficient to aid them with completion of the assignment. There were 14% of students who disagreed and only 5% who strongly disagreed.
- 60% of students agreed and 22% strongly agreed that the guidance given for the project was clear and easy to follow. There were 16% of students that disagreed and only 2% that strongly disagreed.
- 59% of students agreed and 11% strongly agreed that they gained insight into citizen science, while 28% of students disagreed and only 2% strongly disagreed.
- Students were more divided on whether they would prefer further such assignments to more traditional assessment approaches with 37% of students stating that they agreed and 16% strongly agreeing, but 34% disagreeing and 13% strongly disagreeing.
- Most students indicated that the work load was appropriate for the level of credit, with 63% students agreeing and 17% strongly agreeing while 15% disagreed and only 5% strongly disagreed.
- 57% of students agreed and 14% strongly agreed that the assignment was a valuable learning experience, although 25% of students disagreed and a further 5% strongly disagreed.
- Fewer than 50% of students were more motivated than usual in doing this assignment with 38% of students agreeing and 10% strongly agreeing while 41% of students disagreed and 11% strongly disagreed.
- Most students felt that they had made a worthwhile contribution to an important global project. with 62% agreeing and 9% strongly agreeing. In contrast, 25% disagreed and only 3% of students strongly disagreed with the statement. Note that one student omitted to answer this question.
- The majority of students thought that subsequent cohorts of students would be happy doing a similar assignment with 52% of students agreeing and 15% strongly agreeing as opposed to 25% who disagreed and 9% that strongly disagreed.
- Most students felt like they gained some useful transferrable skills from the assignment as outlined in section 5.1. There were 65% that agreed and 20% that strongly agreed while 13% disagreed and only 3% strongly disagreed.

Question 2 (a-g) of the survey asked students to give a score from 1-10 to a list of items, indicating how important each item was in enabling them to successfully complete the assignment. (1 is not important and 10 is very important). There were 152 students that responded to question 2 (a-g). Table 2 presents all the responses to question 2 as a percentage of the total responses. The following are ordered by what the student's perceived to be of most importance based on their responses:

1. The students felt that the clear assignment guidelines were the most important aspect with over 96% of the responses between 6 and 10 and over 77% of responses between 8 and 10.
2. The in-class support ranked second as being most important with students with 93% of responses between 6 and 10 and 75% of responses between 8 and 10.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

3. Most students felt that the lecturer's enthusiasm was third most important and enabled them to successfully complete the assignment. Over 91% of the student's responses were between 6 and 10, and over 72% of the responses were between 8 and 10 for this aspect.
4. Online support for students was the next most important with 68% of responses between 6 and 10 and 45% of responses between 8 and 10.
5. The fact that students knew that they were contributing to a real-world global project ranked as important, with over 67% of responses between 6-10 and over 38% of responses between 8 and 10.
6. The introductory video from Dick Dee of ECMWF was the sixth most important with 43% of responses between 6 and 10 and 16% of responses between 8 and 10.

Question 3 of the survey asked students to indicate three aspects of this assignment that worked well, and 130 students responded. The results were analysed, and some common themes were identified. Over 70 students stated that the supports and guidance that was provided to aid them in completing the assignment worked well. In addition, 51 students stated that they had gained some useful research methods and transferable skills, as presented in section 5.1, from doing this assignment. The students also felt that working in teams was very useful and shared the workload, with 32 students making this statement. However, only 20 students stated that the time given to complete the assignment was adequate.

Question 4 of the survey asked students to indicate three aspects of this assignment that could be improved, and 115 students responded. Again, some common themes were extracted from the responses. In relation to the time allocated to complete the assignment 34 students felt that there was not enough time, felt a bit under pressure and suggested a reduction in the number of stations given to each individual to resolve. There were 29 students who mentioned that they would like clearer guidance and instructions how to use the online resources such the Oscar Surface web tool. There were 26 students who felt that more station list resources and potential online sources to find the stations should be made available to them. In addition, nearly 10% of students said that clear instructions on what to do when a station could not be found should be outlined in the handouts.

The final question of the survey asked students to add any other thoughts/ comments they had about the continuous assessment. Only 37 out of the 152 students responded to this question. There were 21 students who responded with negative comments and 12 with positive comments towards the assignment, the remaining 4 student comments were general. Some of negative comments stated that the assignment was too stressful, time consuming, difficult, boring, frustrating and not enough support was provided. Examples of some of the positive comments are that the assignment was enjoyable, extremely beneficial, rewarding and real. Interestingly, these results show that a minority of students did not enjoy this assignment and decided to express this in the open-ended question rather than respond negatively to questions 1 and 2. However, despite this the overall evidence presented in this section indicates that the project was well received by the students with most of them engaging fully in the process.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

6 Future plans

Currently as part of the C3S Global Land and Marine Observations Database we have inventoried 23,619 sub-daily stations derived from 51 sources, 173,782 daily stations from 137 sources and 85,186 monthly stations from 55 sources. In addition, new sources of data are being acquired all the time which means that the potential issues with resolving station locations may be an ongoing challenge. For example, we are working closely with the C3S Data Rescue Service to ensure all rescued climate data is deposited via the new data discovery and deposition web-based service which we are developing (Noone et al., in prep). Work is also ongoing in collaboration with the European Environmental Agency (EEA) in its capacity as the Copernicus in-situ lead, and with ECMWF in its capacity as the entrusted entity for C3S. Additional data inputs for the database may also be secured based on the recently enacted EEA-EUMETNET (EUMETNET is a grouping of 31 European National Meteorological Services) agreement on data sharing. We have identified several thousand stations across the existing secured sources that require checking and it is all but certain that newly acquired sources will also contain geolocation issues. Therefore, there is likely to be no issue with running the Geo-locate assignment for years to come.

