CAPSTONE PROJECT BY TEAM 9

IDENTIFYING OCEAN SURFACE TEMPERATURE EXTREMES AT 2 LOCATIONS AROUND AUSTRALIA USING DATA SETS SPANNING 27 YEARS

XUANYU JIN (z5155473), SHU WANG (z5211077), YUXI YAN (z5211026)
ZHAOYI XIAO (z5229034), JIAQI ZHOU (z5229080),

School of Mathematics and Statistics
UNSW Sydney

November 2020

Submitted in partial fulfilment of the requirements of the capstone course DATA3001

Plagiarism statement

I declare that this thesis is my own work, except where acknowledged, and has not been submitted for academic credit elsewhere.

I acknowledge that the assessor of this thesis may, for the purpose of assessing it:

- Reproduce it and provide a copy to another member of the University; and/or,
- Communicate a copy of it to a plagiarism checking service (which may then retain a copy of it on its database for the purpose of future plagiarism checking).

I certify that I have read and understood the University Rules in respect of Student Academic Misconduct, and am aware of any potential plagiarism penalties which may apply.

By signing this declaration I am agreeing to the statements and conditions above.

Signed: Jiaqi Zhou Date: 21/11/2020

Signed: Zhaoyi Xiao Date: 21/11/2020

Signed: 金[illegible] Date: 21/11/2020

Signed: Shu Wang Date: 21/11/2020

Signed: 颜雨菡 Date: 21/11/20

Abstract

Marine Heatwaves (MHWs) are important climatic extremes in the oceanic system that have long-period on marine ecosystems and marine biodiversity.[3] Global average marine heatwave frequency and duration are shown to be increased from 1925 to 2016.[2] Measuring the severity of MHWs and growing public awareness is becoming more important. In this project, MHWs happened near the seashore and in the deep ocean in Port Hacking Site and Rottnest Island Site in the time interval from the start of 1993 to the end of 2019 are detected, and main characteristics of these MHWs include frequency, duration, and intensity are compared to investigate how they perform differently in different sites or different locations in the same site and how they have changed over the past 27 years in the same or different ways. The analysis part of this report shows there are increases in both frequency and duration of MHWs in all locations in both sites over the past 27 years, but there is no obvious increase in MHW intensity. MHWs in different locations in the same site perform similarly, but MHWs in different sites perform differently due to the overall temperature differences in different locations around Australia. The result of this project shows that even if MHWs in different locations perform differently, but MHW does happen more frequently and lasts longer in recent years around Australia. Since when ocean temperatures are extremely high for a long time, there are significant impacts on marine ecosystems and economies based on sea lives, this report could help people, especially for people in Australia, raise the general awareness of the severity of MHWs happened around Australia, so people could avoid actions that will cause MHWs and then protect vulnerable marine habitats and resources.

Contents

Chapter 1

Introduction

Marine heatwaves are defined as phenomena where, in the ocean, water temperature keeps remarkably high in a time interval [3]. Over the past several decades, it is monitored that marine heatwaves occur with higher frequency at higher temperature and continue over longer periods.

In 2016, Great Barrier Reef has experienced a marine heatwave which causes a coral bleaching event. Only 70 per cent of the corals still alive after the bleaching [1]. Scientists have to bring more evidence to show that marine heatwaves can impact Australia's coral ecosystem in a long-term and destructive way. As a chain reaction, industries such as fisheries, aquaculture and tourism will finally affect the country's economy[1]

As an abnormal oceanic event, marine heatwaves appear to be an unstable factor and have the potential of undermining the ocean ecosystem and economy based on sea lives. Thus, to help protect the lives in the ocean, knowing more about the characteristics and measuring the severity of marine heatwaves has become an urgent problem now.

This project aims to compare how characteristics of MHWs perform differently in different sites or different locations in the same site, and how characteristics changed in the same way over the past 3 decades.

In this project, 4 boxes are selected from 2 National Reference Stations around Australia, Port Hacking Site (-32.01S to -36.13S, 150.01E to 153.99E, near Sydney) and Rottnest Island Site (-30.51S to -32.99S, 113.01E to 115.79E, near Perth), and all MHWs happened in these boxes from January 1st, 1993 to December 31st, 2019 are detected. Each site has one box selected near the seashore and another box selected in the deep ocean. Characteristics include duration, frequency, and intensity of MHWs in each box are compared overtime to conclude how these characteristics of MHWs have changed over the past 27 years. Also, characteristics of MHWs are compared among boxes to see how MHWs in different locations perform differently. This project uses Python to access the data and uses R to make plots containing tendency lines and fitted lines.

1.1 Introduction to Marine Heatwaves

The definition of marine heatwaves (MHWs) used in this project is defined by Hobday in 2016 [4]. It defined that a marine heatwave is when seawater temperatures exceed a seasonally-varying threshold for at least 5 consecutive days. Successive heatwaves with a gap of 2 days or less are considered as the same event [3].

Some terminologies are needed to understand this definition. The first one is climatological mean, which is the average of all seawater temperatures on a certain date in all years. For example, the temperature used in this project are sea surface temperatures for 27 years, then the climatological mean of January 1st is the average of all SSTs in January 1st in the 27 years. The threshold used in this project is the 90th percentile of SSTs in the same location in the 27 years.

There are some characteristics of MHWs used in this project for analysis. The duration of a MHW is the number of days from the start of the MHW to the end. The frequency of MHWs is the number of MHWs in a certain time block, like annual frequency or frequency in each 3-year time block. Intensity is the difference between the temperature and the climatological mean in degrees Celsius. There are 3 kinds of intensities used in this project. The maximum intensity is the difference between the peak temperature of the MHW and the climatological mean. The mean intensity is the average of all intensities in a MHW. The cumulative intensity is the integral of intensity over the duration of the MHW. The MHW days is the number of days that were part of a MHW in a time block; its value equals to the sum of durations of all MHWs happened in the time block [3].

Chapter 2

Literature Review

It is observed that MHWs can have devastating and long-term impacts on ecosystems, with subsequent socioeconomic consequences. It is found that the frequency and duration of global MHWs increased by 34% and 17% respectively from 1925 to 2016 [2], and this trend is likely to be continued with global warming.[5]

There are three main characteristics of MHWs: duration, intensity, and frequency, and details of them are introduced in Chapter 1. These global MHW properties are calculated by the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation (OI) Sea Surface Temperature (SST) [2]. Due to some property of MHW, there would be some biases, so they give the average difference between two-time slices without making any assumptions about the distribution of the data. There are six century-scale monitoring stations to detect daily SST, but there are also some missing data and gaps. To fill gaps, they calculated differences in mean MHW properties between two 30 years. They built a linear regression model to describe the relationship between MHW properties and annual proxies, which is Properties = beta0 + beta1*proxies + u. The intensity of MHW was modelled by a normal distribution, frequency and duration were approximately Poisson distribution. By satellite record, they built some graph and concluded that the frequency of MHW has a significant increase in the trend of 0.45 annual events per ten years from 1982 to 2016. Between 1982–1998 and 2000–2016, more than 65% of the MHW intensity in the global ocean increased, global MHW intensity increased 0.085°C per decade. Between 1982–1998 and 2000–2016, the average MHW duration increased by 84% of the world's oceans. Since 1982, the average duration of the global ocean has increased by 1.3 days per decade for the duration of MHW. Besides, there was a negative relationship between the average frequency and duration of MHWs, and SST positively correlated with MHW intensity.

Considering the possibility of continued warming of the ocean surface throughout the 21st century, it is predicted that the frequency and duration of MHWs will keep increasing. This will have an impact on marine biodiversity and the goods and services provided by marine ecosystems [2]. Keep a long-term record will boost discovering and improving the influence of MHWs.

Hobday's article points out a hierarchical approach to recognized MHWs according to different levels of intensity.[4] They are Moderate (Category I), Strong (Category II), Severe (Category III) and extreme (Category IV) respectively. To be more specific, Category I is intensity 1–2 times higher than 90th percentile. Category II is 2–3 times higher. Category III is 3–4 times higher while category IV is any MHWs with intensity higher than Category III. The author also recommends to settled the thresholds points for each category as fixed values to avoid being changed by world warms.

Hobday also recommends studying MHWs by comparing data at the same depth. After considering information about MHWs' intensity, duration, frequency and spatial extent, one can conclude about how MHWs have changed. In this case, Schaeffer and Roughan's study will be a great example.[6] In the article 'Subsurface intensification of marine heatwaves off southeastern Australia: The role of stratification and local winds', they use data from sea surface until 100 meters depth to study the relationship between depth and MHWs. This article shows that MHWs are most intense at depth, regularly extend to the bottom of the water column and are driven by downwelling favourable winds during periods of weak stratification.

Chapter 3
Material and Methods

3.1 Description of the Data

Raw files are in NetCDF (nc) format, a self-described, portable, appendable, sharable and archivable data format commonly used in the research of environmental science. Given the relatively small file size, which is 800 GB to 1 GB, it is stored locally on the disk of the group member who is responsible for raw data processing. Three of these nc files are available, representing 3 distinct locations of interest: Port Hacking, Rottnest Island, and Maria Island.

There are 12 variables in each NetCDF file. Details of these variables are shown in "Table 1: Description of Variables" in the Appendix.

3.2 Software

Software used in this project can be classified into 3 categories: version control tools, coding assistance like editor, and teamwork collaboration tool.

Github is selected as the version control tool that is used throughout the project. Although this project itself does not require a large number of files or massive iterations of code, it is still in need of version control tool. For one thing, functions are not written altogether by one person, but section by section by 2 people of the group. Having a version control tool enables both of the group members to have access to the newest version of the code that the other person has written and thus, increases productivity. For another thing, it enables one to retrieve the previous version of the code so that if the current version of code has some bugs that cannot be easily fixed, or has some redundant feature, one could easily go back to a previous version. Github does not differ from other git-based version control tools, like GitLab vastly, but members of the group are more familiar with Github, making Github the choice of version control tool.

As for code assistance tool, Anaconda and Jupyter notebook serve different yet important purposes in doing the project. Anaconda provides a virtual environment so that it is possible to install packages and set up a programming environment without changing anything to the main operating system. Given group members are to cooperate with each other and may have different versions of programming language installed on their computer, and consider the problem of incompatibility, it is necessary to ensure that all members of the group are on the same page, and Anaconda can perform the job as the group requires. Jupyter notebook enables the team to execute the code piecewise and test each part of the code to see if it works as expected. Trello is used as a tool to precipitate the teamwork. It helps to create timelines and schedules to inform individuals of which the part of the mission are they responsible for.

In terms of programming language, Python and R are selected to fulfill different goals. Python is used to read, process and export data, and R is to take the output of Python and produce graphs and other useful statistical data. They are chosen for different reasons.

Starting from R. It is a powerful programming language for statistical usage. In this project, it is used to generate graphs necessary to make the analysis. There are 2 reasons to choose R. Firstly, all group members are subjected to the training of R, making it understandable by the whole team. Although only 2 group members are responsible for programming and producing the graph, the whole group being able to understand the code enables everyone to check for errors, doing supplementary work if the code is not complete, and make suggestions about what could be added to produce a better result. Secondly, R is a free, open-source, and many packages are available. That means R is powerful enough to handle all kind of possible problems. From the very beginning, it is not clear to the group that what kind of problem is to be solved, so having a powerful yet easy-to-use tool is essential to get the work done. In the end, it is proven correct that unforeseeable problems, a formatting issue about plotting, did occur and R is, to some extent, the most efficient solution.

Python is used for data extraction, process and exportation. It is chosen because it fits the raw data stored in the NetCDF (referred as nc) file. While R does have packages that support reading nc files, the speed of performing operations such as accessing specific element, looping through the data set, is slow, and it is complicated to do more advanced programming, such as nested for loop and reading command-line arguments, with R. On the other hand, it is easily achievable with Python. Also, one very important open-source package, which directly calculates the MHWs, is available as a python file. Working with python makes it easier to take the advantage of all available resources.

