COMP777 Major Project

Megan Trethewy (44883005)

2019-11-15

1 Description and Justification

Recently, climate change has become a global phenomenon that is altering the natural of various ecosystems. A typical trend associated with climate change is increased ocean temperatures, which has been significantly observed around the world. Consistent increases in temperature places threats to sensitive ecosystems, such as coral reefs, and could alter the nature of the environment. Marine heatwaves have been defined as extended periods of time (¿5 days) where ocean temperatures are warmer than usual (in the 90th percentile of normal temperature range) (Hobday et al. 2016). With increasing frequencies of marine heatwaves occurring globally, it is crucial for a consistent definition for identifying these events to be developed. It has been recognised that there are some inconsistencies in defining and identifying these ocean events, and Hobday et al (2016) attempts to create a universal metric for future analysis. This article is centred around a Python code which allows users to download data and identify the occurrence of marine heatwaves in specified geographical locations. The code that was incorporated in this article was well constructed and easily accessible through the use of a GitHub repository (Oliver, 2016). The article used elements of the climatological data (such as ocean temperatures) to perform assessments of marine heatwave occurrences across the world's oceans, with the aim of being able to apply the code to a diversity of contexts. This paper has been cited 185 times since it was published in 2016, in the journal Progress in Oceanography (Impact Factor: 3.245), highlighting the strength of the paper and the Python code associated with it.

2 The Original Work

2.1 The Original Code

The article presented by Hobday et al. (2016) was based around a Python script, which conducted the analysis of heatwave events that occurred at 3 locations; Western Australia (Ningaloo Reef, Indian Ocean), the Mediterranean Sea, and the North West Atlantic Ocean. To make this method of data analysis more accessible for wider groups of researchers, Schlegel (2019) created a RStudio script that followed the original Python code by Oliver (2016).

2.2 Running the Original Code

Due to issues installing and running the necessary Python packages, and having previous experience using the RStudio platform, Schlegel (2019) RStudio script was used for this investigation. The original script was run a total of 3 times to assess the reproducibility of each marine heatwave event identified by Hobday et al. (2016). Each of the codes used for the 3 marine heatwave events can be found at https://github.com/MTreth/COMP777$_{M}arineHeatwaves/tree/master.$

2.3 Success of the Replication

The success of the replication was assessed by comparing the graphs in the original article (Figure 1, Hobday et al. 2016) and those created by the RStudio code. Four types of graphs were produced by the code for each of the 3 datasets; the first indicated the specific marine heatwave event and highlighted the temperature of the ocean, the climatology (the average temperatures for that time of year) and the threshold (where temperatures are in the 90th percentile and become considered a marine heatwave), the second highlighted the duration of marine heatwave events (days), the third demonstrated the maximum intensity of the marine heatwaves (how high the temperature was above the threshold) and the fourth showed the cumulative intensity (the temperature x the number of days) (Figure 1). Each of the graphs produced from the RStudio code mirrored those of the original article (Figure 1, Figure 2), and a Cohen's Kappa Agreement Analysis further highlighted the success of the replication, with a high agreement rate (0.47). However, it was noted that in the total number of marine heatwave events in the figures (Figure 2), there were different values identified by Hobday et al. (2016) for each of the 3 locations. This difference occurred as more recent dates (after 2016) were included. Overall, the RStudio script provided by Schlegel (2017) was effective in replicating the original data presented by Hobday et al. (2016).

3 The New Data

3.1 New Data Description

The new data for this investigation was centred around the Great Barrier Reef, off the east coast of Australia (Figure 3). This coral reef ecosystem has been subjected to many marine heatwaves, particularly in recent years as a result of climate change. The marine heatwaves of 2016 and 2017 caused the greatest mass coral bleaching event seen by this ecosystem, wiping out approximately 60 percent of the northern reefs (AIMS, 2019). Due to this significant impact on the Great Barrier Reef, this marine heatwave event became the focus of the new data for this investigation. The use of Schlegel (2019) RStudio script on this location assessed the frequency and intensity of marine heatwaves along the Great Barrier Reef, Australia. By applying this tool to the new dataset, methods to predict the occurrence of damaging marine heatwaves could be better understood, assisting in the development of future strategies to prevent further damage to the coral reef ecosystem.

3.2 Creation of the New Data

For this dataset, sea surface temperatures from the last 40 years was retrieved from the ERDDAP, which is associated with the National Oceanic and Atmospheric Administration (ERDDAP, 2019). This was performed using the RStudio script developed by Schlegel (2019), which extracted all of the ocean temperature data within a specified longitude and latitude. For this new data, which focused on the Great Barrier Reef, the latitude was -10 to -40 S and longitude was 142 to 155 E (Figure 3). The raw data indicated the location, date and temperature for the Great Barrier Reef to be used for marine heatwave event identification.