As alluded to in the introduction, as well as stations incorrectly located over water which are easy to identify, if not to rectify, many could be incorrectly located on land. To identify such land-based locational outliers, we plan to develop a suite of data quality control checking tools. For example, pairwise homogeneity assessment could be used to identify any irregularities in data when compared with data from other stations within a given distance of each other (Dunn et al., 2014; Durre et al., 2008; Menne et al., 2012). This process will be automated as much as possible but there will be a need for visual checks. Stations identified via these approaches could constitute an additional valuable source of data for future iterations of the Geo-locate project providing a greater variety of issues for students and additional tools such as data comparison tools which may enable a more nuanced assessment in future as well as improving learning outcomes for the students.

For example, reanalysis data provides data sets at regular intervals over long periods of time for climate monitoring and research (e.g. Dee et al., 2011) which is produced via data assimilation using a frozen version of a given forecast system. The C3S reanalysis data contains estimates of atmospheric variables such as air temperature, pressure and wind at varying altitudes. Reanalysis also contains surface variables such as rainfall, soil moisture content, and sea-surface temperature. ECMWF reanalysis products provide estimates for all locations on earth, and ERA5 will shortly extend back to 1950, for more information see (https://www.ecmwf.int/en/research/climate-reanalysis). To address the location issues over land, and also to confirm relocation of marine stations, in future project iterations we are looking into the potential to compare the station data with the reanalysis data at the same location or plausible alternative locations provided by the students to identify likely differences due to incorrect station locations. This shall require development of underlying software and a web-based interface to enable the analysis but would add considerable flexibility and data analysis aspects to future assignments.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

7 Conclusions

The Geo-locate project which worked with 2nd year undergraduate Geography students has been successful both in terms of educational outcomes and resulting geolocation issues resolution, with 1926 land-based stations with location issues in the original sources resolved. In addition, the students identified 91 marine stations. This is a significant result as these stations can now be included in the inventory to be assessed for inclusion in the Copernicus Climate Data Store (CDS). Such a result would have taken many person months, if not person years, of service team members effort to achieve and would not have benefitted from multiple independent assessments. Many of these stations are situated in regions where there are sparse observations and the inclusion of these stations in the CDS will allow for a more robust climate assessment in the future. An updated list of all these stations will be made available through the service as metadata, which will also include all the student comments and notes.

The results of the student feedback survey are generally very positive and indicate that most students gained some of the useful transferrable skills outlined in section 5.1 and felt like they were involved in a meaningful real-world project. In addition, overall the students felt that the supports and guidance given were sufficient in helping them complete this assignment. However, we will be reading over all the students comments and suggestions (positive and negative) and will continue to evolve the project to ensure optimal educational outcomes. Based upon the successful educational outcomes and data problem resolutions attained in the first two rounds of the Geo-locate project, we aim to continue the project for many years to come. Finally, we would encourage other organisations to investigate the potential for engaging university students to help resolve similar data issues. Likewise, students can provide assistance with other projects where labour-intensive tasks exist, and the students can gain useful research skills and have the opportunity to work with real data.

Acknowledgements

This project could not have been completed without the help of the Geography undergraduate students. The students who participated are listed in alphabetical order in Supplementary-5 Information of this paper.

8 References

Allan, R., P. Brohan, G. P. Compo, R. Stone, J. Luterbacher, and S. Brönnimann.: The International Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative. Bull. Amer. Meteor Soc., 92, 1421–1425, https://doi.org/10.1175/2011BAMS3218.1, 2011.

Ashcroft, L., R. Allan, H. Bridgman, J. Gergis, C. Pudmenzky, and K. Thornton.: Current climate data rescue activities in Australia. Adv. Atmos. Sci., 33(12), 1323–1324, doi: 10.1007/s00376-016-6189-5 ,2016.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J., and Parrish, J.K.: Next Steps for Citizen Science., 2018 http://science.sciencemag.org/content/sci/343/6178/1436.full.pdf [Accessed August 29, 2018]

Brohan, P. (2016) Free at last. Available at: https://blog.oldweather.org/2016/06/29/free-at-last/ [Accessed 9th May 2019]

Brönnimann, S., Brugnara, Y., Allan, R.J., Brunet, M., Compo, G.P., Crouthamel, R.I., Jones, P.J., Jourdain, S., Luterbacher, J., Seigmund, P., Valente, M.A., Wilkinson., C.W.: A roadmap to climate data rescue services. Geosci Data J.; 5: 28–39. https://doi.org/10.1002/gdj3.56 , 2018.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P. , Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. , Park, B., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J., Vitart, F. : The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteorol. Soc., 137: 553-597. doi:10.1002/qj.828, 2011.