To make full use of Python, 5 modules are used and will be introduced separately: netCDF, numpy, datetime, marineHeatWaves, csv. Only the modules that are used directly will be introduced.

Starting from marineHeatWaves. marineHeatWaves is provided by Hobday et. al. (Hobday et, al, 2016) as a definition of what a MHW is. The function 'detect' is of vital importance to the project. It takes in 2 numpy arrays, one represents time series and another represents locational temperatures, and output detected MHWs. More about this will be introduced in later chapters. numpy is another important module. It serves many purposes: first, since the detect function from marineHeatWaves takes numpy arrays only, numpy is used to convert python list to numpy array; secondly, np.nan is used to hold the place if data is absent at a particular location of a given time; last but not least, the masked array feature is used to determine if a data collected is biased.

netCDF performs a simple role in accessing the raw nc files. It simply opens the nc files so the program can access the data and is able to do further process.

csv is the packages used to export read data as .csv files, and the exported files will be used as input of R scripts to generate plots and other useful statistical information.

datetime is used to transform the ordinal numbers representing the date and time to machine-readable datetime data so that the marineHeatWaves module can process.

3.2.1 Selection of Variables

As mentioned in the description of variables in 3.1, there are 12 variables in total. However, not all variables are useful in this project. The function used to detect marine heatwaves needs 2 input vectors: time and sea surface temperature.

The variable "time" and "sst_dtime" are both used to represent the time, but "sst_dtime" is used to represent the time in a day that the sea surface temperature was measured while "time" is used to represent the day that the sea surface temperature was measured, and "time" is also a coordinate use in other variables. Therefore, "time" is chosen as the time vector.

Variables "lat" and "lon" need to be kept since they are used as coordinates to locate other variables.

The variable "sea_surface_temperature" is also needed since it is the input vector of the marine heatwave detection function. The variable "quality_level" is the overall quality indicators of sea surface temperatures. It is important for data selection and avoiding errors caused by bad data quality. The variable "sses_bias" is an estimate of the bias of the measurement with respect to other measurements made under similar view, quality and geographical conditions [7]. The debiased sea surface temperature, which is the sea surface temperature used in this project, is given by minus the sses_bias from the original sea surface temperature.

3.3 Data Cleaning

Data sets used for this project provide SSTs from 1992-03-21 to 2020-09-17. However, since in the analysis part, the time will be divided into time blocks, and use the function 'blockAverage', which has an option called 'blockLength' in years, to calculate averages of characteristics of MHWs, so days before 1993-01-01 and days after 2019-12-31 are deleted from the time vector to make analysing easier. In the later coding part, the amount of missing data also proved that removing days in 1992 and 2020 is the correct choice.

Since each site covers a very large area with at least 1,5000 pixels in it, not all pixels are used in this project. Boxes are chosen in each site as samples instead, and then use the average of SSTs in the box in each day as the temperature vector to detect MHWs happened in the box.

There are many advantages of using average SSTs in a box rather than SSTs at one pixel to detect MHWs: First, calculating the average could increase the amount of data. This is because when the SST is missing on this day in this pixel, it might not be missing in all pixels in the box. Second, when using SSTs for one pixel to detect MHWs, only SSTs with a quality level larger than or equal to 4, meaning SST with acceptable quality, could be used, but if use SSTs in a larger area, like a box, this standard could be lowered to 3.[8] This enlarges the amount of data as well.

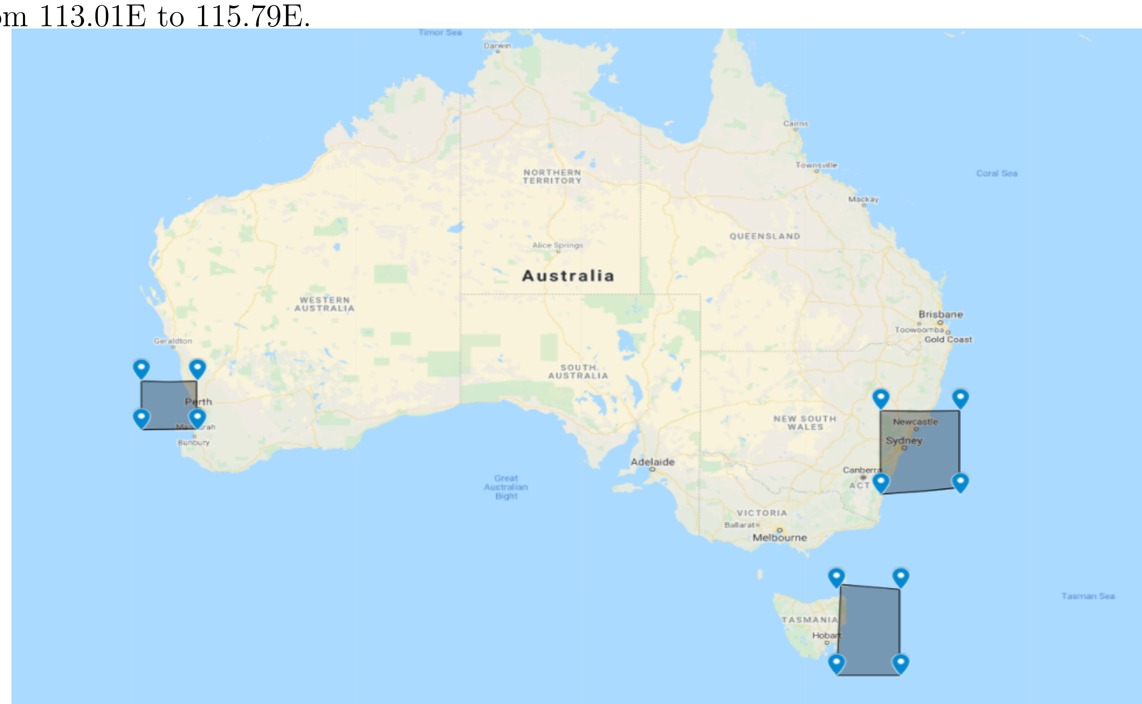
There are 3 rules for choosing the location of boxes: First, choose 2 boxes for each site, and those two boxes should be far from each other so they will not be involved in the same MHW. Second, since each site covers an area from the seashore to the deep ocean, it is desired to choose 2 boxes that one near the seashore and the other in the deep ocean, and investigate if there are differences between MHWs happened near the land and MHWs happened in the deep ocean. The last rule is to find boxes with the largest number of SSTs to minimize errors caused by missing data.

There is still one thing that needed to be thought about is the size of boxes. Since the average of SSTs in the box is used, the number of data increases while the box size increases. However, if the box size is too large, the difference of SSTs in the box will be large as well, then the average SSTs could not represent the actual temperature in the box. Therefore, what is the most appropriate selection of box size is also a problem to solve before starting to detect MHWs.

Locations for the 3 sites used in this project are: Port Hacking Site, which is near Sydney, it covers from-32.01S to -36.13S, and from 150.01E to 153.99E.

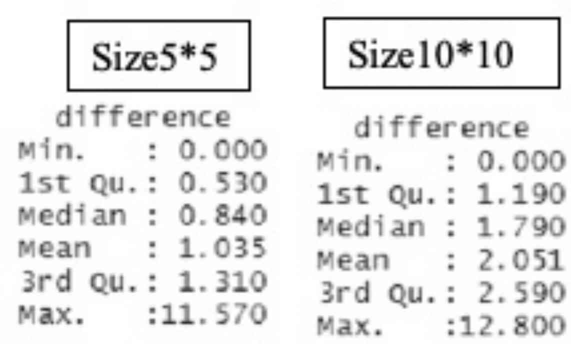
Maria Island Site, which is near Tasmania, it covers from -40.51S to -44.29S, and from 147.81E to 150.99E.

Rottnest Island Site, which is near Perth, it covers from -30.51S to -32.99S, and from 113.01E to 115.79E.



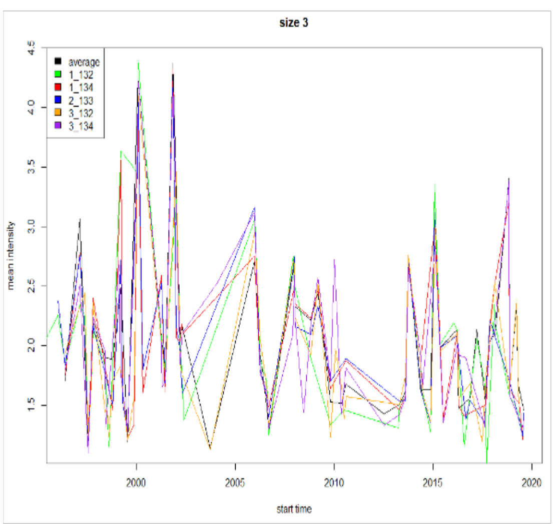
There are 9861 days from 1993-01-01 to 2019-12-31. After thinking about all three rules of choosing the location of boxes that mentioned above, four boxes will be selected from Port Hacking Site and Rottnest Island Site once the size of boxes is decided. No box will be chosen from Maria Island Site since almost all pixels in the deep ocean area in this site have less than 4500 SSTs out of 9861 days, which is less than a half, this number of SSTs is too low to detect MHWs. The location of the first box in Port Hacking Site is near the latitude index 1 and longitude index 132. Pixels near this pixel has about 5700 (57.80%) SSTs out of 9861 days, and about 4700 (82.46%) SSTs out of these 5700 SSTs have a quality level of 3 or higher, which is about 82.5%. This box is used to decide the box size.

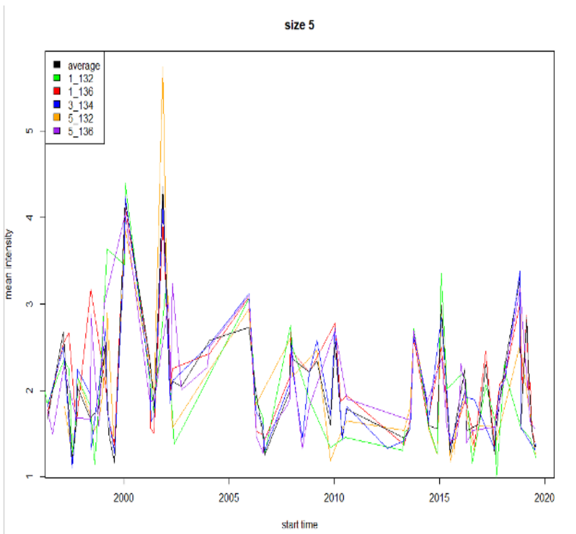
Boxes with size 10*10, 5*5, and 3*3 near the pixel mentioned in the paragraph above are compared to select the most appropriate size of boxes. The first thing to compare is the difference between the highest temperature and the lowest temperature on the same day in each box. In the box with size 10*10, the mean difference of temperatures is 2.051 degrees Celsius, which is even higher than the mean intensity of many MHWs. This means the average of SSTs in the 10*10 box is unrepresentative of all SSTs in the box. Mean difference in the box with 5*5 size is 1.035 degrees Celsius, which is acceptable.



Size5*5	difference	Size10*10	difference
Min.	: 0.000	Min.	: 0.000
1st Qu.	: 0.530	1st Qu.	: 1.190
Median	: 0.840	Median	: 1.790
Mean	: 1.035	Mean	: 2.051
3rd Qu.	: 1.310	3rd Qu.	: 2.590
Max.	:11.570	Max.	:12.800

As mentioned before, an increase in the box size leads to an increase in the number of average SSTs but might cause the loss of MHWs happened in parts of the box at the same time. After calculating the average SSTs in boxes with sizes 3*3 and 5*5, there are 6376 (64.66%) SSTs in 9861 days in the 3*3 box, and 5331 (83.61%) SSTs of them are with a quality level of 3 or higher, while there are 6779 (68.74%) SSTs in 9861 days in the 5*5 box, and 5769 (85.10%) SSTs of them are with a quality level of 3 or higher. Then five pixels, four at corners and one in the centre of the box, are used to detect MHWs happened in each pixel. Figures below show the comparison of mean intensities of MHWs happened in each pixel and MHWs detected by average SSTs in the box.