3.3 Running the New Data through the Original Code

Once the original code had been run for the new data set, 4 graphs (mirroring the general themes of Hobday et al. (2016) (Figure 1)) were created, indicating that the code worked (Figure 4). In terms of the specific marine heatwave event highlighted, the graph indicated the period between both of the 2016 and 2017 events, with warmer then average winter ocean temperatures being noted. It is clear from the top graph (Figure 4) that much of the ocean temperatures recorded during this time were above the 90th percentile (the sections in red), further highlighting the severity of this heatwave. Later analysis did reveal that a previous marine heatwave, in 2004, was the most intense seen on the Great Barrier Reef from all the data points observed. This highlights how this code developed by Oliver (2016) and Schlegel (2019) could be used in future research to identifying similar intense events that could cause significant levels of destruction to coral reef ecosystems. To understand whether the results obtained for this new dataset were true to other resources available, comparisons of both government agency websites and previous literature were performed. The Australian Institute of Marine Science, in association with the Bureau of Meteorology, keep records of previous climatological data, particularly for the marine realm. An ocean temperature graph was conducted for the time when the 2016-2017 marine heatwave occurred on the Great Barrier Reef using this resource (Figure 5), and the general trends in temperature mirror that produced by the RStudio code (Figure 4). To further indicate the sturdiness of the new dataset, DeCarlo and Harrison (2019) conducted an analysis of coral bleaching events and marine heatwaves using Hobday et al (2016) code for the Great Barrier Reef. By comparing the results of their investigation to the new code produced here, it was seen that the number of marine heatwave events along the Great Barrier Reef was consistent across both investigations (Figure 4, Figure 6).

3.4 Future Implications

Marine heatwave events are predicted to become more common as the world continues to suffer from climate change (DeCarlo and Harrison, 2019), with each of the 4 marine heatwave datasets analysed showing an increase in the occurrence of these events in recent decades. With effective tools such as Hobday et al (2016) for identifying these events, better strategies for predicting and preventing the impacts of these events on sensitive ocean ecosystems (such as coral reefs) can be achieved.

4 COMP777 Tools and Techniques Used

4.1 Virtual Machine

For this code, large datasets were downloaded, due to the 40 year time period and large geographical ranges associated with the spread of a marine heatwave event. The use of the Amazon Web Services virtual machine assisted with downloading and running the large datasets associated with the identified script.

4.2 Cohen's Kappa Agreement Analysis

The main way to assess the success of the rerunning the original dataset was to observe both the number of marine heatwave events identified by Hobday et al (2016) and comparing the created graphs (Figure 1, Figure 2). This was conducted by analysing the agreement of two independent observers to understand whether the graphs that were included in the original article and the reproduced graphs mirrored each other. For this, Microsoft Excel 2016 was used to calculate the rate of agreement.

4.3 Downloading Data from Internet Resources

For the creation of new data to be run through the original code, the technique of obtaining data from an internet webpage was used. An RStudio script, developed from Schlegel (2019) "heatwaveR" script, was used to download data regarding the 2016 and 2017 marine heatwave event on the Great Barrier Reef, Australia. This data was gathered from the National Oceanic and Atmospheric Administration (ERDDAP, 2019) , which is the same location the original data was collected from (Hobday et al. 2016). By specifying the longitude and latitude points, as well as the time period, the necessary dates and ocean temperatures were downloaded.

4.4 OpenRefine and Observing the New Data

Once the new data had been downloaded from the internet, it was then exported to OpenRefine. This step was to assess whether the code used above had collected a dataset that could be used by the original RStudio script. By comparing the structure of the new data to the original examples identified in the Hobday et al. (2016), as sense of the effectiveness of the code was gauged. In this particular example, the new data included the specific longitude and latitude coordinates, the date and ocean temperature (degrees C). The overall structure was similar to that of the original datasets, differing with the inclusion of latitude and longitude results for each data point. This tool was also useful in creating a .csv file for the GitHub repository, to display the new data.

5 Supplementary Materials

Each of the figures and RStudio code discussed in this article have been saved as part of the GitHut repository, under SupplementaryMaterials.pdf

https://github.com/MTreth/COMP777$_MarineHeatwaves/tree/master$

6 References

Australian Institute of Marine Science (AIMS). 2019. Coral bleaching events. Viewed 12 November 2019. https://www.aims.gov.au/docs/research/climate-change/coral-bleaching/bleaching-events.html

Decarlo, T, Harrison, H. 2019. An enigmatic decoupling between heat stress and coral bleaching on the Great Barrier Reef. PeerJ. 7, pp. E7473

ERDDA, 2019. ERDDAP ¿ griddap ¿ Data Access Form. Viewed 20 October 2019. https://www.ncei.noaa.gov/erddap/index.html

Hobday et al. 2016. A hierarchical approach to defining marine heatwaves. Progress in Oceanography. 141, pp. 227-238

Oliver, E. 2016. marineHeatWaves. Viewed 10 October 2019.
https://github.com/ecjoliver/marineHeatWaves

Schlegel, R. 2019. heatwaveR. Viewed 10 October 2019.
https://robwschlegel.github.io/heatwaveR/index.html