Dunn, R.J.H, Willett, K.M., Morice, C.P., Parker, D.: Pairwise homogeneity assessment of HadISD. Climate of the Past. 10. 1501-1522. 10.5194/cp-10-1501-2014.

Durre, I., Menne, M.J., Vose, R.S.: Strategies for evaluating quality assurance procedures. J. Appl. Meteor. Climatol., 47, 1785–1791, 2008.

Jacobsen, K., Van Hirtum, L., Amara, M., Beeckman, H., Meeus, S., Vandelook, F., Stoffelen, P., Verbeeck, and Hutkens, K.: Climate data rescue from the Belgian colonial archives: helping to close the data-gap over Central Africa. Early Instrumental Meteorological Series, Conference abstracts, 2018.

Kaspar, F., Tinz, B., Mächel, H., and Gates, L.: Data rescue of national and international meteorological observations at Deutscher Wetterdienst, Adv. Sci. Res., 12, 57-61, https://doi.org/10.5194/asr-12-57-2015 , 2015

Knapp, K.R., Matthews, J.L., Kossin, J.P., Hennon, C.C.: Identification of Tropical Cyclone Storm Types Using Crowdsourcing. Mon. Wea. Rev., 144, 3783–3798, https://doi.org/10.1175/MWR-D-16-0022.1 , 2016.

Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E., Houston, T.G.: An Overview of the Global Historical Climatology Network-Daily Database. J. Atmos. Oceanic Technol., 29, 897–910, https://doi.org/10.1175/JTECH-D-11-00103.1, 2012.

Mitchell, N., M. Triska., A. Liberatore, L., Ashcroft, R., Weatherill, and N. Longnecker.: Benefits and challenges of incorporating citizen science into university education. PLOS ONE, 12, e0186285, https://doi.org/10.1371/journal.pone.0186285 , 2017

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

Noone, S., Berry, D., Dunn, R., Perez-Gonzalez., Kettle, A., Menne, M., Stephens, A., Thorne, P., Tucker, W., Voces, C., Willet, K.: Call for data contributions: A truly integrated global collection of historical surface meteorological observations. In preparation

Phillips, C., Walsh, D., O'Regan, K., Strong, K., Hennon, C., Knapp, K., Murphy, C., Thorne, P.W.: Assessing Citizen Science Participation Skill for Altruism or University Course Credit: A Case Study Analysis Using Cyclone Center. Citizen Science: Theory and Practice. 3(1), p.6. DOI: http://doi.org/10.5334/cstp.111, 2018.

Ryan, C., Duffy, C., Broderick, C., Thorne, P.W., Curley, M., Walsh, S., Daly, C., Treanor, M., and Murphy, C.: Integrating data rescue into the classroom. Bull. Amer. Meteor. Soc., 0, https://doi.org/10.1175/BAMS-D-17-0147.1, 2018.

Thorne, P.W., Allan, R., Ashcroft, L., Brohan, P., Dunn, R., Menne, M., Pearce, P., Picas, J., Willett, K., Benoy, M., Bronnimann, S., Canziani, P., Coll, J., Crouthamel, R., Compo, G., Cuppett, D., Curley, M., Duffy, C., Gillespie, I., Guijarro, J., Jourdain, S., Kent, E., Kubota, H., Legg, T., Li, Q., Matsumoto, J., Murphy, C., Rayner, N., Rennie, J., Rustemeier, E., Slivinski, L., Slonosky, V., Squintu, A., Tinz, B., Valente, M., Walsh, S., Wang, X., Westcott, N., Wood, K., Woodruff, S., Worley, S.: Towards an integrated set of surface meteorological observations for climate science and applications. Bull Am Meteorol Soc 98:2689–2702, https://doi.org/10.1175/BAMS-D-16-0165.1, 2018.

van Eggeraat, C.: Second Year Geography 2018/19 handbook, available at: https://www.maynoothuniversity.ie/sites/default/files/assets/document/2nd%20year%20guide%2018-19_0.pdf, 2018.

Vitone, T., K. A. Stofer, M. S. Steininger, J. Hulcr, R. Dunn, and A. Lucky.: School of ants goes to college: Integrating citizen science into the general education classroom increases engagement with science. J. Sci. Commun., 15, 1–24, 2016.