According to line charts above, lines of mean intensities in 5 pixels in size 5*5 box have a more similar shape to the average line. Also, 89 different MHWs are detected in 5 pixels in 3*3 box, and 52 (58.42%) of them are detected by average

SSTs, while 97 different MHWs are detected in 5 pixels in 5*5 box, and 58 (59.80%) of them are detected by average SSTs. Both the larger number of data and the lower portion of the loss of MHWs are related to the box size 5*5, so the most appropriate size of boxes is decided as 5.

Since the box size is 5, the first box has latitude index from 1 to 5 and longitude index from 132 to 136. There are 10408 days from 1993-03-21 to 2020-09-17 and after calculating average SSTs in this box, 7002 (67.28%) days out of 10408 days have SSTs. There are 286 days from 1992-03-21 but only 108 (37.76%) days have SST records, and there are 115 (43.89%) days out of 262 days from 2020-01-01 to 2020-09-17 have SST records. Deleting days in 1992 and 2020 makes the portion of data increase from 67.28% to 68.75%, so it is the correct choice.

As mentioned before, the average of SSTs in the box in a day will be used as the SST in this day, and the list of averages of SSTs will be used as the temperature vector to detect MHWs. However, the average is not calculated by all SSTs in the box in a day, but use some rules to control the quality of data. Since the box size is 5, there are at most 25 SSTs in the box in a day. If there are no SSTs in the box in a day, then the SST for this day is set to be "nan"; if there are N SSTs but there is not SST with a quality level larger than or equal to 3, then the average SST for this day is the average of those N SSTs because it is better to have low-quality data than no data; if there are N SSTs and M of them have quality level larger than or equal to 3, then the average SST for this day is the average of those M good quality SSTs (N and M represent positive integer numbers here).

Information of 4 boxes chosen from Port Hacking Sites and Rottnest Island Site are shown in "Table 2: Selection of Boxes" in the appendix.

The "Selection of Boxes" table shows there are still many gaps after calculating averages of SSTs in the box. Consider the definition of MHW, "An MHW event occurs when the threshold temperature exceeds 90%, lasts for 5 days or more, and it is more than 2 days from the next MHW. (Hobday, 2016)", gaps are divided into 3 different types according to their length and their influence on the detection of MHWs:

1.Gaps no more than 2 days

2.Gaps more than 2 days but less than 5 days

3.Gaps of 5 days or more

Gaps no more than 2 days will be filled by the interpolation. There are 3 different situations of this type of gaps different by their positions compared to a MHW: a gap before a MHW, which might be part of a MHW or not, using interpolation here might cause the loss of the first day or the first two days of this MHW; a gap in a MHW, which will not affect the length of this MHW since the temperature generated by interpolation between two temperatures in a MHW must be a temperature above the threshold, also, two MHWs with a no more than 2 days' gap between them will be considered as the same event; a gap after a MHW, which might cause the loss of the last day or the last two days of this MHW. Gaps in the first or the last situation might have effects on the detection of MHWs since if the missing days are part of a short MHW, but the temperatures in gaps after interpolation are not above the threshold, then this MHW might not be detected as a MHW since it might have only 3 or 4 consecutive days with temperatures above the threshold.

Situations for gaps with length larger than two days but less than 5 days are almost the same as these for gaps no more than 2 days, and gaps in this type are filled by interpolation as well. A gap in this type before or after a MHW might cause the loss of a MHW or a reduction in the duration of a MHW.

Gaps of 5 or more days will be filled by climatological means in these days. Since 5 consecutive days with temperatures above the threshold will be detected as a MHW, if there are 5 consecutive missing days, it is hard to say if there is a MHW happen in this gap. Filling by climatological means makes sure the temperatures in this gap will not exceed the threshold and filled values of temperatures will not affect the calculation of climatological means.

Numbers of gaps in different types are shown in "Table 3: Number of gaps" in the appendix.

3.4 Pre-processing Steps

To retrieve valuable information from raw data, pre-processing is needed to isolate variables with interest and detect MHWs from those variables. The detected MHWs is then saved into .csv files to be readable by R for later usage.

The function used to find MHWs is function detect() from marineHeatWaves module and it takes 2 main input numpy arrays: t, the time vector in datetime format, and temp, the temperature vector assumed degree Celsius.

To match the input requirement of detect() function, relative data are extracted and stored into python list and converted into numpy array. Number of maximum days over which to interpolate pad missing data specified as nans in input temp time series, the option maxPadLength, is selected to be 4 for reason mentioned above.

One factor to notice is that sea surface temperature, or simply SST, from raw data are stored in Kelvin rather than Degree Celsius. To convert it into Degree Celsius, it is needed subtract 273.15 from the original number.

3.5 Assumptions

The assumption made in this project is temperatures within a box on the same day are similar, and the average of SSTs in this day could represent the temperature in the box in this day.

Chapter 4

Exploratory Data Analysis

After data pre-processing, there will be 16 .csv files produced for R to make statistic summary and plotting by the Python program. These files represent data for 4 specific locations from 2 sites. Maria Island is not selected for the reasons mentioned above. For 1 such location, 4 files are available. One of them lists all MHWs of that location will all features, the other 3 contains similar information but with frequency-related data. They will be introduced separately.

For the file with all MHWs, it contains 8 columns of data: index_number, time_start, time_end, duration, peak_time, max_intensity, mean_intensity, cumulative_intensity. Index_number is a reference of the index number a certain MHW, and it serves as an identifier and is insignificant in analysis; time_satrt and time_end are the starting time and end time of that MHW respectively; durations is how long the MHW has last; peak_time is when the MHW reaches maximum intensity (intensity with the largest value); max_intensity is the value of the maximum intensity; cumulative_intensity is the sum of all intensity values of intensities of every day within that MHW period; and mean_intensity, the average value of the intensity, which is calculated using cumulative_intensity to be divided by duration. Numerical summaries of columns in these files are in "Table 4: Numerical Summaries" in the appendix.

Those 3 files with frequency information are similar to each other, but the frequency is calculated in different ways. In the context of this project, "frequency of MHWs" is defined as "number of MHWs occurred within a period of time". The difference in those 3 files is the time interval chosen. Frequency is calculated in 3 different ways: divide the 27-years data by 1, into 27 blocks; divide the 27-years data by 3, into7 blocks; and divide the 27-years data by 9, into 3 blocks, and calculate numbers of MHWs within those intervals.

Those .csv files are then supplemented to R to produce graphs used for analysis. Plots are available in later sections.

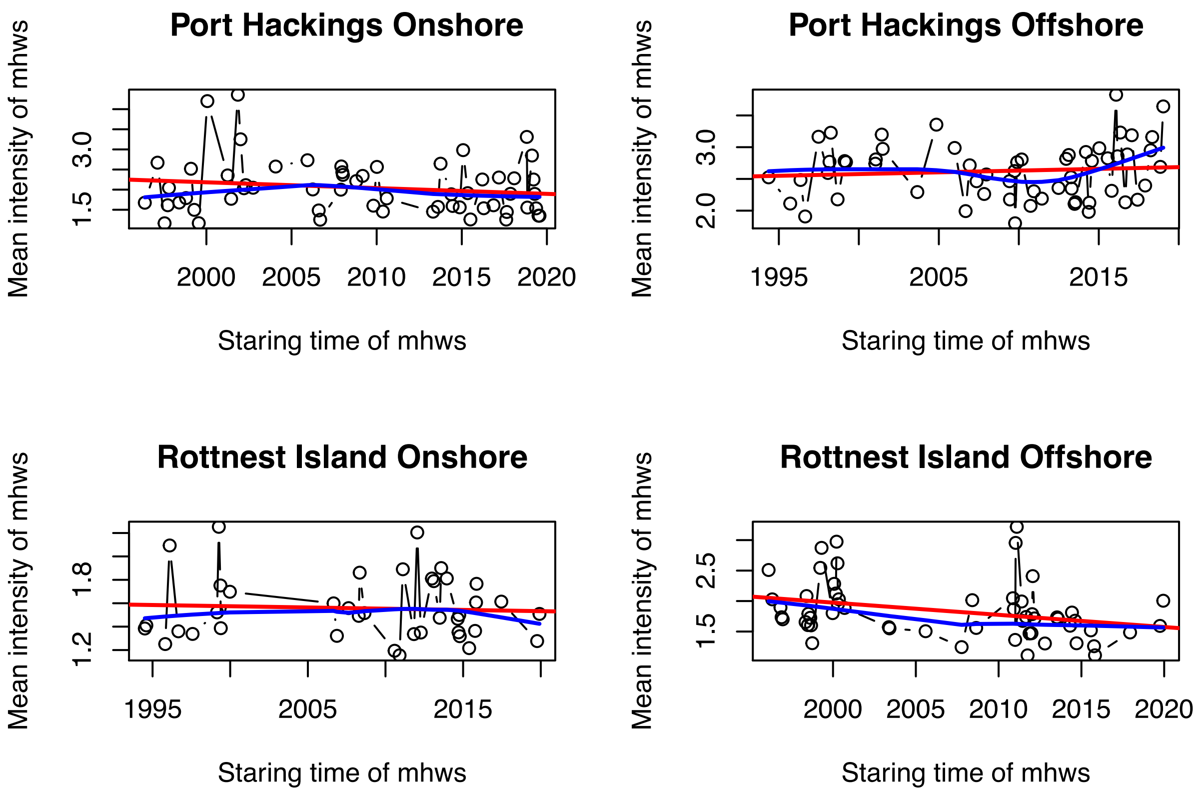
Chapter 5

Analysis and Results

There are eight figures that have been made to investigate the key characteristics of MHWs. At each site and as for each MHW, features include duration, maximum intensity, and mean intensity of MHWs are calculated, and counts of MHWs concerning per year, per 3 years and 9 years are also calculated as frequency.

This part of the report focuses on analyzing the overall trend of each characteristic at the different monitoring site, both onshore and offshore. The red line in the graphs are the fitted regression lines indicating the overall tendency of the data, and the blue line is the local regression line showing the moving average of the data locally.