World Meteorological Organization: Guidelines on Best Practices for Climate Data Rescue, WMO-No. 1182, WMO, Geneva, https://library.wmo.int/doc_num.php?explnum_id=3318, 2016.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

Figures, tables and captions

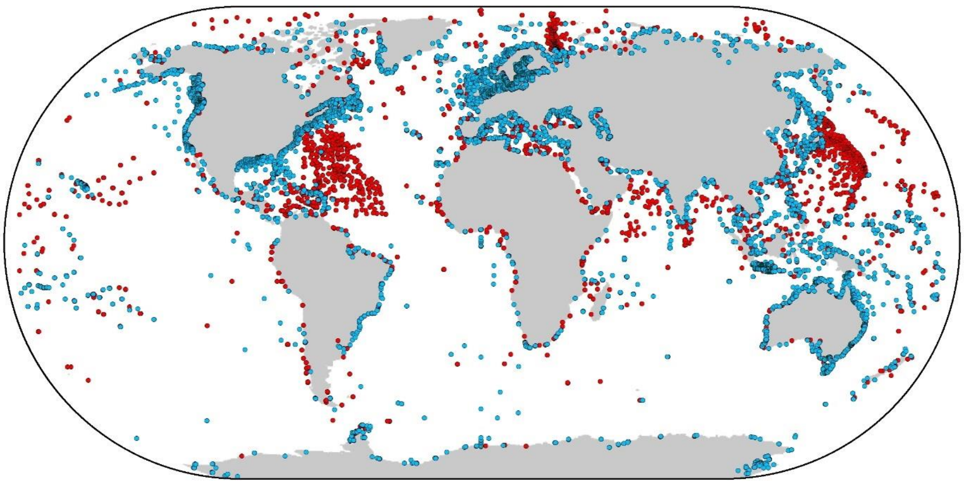

Figure 1 Map location of the stations identified with location issues (blue dots denote daily stations and red dots denote monthly stations).

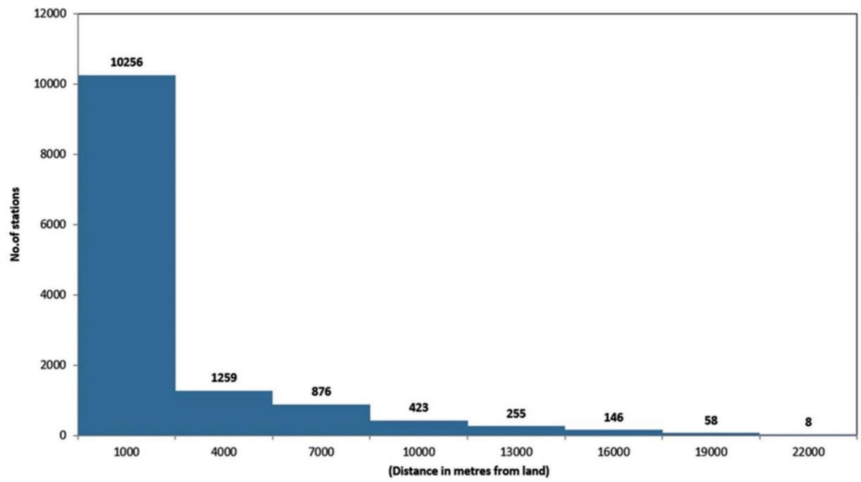

Figure 2 Histogram of the number of stations in each bin based on the distance in meters of each station from land.

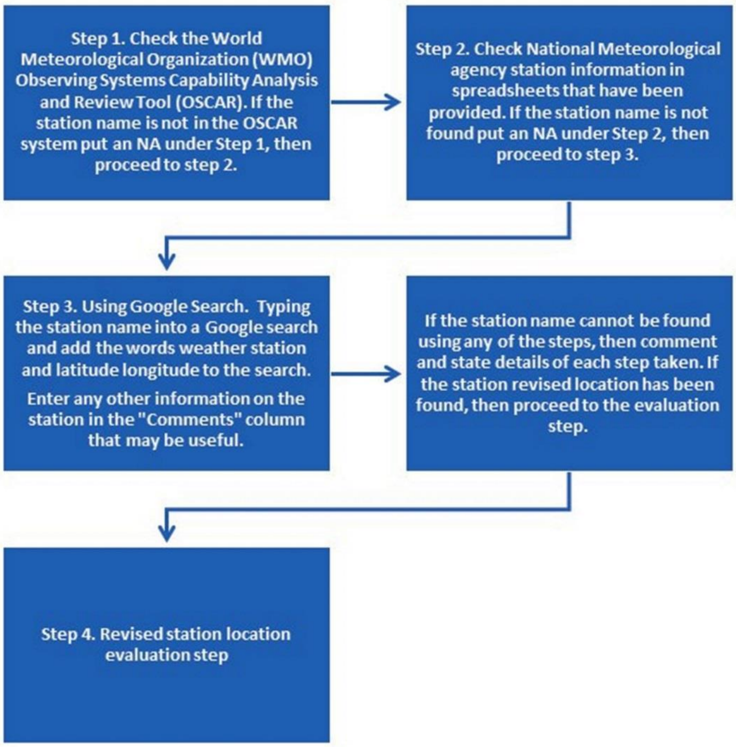

Figure 3 Workflow summary of the guidance and steps required to complete the assignment.

REVISED_LATITUDE	REVISED_LONGITUDE	REVISED_ELEVATION _METRES_ASL	REVISED_ COUNTRY	REVISED_ INFORMATION _SOURCE	STEP 1	STEP 2	STEP3	COMMENTS	Web address	Verified Results
38.02916667	138.24	5.5	Japan	OSCAR	Navigated to the OSCAR webpage and entered my station name in to the dropdown menu. This station is on the OSCAR system.	NA	NA	Found the weather station on the OSCAR system. It is located on an island off the coast of Japan	https://oscar.wmo.int/surface//index.html#/search/station/stationReportDetails/7185	In order to verify my results I navigated to the Japan Meteorological Agency however, I was unable to loacate coordinates for this station on their site. I then located the station on the global historical climatological network list to try and verify my results by that. However, the logitude and latitude provided by the list was not the coordinates of my revised location nor were they the coordinates of the original location although they were closer to the revised coordinates and the elevation was also much higher than that of the original and revised elevation. Therefore, I was unable to verify my results.
39.07372	26.887984	3	Turkey	OSCAR	Navigated to the OSCAR webpage and entered my station name in to the dropdown menu. This station is on the OSCAR system.	NA	NA	Found the weather station on the OSCAR system. It is located in the country of Turkey	https://oscar.wmo.int/surface//index.html#/search/station/stationReportDetails/4000	In order to verify my results I navigated to the Turkish Ministry of Agriculture: General Directorate of Meteorology which can be accessed via the following URL https://mgm.gov.tr/. Under the subtitles 'predictions' then 'city and town centre estimates' I searched for my station 'Dikili' and there was only one station of that name in Turkey. The latitude, longitude and altitude of the station matched those of my revised coordinates excecpt that they had been rounded to four and three decimal places respectively. Therefore, I verified my results.
1.1333	127.5	16	Indonesia	Gladstonefamily Weather	Navigated to the OSCAR webpage and entered my station names in to the dropdown menu. The station could not be located on the OSCAR system. I also entered different variations of the name incase of spelling mistakes and there was still no data on the station.	I then inserted the coordinates into google maps and established that the station was loacted on Indonesia. I then navigated to both the global historical climatological network list and the GHCNM list and was unable to locate my station on any of these. I also attempted to locate the station on the Indonesian Meteorology, Climatology and Geophysical Agency, but was unable to locate it there either.	Upon the failure of locating the station through any meteorological website I then preformed a google search of the station name and established that the station was most likely located in an airport by the name of Jailolo/Kuripasai. Through this google seach I found my revised coordinates for this station on the Gladstonefamily Weather website. The link to this was a result of my google search.	This station posed difficulties as I was unable to find it on OSCAR or any of the station lists provided to us. I was also unable to find it on the Indonesian Meteorology, Climatology and Geophysical Agency as the website would not open on my home computer or the several computers on campus which I tried to open it on. I am unsure as to why this is so it was partiularly difficult to find coordinates. I am unsure as to the validity of the website from which I took my revised coordinates from, however, they do place the station on land and it is relatively close to the original coordinates.	https://webcache.googleusercontent.com/search?q=cache:d6_8ftHvvh8J:https://weather.gladstonefamily.net/site/97428+&cd=1&hl=en&ct=clnk&gl=ie	I was unable to verify my results for this station as I could not find the longitude and latitude for this station on any other site or lists other than those that were used as part of my steps to establishing the revised coordinates.

Figure 4 Snapshot of the example student sheet with details of each steps taken to resolve the station location issues.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

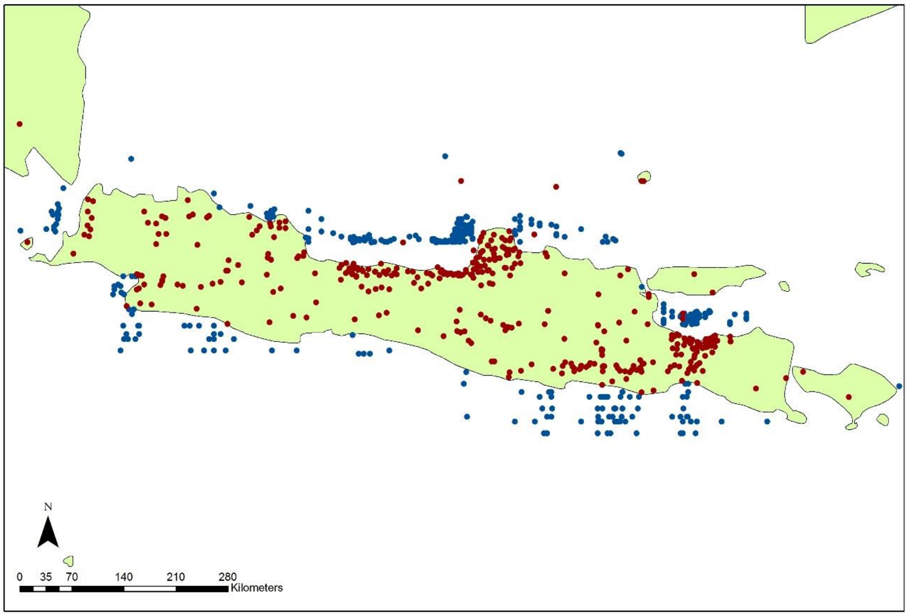

Figure 5 Map of the original station location situated in the ocean (blue dots) and the updated resolved station location (red dots) for Java and parts of Indonesia.