5.1 Figure1 “plot of mean intensity against date”



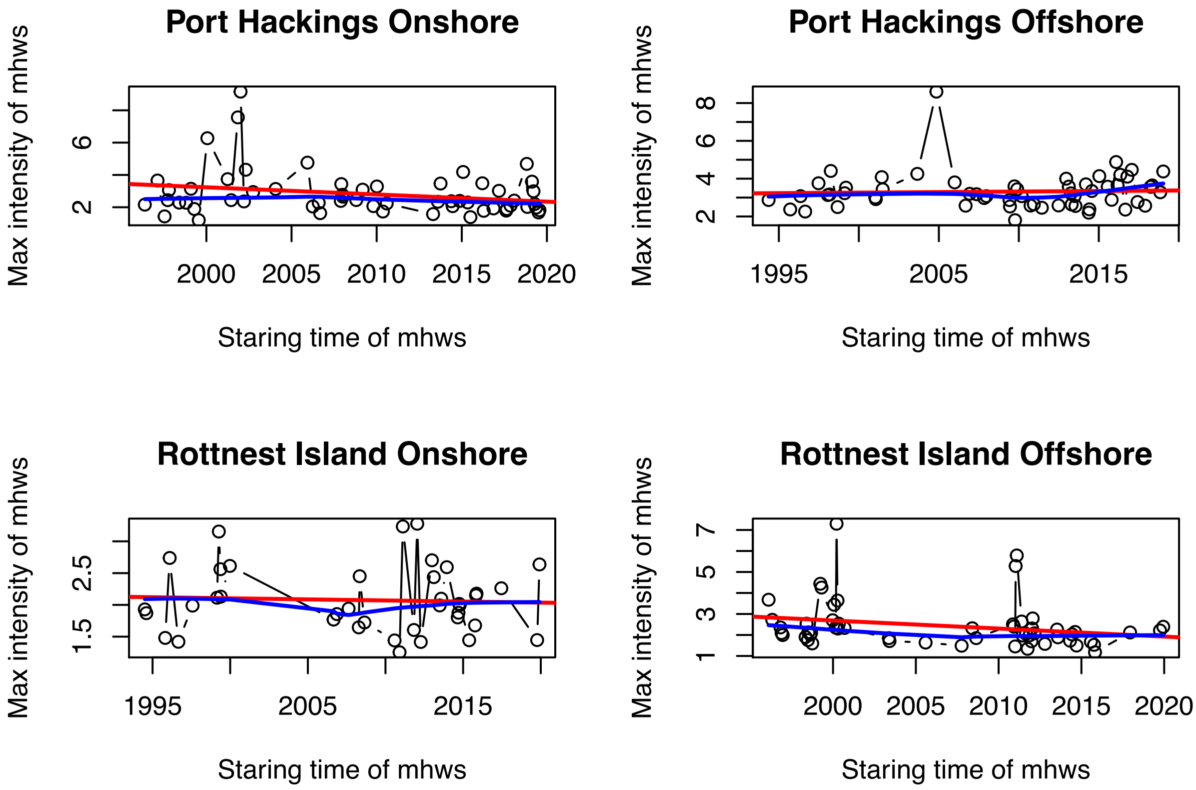
[A port hacking onshore; B port hacking offshore; C Rottnest island onshore; D Rottnest island offshore] When looking at how mean intensity changed in the past 3 decades at Port Hackings and Rottnest Island, it could be noticed that both Port

Hackings and Rottnest Island onshore boxes have large variations in the mean intensity of MHWs. From figure1.a and figure1.c, it is clear how onshore location related together. It is shown that there is a point with extreme high mean intensity at both Port Hacking and Rottnest Island in 2000, which is far away from the fitted line. However, the maximum mean intensity of the two sites has a large difference, 4.4 ºC for Port Hackings and 2.22 ºC for Rottnest. Besides, the average mean intensity at Port Hackings is approximately equal to the maximum mean intensity at Rottnest. It is deduced that this phenomenon is caused by geographical factors. Apart from that, there is also a high mean intensity point in 2002 at Port Hacking. As for Rottnest Island, there was also a point away from the fitted line, with 2.2 ºC of mean intensity. To sum up the tendency respectively, there is a decreasing fitted line along with year goes, since, after the peak in 2002, the mean intensity of MHWs are stably distributed along the fitted line. However, because of a sub-peak in 2012, the fitted line for Rottnest goes horizontally. Two graphs have a similar smooth line, the relatively low intensity in the 1990s, then increases in the 2000s and decreases in 2010s.

By comparing offshore Port Hackings and offshore Rottnest Island, the overall tendency of their mean intensity graph is different. In figure 1.b, the red fitted line is slightly increasing. Thus, it is predicted that the mean intensity of MHWs at offshore Port Hackings will keep climbing. However, the fitted line for offshore Rottnest Island has a linearly decreasing tendency in figure 1.d. One similar factor between these two graphs is that they both have a good interval (low variation). It is clearly shown that there is a stable mean intensity trend before 2010 in figure 1.b, but it reached a nadir in 2010, the mean intensity of MHWs at offshore Port Hackings achieved 1.5 ºC. After 2010, the MHWs mean intensity kept increasing and reached a peak (3.3 ºC) in 2016. As for figure 1.d, it has a fluctuating trend before 2000 and achieved a peak of 3.5 ºC until 2011. However, the MHWs mean intensity at Rottnest Island had a sharp decline in 2012, reached a nadir at 1 ºC, then kept dropping.

Figure1.a and Figure1.b show the mean intensity of marine heatwaves onshore and offshore of Port Hacking in this period. Smooth lines in both figures indicate a small upward trend from 1993 to 2005, and a slight downward trend appears in the following 5 years. Then the smooth line of offshore goes upward while onshore's smooth line continues downward and then it becomes flat. This causes two different fitted lines with opposite slopes. Figure1.c and Figure1.d show the mean intensity of marine heatwaves onshore and offshore of Rottnest Island in this period. The line chart in both figures shows a small downward trend between 2000 to the year 2007. And a slight upward trend appears in the year 2013. Then a smooth line of onshore goes downward while offshore's smooth line becomes flat. This causes two different fitted lines with different slopes.

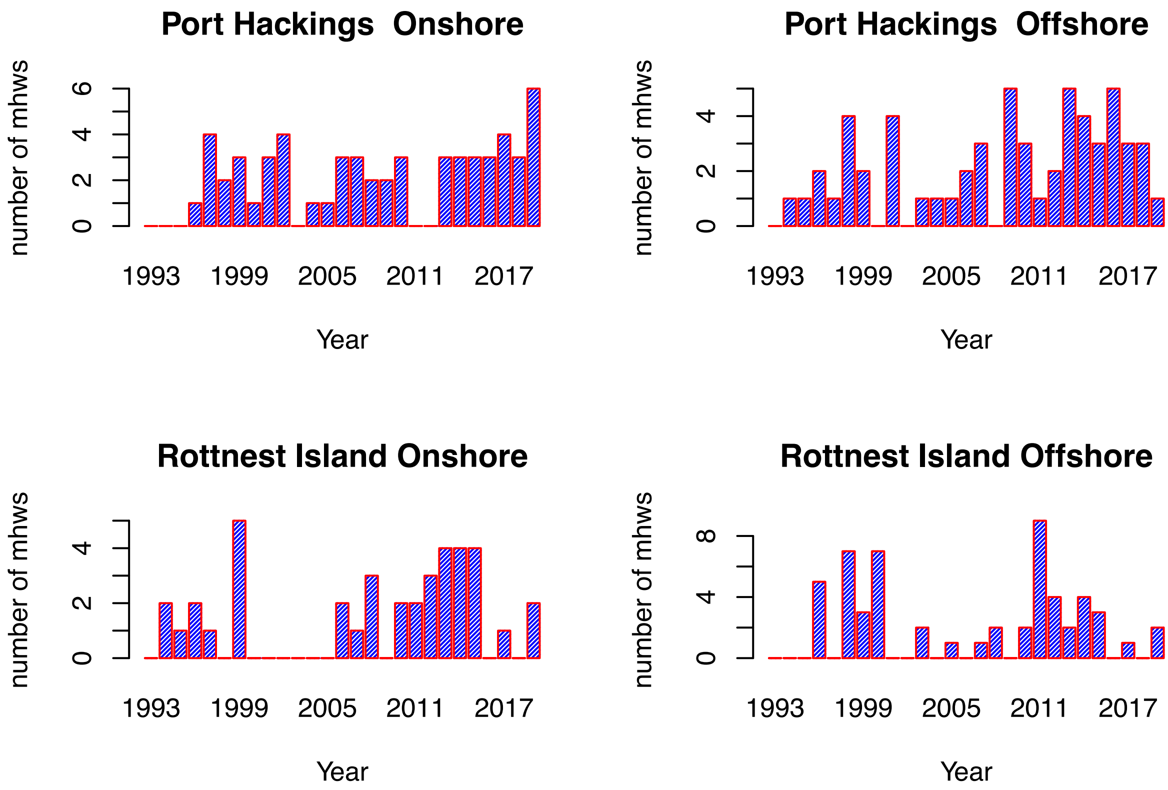
5.2 Figure2 “plot of max intensity against date”



[A port hacking onshore; B port hacking offshore; C Rottnest island onshore; D Rottnest island offshore]

Similar to mean intensity, even though the maximum intensity all occurred after 2000, 3 of the locations end with a decreasing tendency. Fitted line in figure 2.b almost displays a horizontal trend. When looking at Rottnest Island, it still has the lowest maximum intensity for both onshore and offshore. Figure 2.c shows the onshore location only reach 3.3 ºC as its maximum intensity, which is almost 3 times lower than Port Hacking. From figure 2.a and figure 2.c, it is clearly shown that there is a large variation in the maximum intensity of MHWs at Port Hackings onshore box, but a relatively good variation at Rottnest Island onshore box. There is almost no similarity by comparing these two graphs, apart from the lowest temperature of maximum intensity, which is around 1.4 ºC. As figure2.a shown, the overall fitted line is still declining along time goes, but the fitted line of maximum intensity at Rottnest Island in figure2.c keep going horizontally. As for the smooth lines of these two graphs, that of Port Hackings starts at relatively low max intensity in the 1990s, and climbs in the 2000s, then drops in the 2010s. However, the graph for the onshore site of Rottnest Island gives a concave line for a smooth line, the pit is between 2007 and 2008. The unique peak of maximum intensity for onshore Port Hackings appears in 2002, reaches 9 ºC which is much higher than the average fitted line. This phenomenon leads to a peak of mean intensity in 2002. The maximum intensity of Rottnest Island in 2000 and 2012 reaches about 3.2 ºC, which is much higher than normal max intensity, which results in the peak and sub-peak for the mean intensity graph. Both MHWs max intensity at offshore Port Hackings and Rottnest Island are similar to the above mean intensity graphs. In figure 2.b, the red fitted line is slightly climbing, which is associated with the mean intensity of MHWs. As for offshore Rottnest Island, there is a negative relationship between MHWs maximum intensity and years. The peak in figure 2.b reached 8.5 ºC in 2005, which is one Celsius higher than the peak in figure 2.d. However, it is worth noting that both two graphs touched the nadir around 2010.

5.3 Figure3 "MHW per year"

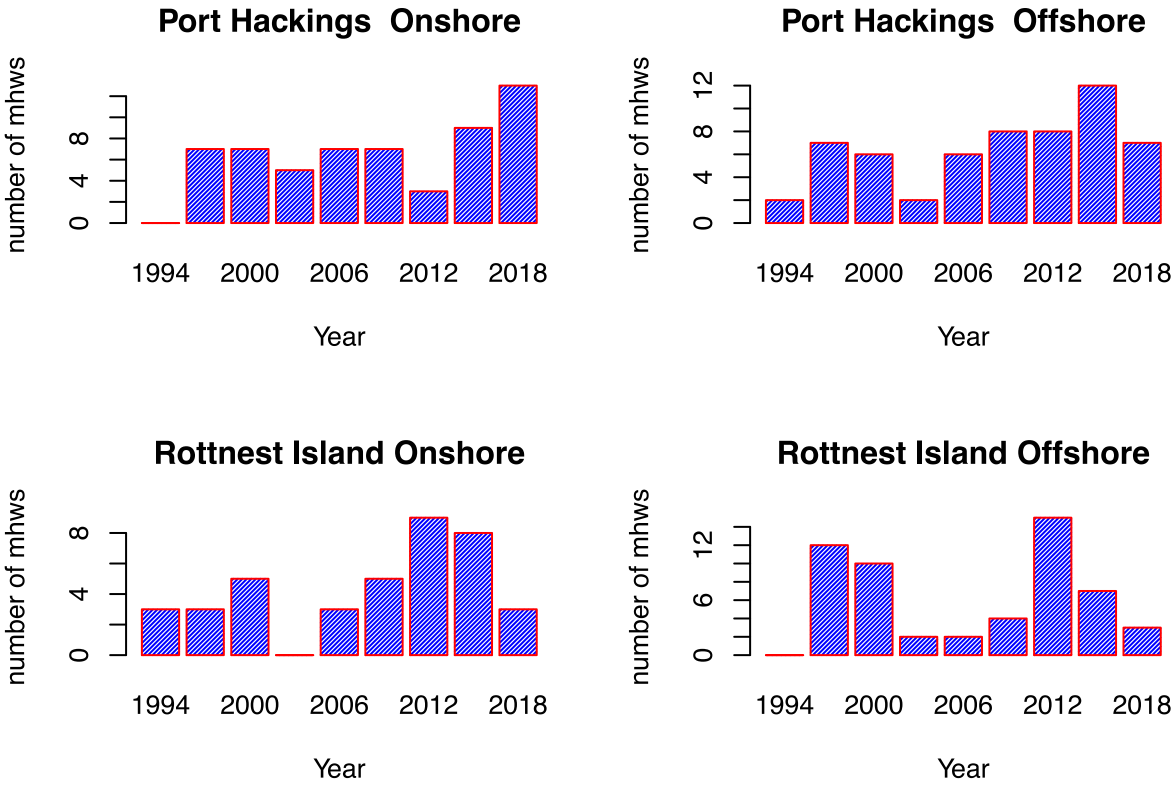


[A port hacking onshore; B port hacking offshore; C Rottnest island onshore; D Rottnest island offshore]