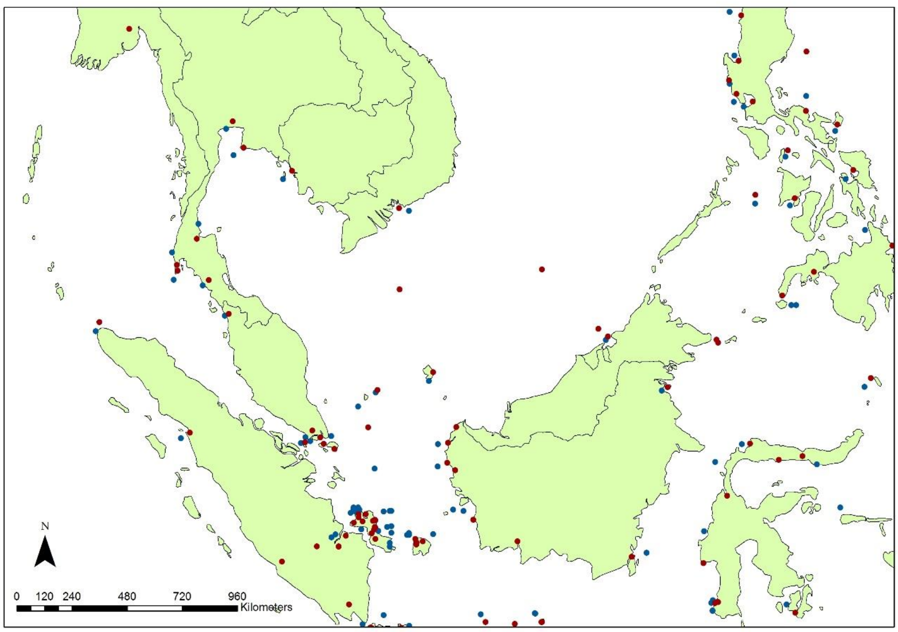

Figure 6 Map of the original station location situated in the ocean (blue dots) and the updated resolved station location (red dots) for Malaysia, Sumatra, Kalimantan, Sulawesi and parts of the Philippines, Thailand, Vietnam and Cambodia.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

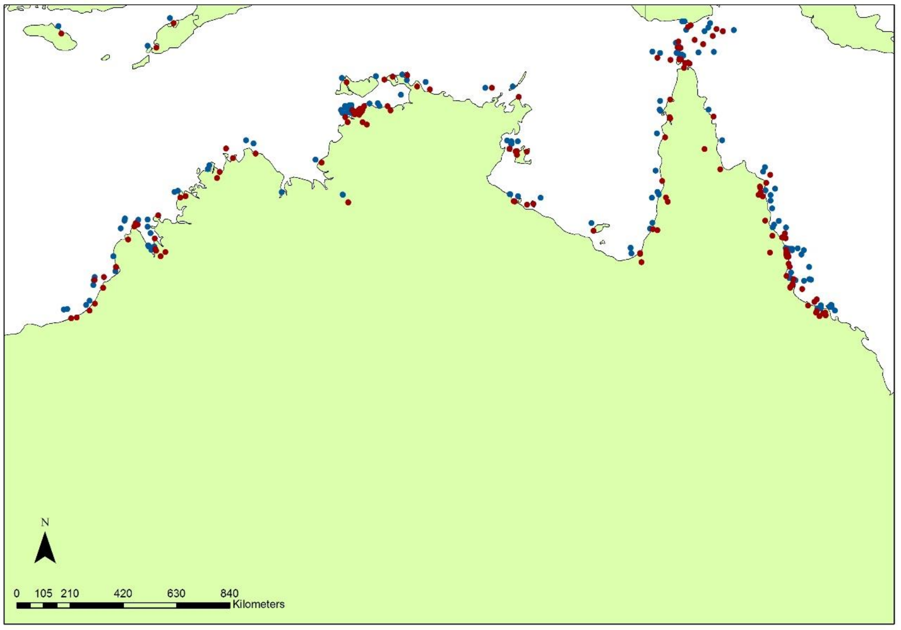

Figure 7 Map of the original station location situated in the ocean (blue dots) and the updated resolved station location (red dots) for Northern Australia.

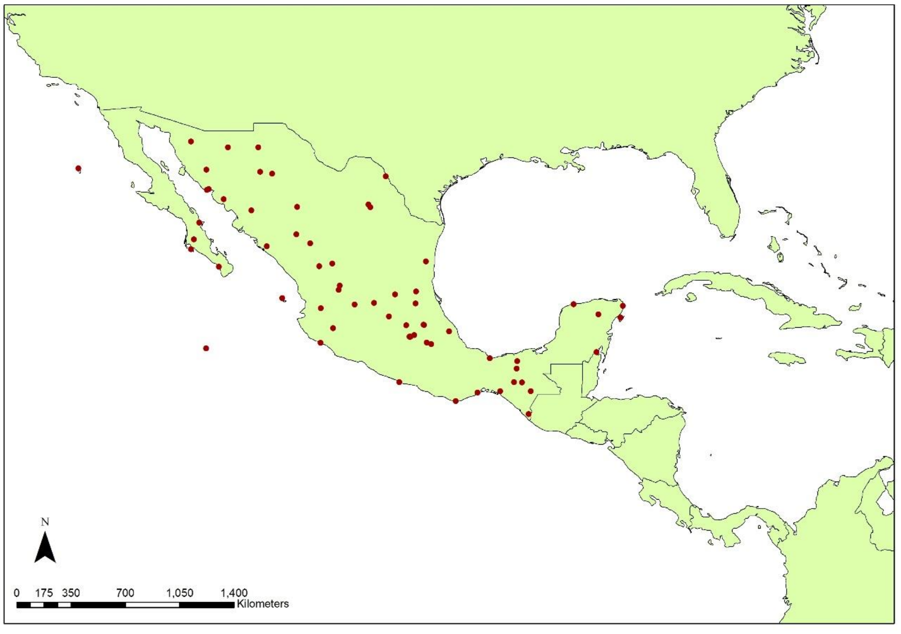

Figure 8 Map of the updated station location (red dots) for Mexico no original location coordinates were available.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

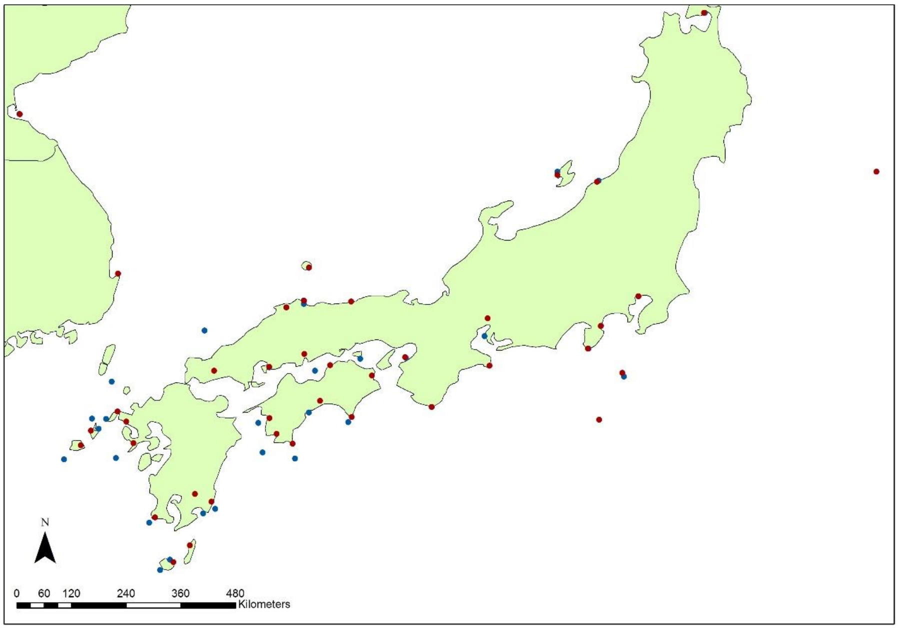

Figure 9 Map of the original station location situated in the ocean (blue dots) and the updated resolved station location (red dots) for Japan.

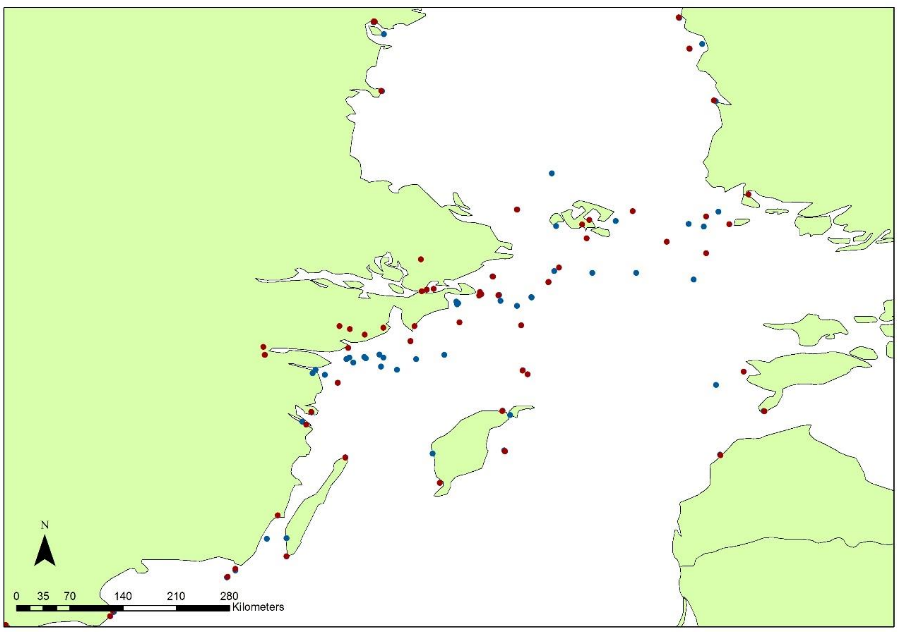

Figure 10 Map of the original station location situated in the ocean (blue dots) and the updated resolved station location (red dots) for parts of Northern Europe.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