From Figure 3, it is clear that Port Hacking has experienced a more frequent MHWs when compared with Rottnest Island. Moreover, figure 3.a shows that in 2019, port Hacking has 6 MHWs at an onshore location, while only 1 MHWs happened at the offshore place. When looking at Rottenest Island, after experienced 5 times MHWS, which is the most frequent MHWs year. From 2000 to 2005, there are no MHWs in Figure 3.c at all. The annual frequencies of MHWs at Port Hacking onshore box and Rottnest Island onshore box are both around three events. There are five years with zero MHW at Port Hackings and eleven years at Rottnest Island, this may be caused by missing data record or just no MHW. Neither of the two bar charts shows an obvious trend. The maximum frequency of MHWs at Port Hacking appears in 2019, and that at Rottnest Island appears in 1999, which are 6 events and 5 events respectively.

By comparing figure 3.b and figure 3.d, it is clear to notice that the total number of MHWs at offshore Port Hackings is much more than the total number of MHWs at offshore Rottnest Island. There is no MHW in 1993, 1999, 2002, and 2008 in figure 3.b, yet the number of no MHWs in figure 3.d is triple. There is one peak at offshore Rottnest box in 2011, reached 9 events, which is higher than the peak at offshore Port Hackings (5 events). From figure 3.b, it is noticed that not only the intensity at offshore Port hackings is predicted to increase, but also frequency.

5.4 Figure4 "MHW 3 years"

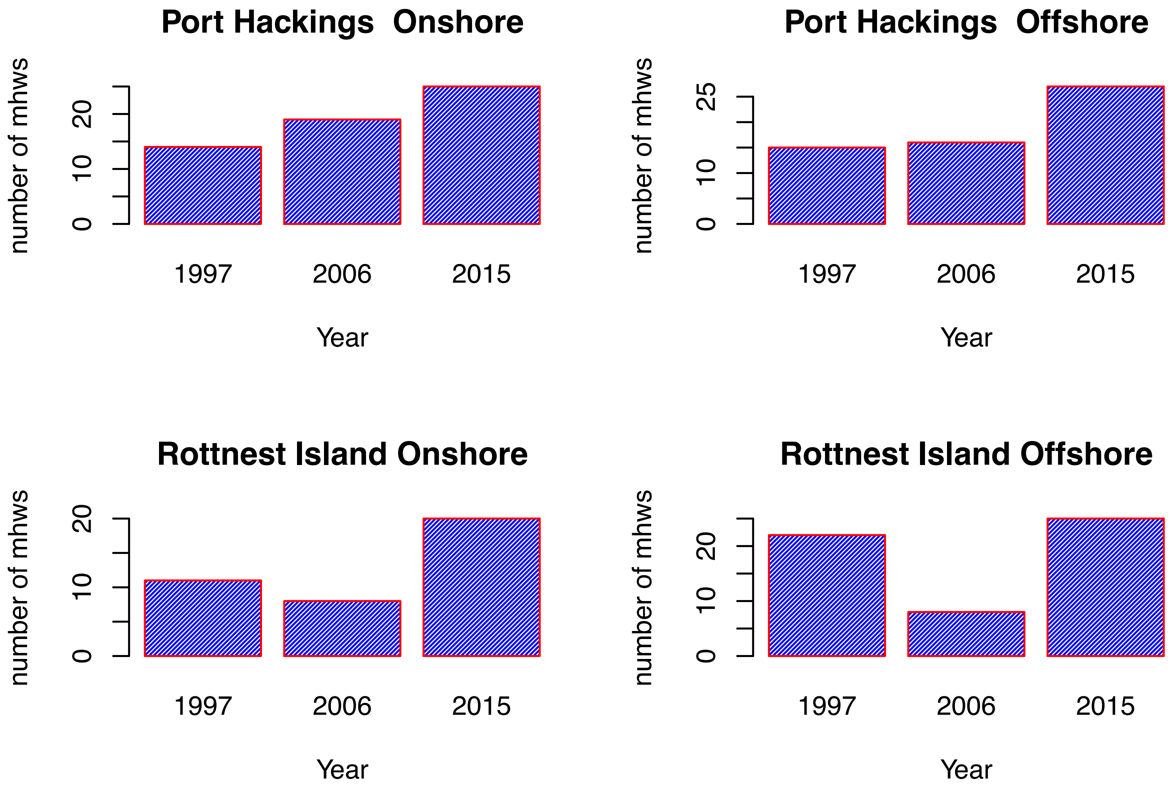


[A port hacking onshore; B port hacking offshore; C Rottnest island onshore; D Rottnest island offshore]

In Figure 4.a and 4.d, no MHW has been recorded between 1993 and 1995, and no MHW at Rottnest onshore from 2002 to 2004. From 2017 to 2019, there are 11 MHW events at Port Hacking, which is the ceiling of figure 4.a. As for Rottnest Island, the peak appears between 2011 and 2013, reaches 9 events. The averages of the frequency of MHWs per three years at Port Hacking and Rottnest are 7 and 5 events respectively.

Moreover, for offshore location, it is clearly shown that the frequency of MHWs per three years at offshore Port Hackings is more stable than that at offshore Rottnest Island, since no zero MHW event. The maximum number of MHWs at Rottnest is two events more than that at Port Hackings, reached 14 events per three years. Even if the distributions of the two graphs are slightly different, the average number of MHWs per three years is similar, which is about 6 events.

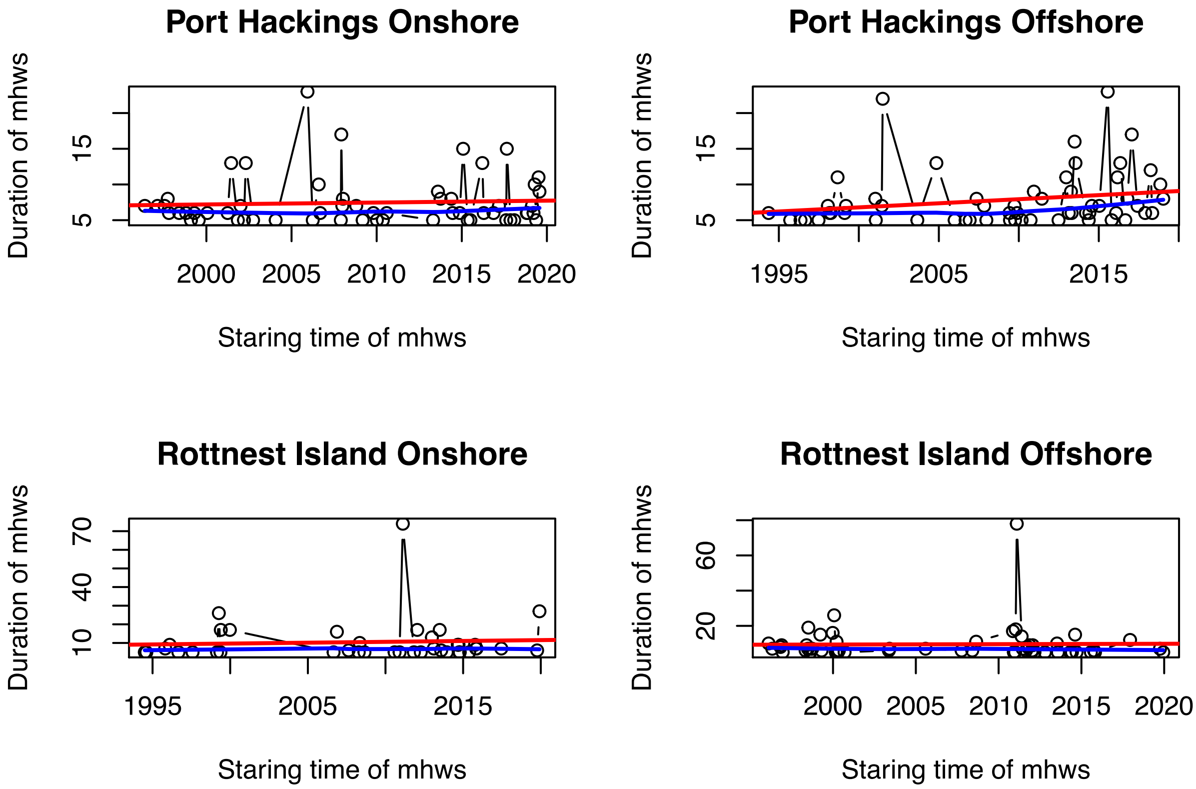
5.5 Figure5 "MHW 9 years"



[A port hacking onshore; B port hacking offshore; C Rottnest island onshore; D Rottnest island offshore]

Bar charts of the frequency of MHWs every 9 years are also computed. The tendency of MHWs' frequency can be found easily when seen in Figure 5. Due to the period limit, only 3 bars can be produced. The distributions of the first 9 years and last 9 years at two onshore boxes are quite similar. The peaks for Port Hackings and Rottnest are 25 events and 20 events, respectively. The pit (14 events) at Port Hackings appears in the first 9 years, but the pit of figure 5.c is from 2002 to 2010, which is 7 MHW events. From these two charts, it is estimated that the frequencies at the two sites would keep increasing. Figure 5.b and figure 5.d are bar charts to compute the number of MHWs per nine years. In the first nine years, it is reported that the frequency of nine years at Rottnest Island was 7 events more than the frequency of MHWs in the first nine years at Port Hackings. However, it reversed between 2002 and 2010, the frequency at Port Hackings is 8 events more than the frequency at Rottnest. Until the last nine years, the frequency of MHWs is the same in both places. Since there are only a few MHWs from 2000 to 2009 at Rottenness Island, so there is no clear linear relationship.

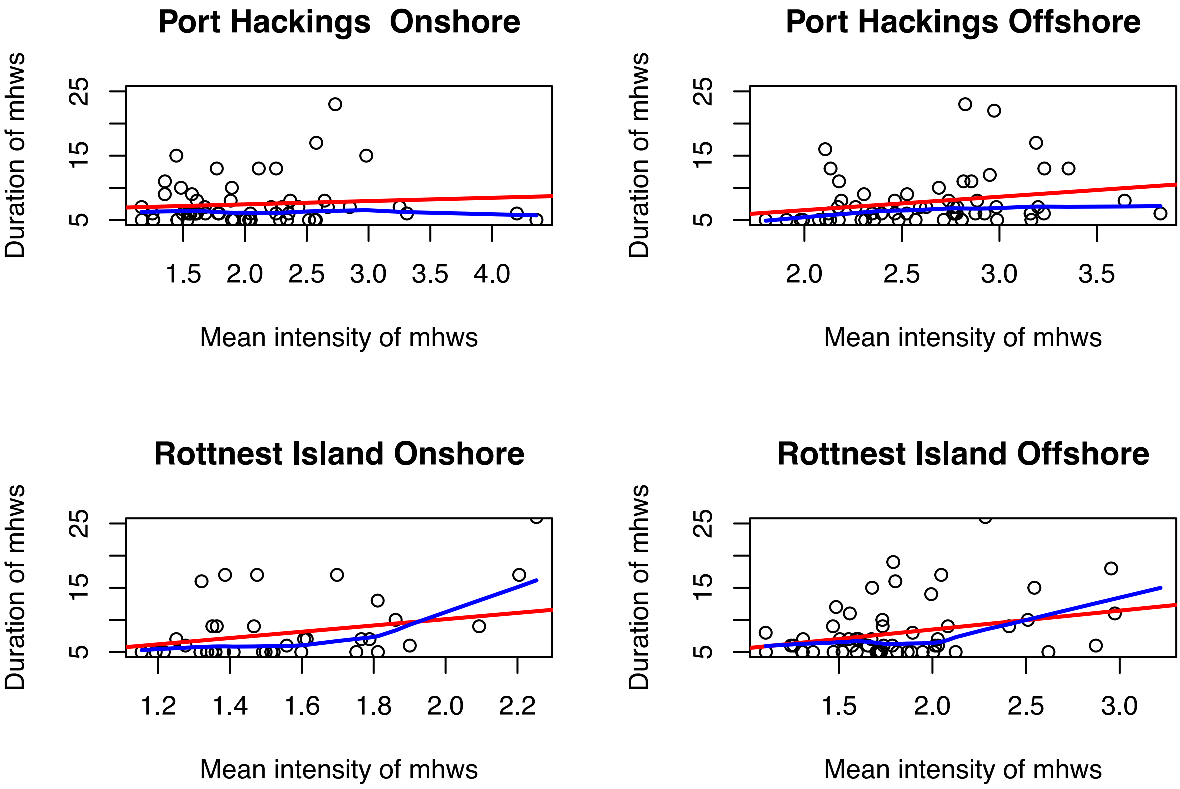
5.6 Figure6 "plot of duration against date"



[A port hacking onshore; B port hacking offshore; C Rottnest island onshore; D Rottnest island offshore]

Then it comes to the duration of MHWs. From figure 6, only 6.b shows a positive relationship over time, the other 3 locations all have a quit stable tendency, which is horizontally changed. From figure 6.a, it is clearly shown that it has a large variation in MHWs duration, even if its fitted line and smooth line go smoothly, the average duration is around 7 days. As for Rottnest Island, the variation is relatively small, except for one MHW longer than 70 days in 2012, which is estimated as an outlier. The fitted line of MHWs duration at Rottnest goes slightly increasing along 10 days. The peak of onshore Port Hacking MHWs duration appears in 2006, which lasts around 25 days. Both graphs reach their nadir at 5 days, due to the definition of MHWs. For figure 6.b, there is a positive relationship between duration and years, but at offshore Rottnest Island, the fitted line goes horizontally. Besides, the interval for figure 6.d is much larger than figure 6.b, since there is one event with 80 days duration in 2012 while most of the MHWs have a duration less than 20 days. On the right side of figure 6.b, events are evenly distributed, but in figure 6.d, most of them were distributed at the bottom. Hiding the peak point in 2012, distributions of the two graphs are slightly similar.

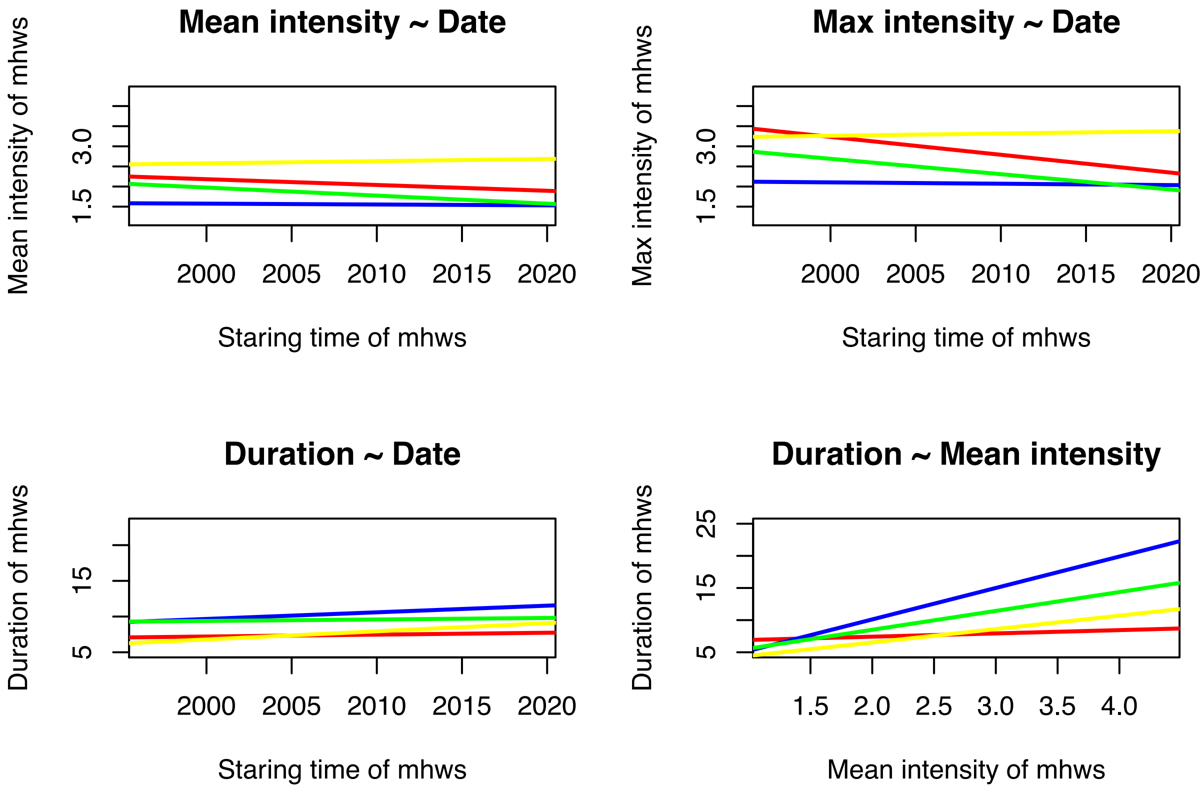
5.7 Figure7 “plot of duration against intensity”



[A port hacking onshore; B port hacking offshore; C Rottnest island onshore; D Rottnest island offshore]

Moreover, Figure7 shows the relationship between duration and mean intensity of MHWs. Both graphs at onshore Port Hackings and Rottnest Island have a positive linear relationship according to red fitted lines. Due to the high variation in duration at Port Hackings, many more points concentrate at the bottom left area but above the fitted line than that at Rottnest. Obviously, even lots of events have a high mean intensity, they still lie on the bottom of graphs in figure 7.a and figure 7.c. However, there is one event with both high mean intensity and long duration happened in 2012, which leads to a climbing deviation for red fitted line and blue smooth line in figure 7.c. Figure 7.b and figure 7.d describe the relationship between the duration of MHWs and the mean intensity of MHWs at offshore Port Hackings and offshore Rottnest Island. It is worth to mention that they are quite similar. Both of them have a duration interval between 0 and 25 days, a mean intensity interval between 0 and 3.5 ^{o}C, as well as the increasing fitted line.

5.8 Figure8



[E: plot of mean intensity against date; F:plot of max intensity against date; G:plot of duration against date; H:plot of duration against intensity]

In Figure8.e, the red line has a negative slope, and the yellow line has a positive slope. And total mean intensity offshore is above onshore's though their starting points are close to each other. Smooth lines and fitted lines in figures of "maximum intensity VS date" have similar shapes as "mean intensity VS date" but with higher intensity. In Figure8.f, the red line starts with a higher point than the yellow line. As the red line decreasing and the yellow line increasing, they across at the year 2000. After this point, the total max intensity offshore is above onshore's.

In Figure8.e, the blue line has a flat slope, and the green line has a negative slope. And total mean intensity onshore is above offshore's. Smooth lines and fitted lines in figures of "maximum intensity VS date" have similar shapes as "mean intensity VS date" but with higher intensity. In Figure8.f, the blue line starts with a lower point than the green line. As the green line decreasing and the blue line remaining flat, they across at the year 2017. After this point, the total max intensity onshore is above offshore's.

Chapter 6

Discussion

Through the study, the three characteristics of MHWs did not show an obvious trend for the past 27 years, but there are some extreme values of intensity and duration that needed more attention.

According to Figure 1 in Chapter 5, from 2001 the mean intensity of the marine heatwaves has increased dramatically, and at the end of 2001 and the start of 2002, there are two MHWs with maximum intensities of 7.56 and 9.15 degrees Celsius happened in the Port Hacking onshore box, which is much higher than the maximum intensities of other MHWs in this box. According to the website "Conversation", years 2000 to 2009 was the warmest decade since modern record began. The main reason for this unpredicted warmth is greenhouse gases produced by human beings[9]. Warming has been dominated by increased carbon dioxide levels. In the long term, scientists have already proved that the surface of the earth, as measured by global mean temperature, has warmed by about one degree Celsius during the past century.[9] From Figure 3 – 5 in Chapter 5, the overall tendency also indicates that unusually hot days and nights will happen more and more often.

Moreover, the records of intensity show a periodical pattern. According to a meteorology study in Australia climate, temperatures are impacted by La Nina and El Nino periodically. [15] Nino and La Nina are the opposite phases of the El Niño–Southern Oscillation (ENSO). La Nina is referred to as the cold phase of ENSO and El Nino is the warm phase of ENSO. These deviations have large impacts on ocean processes and caused circulation changes in sea-water temperatures. They typically last 9 to 12 months, but some can last for years. The frequency of La Nina and El Nino to occur is irregular, but they averagely happen in every 2 to 7 years, and La Nina happened after El Nino mostly.

According to the Australian government— Bureau of meteorology "La Nina — Detailed Australian Analysis(refference format wrong)": 1998-2001, 2007-2009 and 2010-2012 are La Nina years[10] [11] [12]; 1997-1998, 2009-2010 and 2015-2016 are El Nino years. There was a gap of 6 years in Rottnest Island onshore from 1999 to 2006, during which there was not a single MHW event detected. The period of this gap matches La Nina years in Australia, which can explain why no MHW occurs during these 6 years. "Marine Climate Change in Australia" written by Ming Feng, Evan Weller, Katy Hill refers[13] that extremely high temperatures over the past 30 years have mostly occurred in autumn-winter of the west coast, and increase at about 0.02-0.035°C per year, while there was little or no increase in temperatures in the spring-summer period compared to other years. This has caused a delay in seasonal cycles of surface temperatures of the west coast by 10 to 20 days.This explains that the impact of La Nina phenomenon on the sea surface temperature is not only on La Nina years but also on upcoming years.

The existence of La Nina and El Nino years also explains extreme values of duration. According to Figure6.c and Figure6.d in Chapter 5, there are 2 MHWs have especially long duration in 2011. The onshore box of Rottnest Island has a duration of 74 days with only 7 days' data missing, and the offshore box has a duration of 78 days with only 13 days' data missing, and the longest gap in these 2 MHWs is of 2 days, which means even if the temperatures in gaps are below the threshold, they are still considered as parts of this MHW by definition. Therefore, these 2 especially long MHWs are not mistakenly detected due to the missing temperatures but exist in fact. Thomas Werngerg confirms this unusual event in "Marine heatwave drives collapse of kelp forests in Western Australia".[14] In 2011, Western Australia experienced an extreme MHW, affecting larger than 2,000 km of coastline for more than 10 weeks. According to the research, almost half of the kelp perished and turf algae expanded rapidly. Even though the water temperature has already returned to pre-MHW levels, the kelp forests have not recovered even after 8 years.