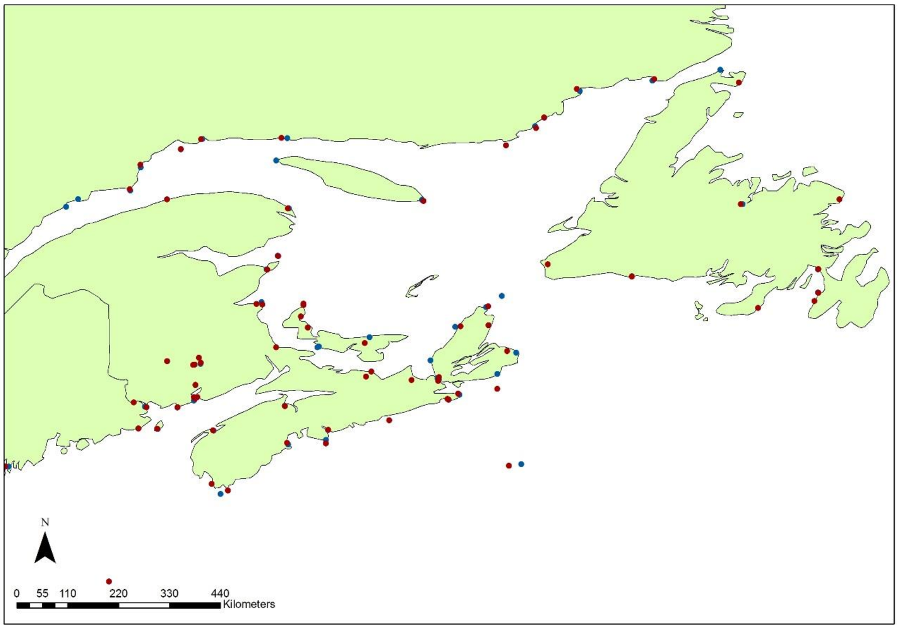

Figure 11 Map of the original station location situated in the ocean (blue dots) and the updated resolved station location (red dots) for eastern parts Canada.

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

Table 1 Results of questions (1a-1k) of the survey which asked the students to indicate the extent to which they agree or disagree with specific statements. The frequency of responses and the percentage of the total responses to each question are presented (152 students participating in the survey)

Question	Strongly disagree		Disagree		Agree		Strongly agree	
	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
(1a) I found the assignment provided insights into some of the issues with data	1	*1%*	8	*5%*	112	*74%*	31	*20%*
(1b) Supports provided were enough to aid completion of the assignment	7	*5%*	22	*5%*	102	*67%*	21	*14%*
(1c) The steps and guidance given were clear and easy to follow	7	*5%*	22	*15%*	102	*67%*	21	*13%*
(1d) The assignment provided me with an insight into the power of citizen science	3	*2%*	42	*28%*	90	*59%*	17	*11%*
(1e) I would prefer to participate in assignments like this over other, more traditional, types of assignments	3	*2%*	25	*16%*	91	*60%*	33	*22%*
(1f) The work load was appropriate to the level of credit given	7	*5%*	23	*15%*	96	*63%*	26	*17%*
(1g) Overall, I found the assignment to be a valuable learning experience	3	*2%*	42	*28%*	90	*59%*	17	*11%*
(1h) I was more motivated than is usual for me in doing this assignment	20	*13%*	51	*34%*	57	*38%*	24	*16%*
(1i) I feel that I have made a worthwhile contribution to an important global project.	7	*5%*	23	*15%*	96	*63%*	26	*17%*
(1j) I think next year's students would be happy if a similar assignment were run again	7	*5%*	38	*25%*	86	*57%*	21	*14%*
(1k) I think that I gained useful transferrable skills from this assignment	16	*11%*	63	*41%*	58	*38%*	15	*10%*

https://doi.org/10.5194/gc-2019-10
Preprint. Discussion started: 2 July 2019

Table 2 Results of question 2 (a-g) of the survey which asked students to give a score from (1-10) to a list of items, indicating how important each item was in enabling them to successfully complete the assignment. The table shows the percentage of the 152 responses to each question. (1 is not important and 10 is very important)

	2a. Lecturer's enthusiasm	2b. Clear assignment guidelines	2c. Online forum support	2d Lecture outlining the global project	2e Video from Dick Dee (ECMWF)	2f. In class support	2g. Knowing this was helping a global project
1 Least Important	0%	0%	7%	9%	15%	0%	5%
2nd choice	1%	0%	3%	4%	6%	0%	3%
3rd choice	0%	0%	4%	12%	13%	1%	4%
4th choice	2%	1%	7%	8%	7%	1%	7%
5th choice	7%	3%	11%	12%	17%	5%	15%
6th choice	7%	6%	9%	12%	14%	7%	13%
7th choice	12%	13%	14%	16%	13%	11%	16%
8th choice	30%	16%	16%	13%	9%	18%	17%
9th choice	16%	16%	13%	8%	5%	27%	12%
10 Most important	26%	45%	16%	8%	2%	30%	9%