Chapter 7

Conclusion and Further Issues

Though out the whole analysis, it is found that for both Port hacking and Rottnest Island site, the characteristics of MHWs on both onshore box and the offshore box does not display any significant difference that will affect the analysis. Minor variations do exist but have no impact on the overall tendency of the changes of MHWs. But data on Port Hackings differs from data from Rottnest Island vastly, suggesting that MHWs could be affected by geographical location.

Despite the existence of limited abnormally high intensity on both Rottnest Island and Port hacking site, the overall tendency remains steady and does not show any trend of increasing. However, it is worth notice that if one is to look at the intensity from 2000 and onward only, there is a tendency of increment in both the max intensity and the mean intensity at Port Hackings site, suggesting a possibility of future increases. Future works could be done to monitor the value of that site.

On both sites, there is a general tendency for the duration of MHWs to increase. It is worth pointing out that there exists a MHW with duration longer than 70 days. In the analysis, although it is not a measurement error, it is regarded as the outlier and removing it does change the conclusion.

As for frequency, since it is measured at different scales, the result differs based on the scale of measurement. In general, the frequency on Port Hackings is steady with a tendency of slightly increasing. On the other hand, since there are gaps of MHWs for 6 consecutive years on Rottnest Island, the conclusion is likely to be affected by the scales of analysis. if the frequency is calculated for every year or every 10 years, there is no significant trend that could be used for analysis. but if the frequency is calculated for every 3 years, the frequency histogram displays a periodical tendency, where the value of frequency oscillates up and down. Due to the limitation of numbers of data available, this potential periodical feature of frequency cannot be investigated future. Further work could be done in this area to investigate this matter.

References

[1] Earth Systems and Climate Change Hub 2019. Dr.Kathleen McInnes, CSIRO. 2019, April 5. "Marine heatwaves: changes, causes and impacts". Earth Systems and Climate Change Hub.

[2] Hobday, A.J., E.C.J. Oliver, A. Sen Gupta, J.A. Benthuysen, M.T. Burrows, M.G. Donat, N.J. Holbrook, P.J. Moore, M.S. Thomsen, T. Wernberg, and D.A. Smale. 2018. "Categorizing and naming marine heatwaves". Oceanography 31(2)

[3] All about MHWs. (n.d.). "Marine Heatwaves Explained," All about MHWs. Retrieved September 18, 2020.

[4] Hobday, A.J., Alexander, L.V., Perkins, S.E., Smale, D.A., Straub, S.C., Oliver, E.C.J., Benthuysen, J.A., Burrows, M.T., Donat, M.G., Feng, M., Holbrook, N.J., Moore, P.J., Scannell, H.A., Sen Gupta, A. and Wernberg, T. 2016. "A Hierarchical Approach to Defining Marine Heatwaves". Progress in Oceanography, [online] 141(141), pp.227–238.

[5] Oliver, E.C.J., Donat, M.G., Burrows, M.T., Moore, P.J., Smale, D.A., Alexander, L.V., Benthuysen, J.A., Feng, M., Sen Gupta, A., Hobday, A.J., Holbrook, N.J., Perkins-Kirkpatrick, S.E., Scannell, H.A., Straub, S.C. and Wernberg, T. 2018, "Longer and more frequent marine heatwaves over the past century," Nature Communications, vol. 9, no. 1.

[6] Schaeffer, A. and Roughan, M. (2017). Subsurface intensification of marine heatwaves off southeastern Australia: The role of stratification and local winds. Geophysical Research Letters, 44(10), pp.5025–5033.

[7] Australia's Integrated Marine Observing System n.d., "Description of IMOS" Australia's Integrated Marine Observing System.

[8] Australia's Integrated Marine Observing System n.d., "How to read the data," Australia's Integrated Marine Observing System.

[9] Neil Plummer, Blair Trewin, David Jones, Karl Braganza, Rob Smalley. 2013, "What's causing Australia's heat wave?," The Conversation.

[10] Australian government Bureau of Meteorology 2016, "What is La Niña and how does it impact Australia?," Australian government Bureau of Meteorology.

[11] Australian government Bureau of Meteorology n.d., "La Niña – Detailed Australian Analysis," Australian government Bureau of Meteorology.

[12] Heidemann, Hanna 2019, "Marine Heat Waves and the Influence of El Niño off Southeast Queensland, Australia," Frontiers in Marine Science, vol. 6.

[13] Ming Feng, Evan Weller, Katy Hill.2009,"Marine climate change in Australia[Impact and Adaptation Response],"The Leeuwin Current

[14] Wernberg T (2020) Marine heatwave drives collapse of kelp forests in Western Australia. In: Canadell JG, Jackson RB (eds) Ecosystem Collapse and Climate Change. Ecological Studies, Springer-Nature, accepted 1 Sept 2019.

[15] National ocean service n.d., "What are El Nino and La Nina?," National ocean service.

[16] Benthuysen, J.A., Oliver, E.C.J., Chen, K. and Wernberg T. 2020, "Editorial: Advances in Understanding Marine Heatwaves and Their Impacts," Frontiers in Marine Science.

[17] Holbrook, N.J., Sen Gupta, A., Oliver, E.C.J., Hobday, A.J., Benthuysen, J.A., Scannell, H.A., Smale, D.A. and Wernberg, T. 2020, "Keeping pace with marine heatwaves," Nature Reviews Earth and Environment, vol. 1, no.9, pp.482–93.

[18] Lee, K.A., Roughan, M., Malcolm, H.A. and Otway, N.M. 2018, "Assessing the Use of Area- and Time-Averaging Based on Known De-correlation Scales to Provide Satellite Derived Sea Surface Temperatures in Coastal Areas," Frontiers in Marine Science, vol.5.

[19] Mazdiyasni, O., Sadegh, M., Chiang, F. and AghaKouchak, A. 2019, "Heat wave Intensity Duration Frequency Curve: A Multivariate Approach for Hazard and Attribution Analysis," Scientific Reports, vol. 9, no. 1.

[20] Thompson, P., Baird, M., Ingleton, T. and Doblin, M. 2009, "Long-term changes in temperate Australian coastal waters: implications for phytoplankton," Marine Ecology Progress Series, vol. 394, pp. 1–19.

[21] Sparnocchia, S., Schiano, M.E., Picco, P., Bozzano, R. and Cappelletti, A. 2006, "The anomalous warming of summer 2003 in the surface layer of the Central Ligurian Sea (Western Mediterranean)," Annales Geophysicae, vol.24, no.2, pp.443–52.

[22] Oliver, E.C.J., Benthuysen, J.A., Bindoff, N.L., Hobday, A.J., Holbrook, N.J., Mundy, C.N. and Perkins-Kirkpatrick, S.E. 2017, "The unprecedented 2015/16 Tasman Sea marine heatwave," Nature Communications, vol.8, no. 1.

[23] Chen, K., Gawarkiewicz, G.G., Lentz, S.J. and Bane, J.M. 2014, "Diagnosing the warming of the Northeastern U.S. Coastal Ocean in 2012: A linkage between the atmospheric jet stream variability and ocean response," Journal of Geophysical Research: Oceans, vol. 119, no. 1, pp. 218–27.

[24] Australia's Integrated Marine Observing System n.d., "SRS-SST Data," Australia's Integrated Marine Observing System.

[25] Australian government Bureau of Meteorology 2011, "Annual Australian Climate Statement 2010," Australian government Bureau of Meteorology

[26] Australian Government Bureau of Meteorology n.d., "Climate Glossary - Seasons," Commonwealth of Australia Bureau of Meteorology.

[27] Nogrady, B. 2013, "Late 20th century hottest in over 1000 years," ABS Science.

[28] Earth Institute Columbia University 2010, "2000-2009: The Warmest Decade," Earth Institute Columbia University

[29] Vorwerk, R. 2017, "El Niño and La Niña – por que no los dos?! (Actually, that would be very bad) - CSIROscope," CSIROscope.

Appendix

All the code and CSV files have been uploaded to this link.
https://github.com/GrassssC/Team-9-Marine-Heat-Waves

Codes

```
Port_box1 = read.csv("Port_Hacking_onshore_analys.csv",header = TRUE)
P1_freq_9year = read.csv("average_9_year_Port_Hacking_onshore_analys.csv",head
P1_freq_3year = read.csv("average_3_year_Port_Hacking_onshore_analys.csv",head
P1_freq_peryear = read.csv("average_1_year_Port_Hacking_onshore_analys.csv",he

Port_box2 = read.csv("Port_Hacking_offshore_analys.csv",header = TRUE)
P2_freq_9year = read.csv("average_9_year_Port_Hacking_offshore_analys.csv",hea
P2_freq_3year = read.csv("average_3_year_Port_Hacking_offshore_analys.csv",hea
P2_freq_peryear = read.csv("average_1_year_Port_Hacking_offshore_analys.csv",h

Rottnest_box1 = read.csv("Rottnest_Island_onshore_analys.csv",header = TRUE)
R1_freq_9year = read.csv("average_9_year_Rottnest_Island_onshore_analys.csv",h
R1_freq_3year = read.csv("average_3_year_Rottnest_Island_onshore_analys.csv",h
R1_freq_peryear = read.csv("average_1_year_Rottnest_Island_onshore_analys.csv"

Rottnest_box2 = read.csv("Rottnest_Island_offshore_analys.csv",header = TRUE)
R2_freq_9year = read.csv("average_9_year_Rottnest_Island_offshore_analys.csv",
R2_freq_3year = read.csv("average_3_year_Rottnest_Island_offshore_analys.csv",
R2_freq_peryear = read.csv("average_1_year_Rottnest_Island_offshore_analys.csv
```

Figure1 "plot of mean intensity against date"

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))

plot(x=as.Date(Port_box1$time_start), y=Port_box1$mean_intensity,xlab="Staring
P1_mean_intensity_date <-lm(Port_box1$mean_intensity ~ as.Date(Port_box1$time_
abline(P1_mean_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Port_box1$time_start),Port_box1$mean_intensity), col = "b

plot(x=as.Date(Port_box2$time_start), y=Port_box2$mean_intensity,xlab="Staring
P2_mean_intensity_date <-lm(Port_box2$mean_intensity ~ as.Date(Port_box2$time_
abline(P2_mean_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Port_box2$time_start),Port_box2$mean_intensity), col = "b

plot(x=as.Date(Rottnest_box1$time_start), y=Rottnest_box1$mean_intensity,xlab=
```

```
R1_mean_intensity_date <-lm(Rottnest_box1$mean_intensity ~ as.Date(Rottnest_bo
abline(R1_mean_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Rottnest_box1$time_start),Rottnest_box1$mean_intensity),

plot(x=as.Date(Rottnest_box2$time_start), y=Rottnest_box2$mean_intensity,xlab=
R2_mean_intensity_date <-lm(Rottnest_box2$mean_intensity ~ as.Date(Rottnest_bo
abline(R2_mean_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Rottnest_box2$time_start),Rottnest_box2$mean_intensity),
```

Figure2 “plot of max intensity against date”

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))

plot(x=as.Date(Port_box1$time_start), y=Port_box1$max_intensity,xlab="Staring
P1_max_intensity_date <-lm(Port_box1$max_intensity ~ as.Date(Port_box1$time_st
abline(P1_max_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Port_box1$time_start),Port_box1$max_intensity), col = "bl

plot(x=as.Date(Port_box2$time_start), y=Port_box2$max_intensity,xlab="Staring
P2_max_intensity_date <-lm(Port_box2$max_intensity ~ as.Date(Port_box2$time_st
abline(P2_max_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Port_box2$time_start),Port_box2$max_intensity), col = "bl

plot(x=as.Date(Rottnest_box1$time_start), y=Rottnest_box1$max_intensity,xlab="
R1_max_intensity_date <-lm(Rottnest_box1$max_intensity ~ as.Date(Rottnest_box1
abline(R1_max_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Rottnest_box1$time_start),Rottnest_box1$max_intensity), c

plot(x=as.Date(Rottnest_box2$time_start), y=Rottnest_box2$max_intensity,xlab="
R2_max_intensity_date <-lm(Rottnest_box2$max_intensity ~ as.Date(Rottnest_box2
abline(R2_max_intensity_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Rottnest_box2$time_start),Rottnest_box2$max_intensity), c
```

Figure3 “MHW per year”

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))
barplot(as.table(setNames(P1_freq_peryear$frequency,P1_freq_peryear$start_year

barplot(as.table(setNames(P2_freq_peryear$frequency,P2_freq_peryear$start_year

barplot(as.table(setNames(R1_freq_peryear$frequency,R1_freq_peryear$start_year

barplot(as.table(setNames(R2_freq_peryear$frequency,R2_freq_peryear$start_year
```

Figure4 “MHW 3 years”

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))


barplot(as.table(setNames(P1_freq_3year$frequency,P1_freq_3year$centre_year)),
```

```
barplot(as.table(setNames(P2_freq_3year$frequency,P2_freq_3year$centre_year)),

barplot(as.table(setNames(R1_freq_3year$frequency,R1_freq_3year$centre_year)),

barplot(as.table(setNames(R2_freq_3year$frequency,R2_freq_3year$centre_year)),
```

Figure5 “MHW 9 years”

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))

barplot(as.table(setNames(P1_freq_9year$frequency,P1_freq_9year$centre_year)),

barplot(as.table(setNames(P2_freq_9year$frequency,P2_freq_9year$centre_year)),

barplot(as.table(setNames(R1_freq_9year$frequency,R1_freq_9year$centre_year)),


barplot(as.table(setNames(R2_freq_9year$frequency,R2_freq_9year$centre_year)),
```

Figure6 “plot of duration against date”

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))

plot(x=as.Date(Port_box1$time_start), y=Port_box1$duration,xlab="Staring time
P1_duration_date <-lm(Port_box1$duration~ as.Date(Port_box1$time_start))
abline(P1_duration_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Port_box1$time_start),Port_box1$duration), col = "blue",

plot(x=as.Date(Port_box2$time_start), y=Port_box2$duration,xlab="Staring time
P2_duration_date <-lm(Port_box2$duration~ as.Date(Port_box2$time_start))
abline(P2_duration_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Port_box2$time_start),Port_box2$duration), col = "blue",

plot(x=as.Date(Rottnest_box1$time_start), y=Rottnest_box1$duration,xlab="Stari
R1_duration_date <-lm(Rottnest_box1$duration~ as.Date(Rottnest_box1$time_start
abline(R1_duration_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Rottnest_box1$time_start),Rottnest_box1$duration), col =

plot(x=as.Date(Rottnest_box2$time_start), y=Rottnest_box2$duration,xlab="Stari
R2_duration_date <-lm(Rottnest_box2$duration~ as.Date(Rottnest_box2$time_start
abline(R2_duration_date, col = "red", lwd = 2, lty = 1)
lines(lowess(as.Date(Rottnest_box2$time_start),Rottnest_box2$duration), col =
```

Figure7 “plot of duration against intensity”

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))

plot(x=Port_box1$mean_intensity, y=Port_box1$duration,xlab="Mean intensity of
```

```
P1_mean_intensity_duration <-lm(Port_box1$duration ~ Port_box1$mean_intensity,
abline(P1_mean_intensity_duration,col = "red", lwd = 2, lty = 1)
lines(lowess(Port_box1$mean_intensity,Port_box1$duration), col = "blue", lwd =

plot(x=Port_box2$mean_intensity, y=Port_box2$duration,xlab="Mean intensity of
P2_mean_intensity_duration <-lm(Port_box2$duration ~ Port_box2$mean_intensity,
abline(P2_mean_intensity_duration,col = "red", lwd = 2, lty = 1)
lines(lowess(Port_box2$mean_intensity,Port_box2$duration), col = "blue", lwd =

plot(x=Rottnest_box1$mean_intensity, y=Rottnest_box1$duration,xlab="Mean inten
R1_mean_intensity_duration <-lm(Rottnest_box1$duration ~ Rottnest_box1$mean_in
abline(R1_mean_intensity_duration,col = "red", lwd = 2, lty = 1)
lines(lowess(Rottnest_box1$mean_intensity,Rottnest_box1$duration), col = "blue

plot(x=Rottnest_box2$mean_intensity, y=Rottnest_box2$duration,xlab="Mean inten
R2_mean_intensity_duration <-lm(Rottnest_box2$duration ~ Rottnest_box2$mean_in
abline(R2_mean_intensity_duration,col = "red", lwd = 2, lty = 1)
lines(lowess(Rottnest_box2$mean_intensity,Rottnest_box2$duration), col = "blue
```

Figure8

```
layout(matrix(c(1,2,3,4),nr=2,byrow=T))

plot(x=as.Date(Port_box1$time_start), y=Port_box1$mean_intensity,xlab="Staring
abline(P1_mean_intensity_date, col = "red", lwd = 2, lty = 1)
abline(P2_mean_intensity_date, col = "yellow", lwd = 2, lty = 1)
abline(R1_mean_intensity_date, col = "blue", lwd = 2, lty = 1)
abline(R2_mean_intensity_date, col = "green", lwd = 2, lty = 1)

plot(x=as.Date(Port_box1$time_start), y=Port_box1$mean_intensity,xlab="Staring
abline(P1_max_intensity_date, col = "red", lwd = 2, lty = 1)
abline(P2_max_intensity_date, col = "yellow", lwd = 2, lty = 1)
abline(R1_max_intensity_date, col = "blue", lwd = 2, lty = 1)
abline(R2_max_intensity_date, col = "green", lwd = 2, lty = 1)

plot(x=as.Date(Port_box1$time_start), y=Port_box1$duration,xlab="Staring time
abline(P1_duration_date, col = "red", lwd = 2, lty = 1)
abline(P2_duration_date, col = "yellow", lwd = 2, lty = 1)
abline(R1_duration_date, col = "blue", lwd = 2, lty = 1)
abline(R2_duration_date, col = "green", lwd = 2, lty = 1)

plot(x=Port_box1$mean_intensity, y=Port_box1$duration,xlab="Mean intensity of
abline(P1_mean_intensity_duration, col = "red", lwd = 2, lty = 1)
abline(P2_mean_intensity_duration, col = "yellow", lwd = 2, lty = 1)
abline(R1_mean_intensity_duration, col = "blue", lwd = 2, lty = 1)
abline(R2_mean_intensity_duration, col = "green", lwd = 2, lty = 1)
```

Tables

Box Index:

1: Port Hacking Onshore Box
2: Port Hacking Offshore Box
3: Rottnest Island Onshore Box
4: Rottnest Island Offshore Box

Table1:Description of variables:

Index	Variable name	Dimensions	Coordinates	Type	Units	Long name	Comment
1	lat	lat	\	float32	degrees_north	latitude	Latitudes for locating data
2	lon	lon	\	float32	degrees_east	longitude	Longitudes for locating data
3	time	time	\	int32	seconds since 1981-01-01 00:00:00	reference time of sst file	A typical reference time for data
4	sea_surface_temperature	time,lat,lon	time,lat,lon	float32	kelvin	sea surface skin temperature	The skin temperature of the ocean at a depth of approximately 10um
5	sst_dtime	time,lat,lon	time,lat,lon	float32	second	time difference from reference time	time plus sst_dtime gives seconds after 00:00:00 UTC January 1, 1981
6	dt_analysis	time,lat,lon	time,lat,lon	float32	kelvin	deviation from last SST analysis	The difference between this SST and the previous day's SST
7	satellite_zenith_angle	time,lat,lon	time,lat,lon	float32	angular_degree	satellite zenith angle	The satellite zenith angle at the time of the SST observations
8	l2p_flags	time,lat,lon	time,lat,lon	int16	\	L2P flags	These are used to reflect various ancillary conditions. This allows decisions on data quality related to ancillary fields to be made by reading a bit array rather than processing the ancillary fields.
9	quality_level	time,lat,lon	time,lat,lon	int8	\	quality level of SST pixel	These are the overall quality indicators and are used for all GHRSST SSTs. In this case they are a function of distance to cloud, satellite zenith angle, and day/night. Flag values: 0 = no data, 1 = bad data, 2 = worst quality, 3 = low quality, 4 = acceptable quality, 5 = best quality.
10	sses_bias	time,lat,lon	time,lat,lon	float32	kelvin	SSES bias estimate	Bias estimate derived from L2P bias
11	sses_standard_deviation	time,lat,lon	time,lat,lon	float32	kelvin	SSES standard deviation estimate	Standard deviation estimate derived from L2P standard deviation
12	sses_count	time,lat,lon	time,lat,lon	float32	count	SSES count	Weighted representative number of swath pixels.

Table2: Selection of boxes:

Box Index	Latitude	Longitude	N SSTs	N good SSTs	Ave. SSTs	Ave. good SSTs
1	1..5	132..136	5691	4701	6779	5769
2	201..205	195..199	4849	3347	6536	4808
3	90..94	128..132	6478	5396	7657	6491
4	0..4	25..29	5623	4301	7116	5586

Latitude: The range of latitude index (index of the variable "lat") in the box.

Longitude: The range of longitude index (index of the variable "lon") in the box.

N SSTs: Average of numbers of SSTs at all pixels in the box.

N good SSTs: Average of numbers of SSTs with a quality level 3 or higher at all pixels in the box.

Ave. SSTs: Number of SSTs in the box after filling gaps with spatial average.

Ave. good SSTs: Number of SSTs with a quality level 3 or higher in the box after filling gaps with spatial average.

Table3: Number of Gaps:

Box Index	Missing days	Gaps	Max gap length	Small gaps	Medium gaps	Large gaps
1	3082	1669	19	1318	266	85
2	3324	1859	21	1497	282	80
3	2208	1427	17	1260	135	32
4	2745	1639	17	1382	212	45

Missing days: Number of days with no SST records out of 9861 days.

Gaps: Number of Gaps in SSTs from 1993-01-01 to 2019-12-31.

Max gap length: The length of the largest gap in days.

Small gaps: Number of gaps with length <= 2 days.

Medium gaps: Number of gaps with length > 2 days and <= 4 days.

Large gaps: Number of gaps with length >= 5 days.

Table4:Numerical Summaries

Box Index	variable name	unit	number of obs.	minimum	1st Quantile	median	mean	3rd Quantile	Maximum	standard deviation
1	duration	day	58	5	5	6	7.448	8	23	3.509
1	max intensity	°C	58	1.2	2.025	2.382	3.822	3.14	9.155	1.409
1	mean intensity	°C	58	1.164	1.555	1.904	2.049	2.363	4.359	0.672
1	cumulative intensity	°C	58	5.83	9.574	12.43	15.494	18.871	62.808	10.02
2	duration	day	58	5	5	6	7.828	8	23	3.962
2	max intensity	°C	58	1.805	2.635	3.169	3.313	3.646	8.604	0.964
2	mean intensity	°C	58	1.801	2.27	2.657	3.626	2.883	3.826	0.438
2	cumulative intensity	°C	58	9.007	13.688	17.387	20.953	22.901	65.406	12.243
3	duration	day	39	5	5	6	10.46	9.5	74	11.762
3	max intensity	°C	39	1.255	1.7	1.992	2.075	2.444	3.277	0.515
3	mean intensity	°C	39	1.155	1.351	1.499	1.556	1.759	2.253	0.271
3	cumulative intensity	°C	39	5.773	7.257	9.349	17.432	18.726	139.881	22.758
4	duration	day	55	5	5	7	9.527	9.5	78	10.311
4	max intensity	°C	55	1.177	1.803	2.119	2.429	2.522	7.297	1.119
4	mean intensity	°C	55	1.109	1.556	1.733	1.838	2.021	3.218	0.468
4	cumulative intensity	°C	55	5.553	8.967	11.146	19.943	17.576	250.994	33.362