5011CEM BIG DATA PROGRAMMING PROJECT

EUROPEAN OZONE VISUALISATION REPORT

Submission Date: 24th April 2020

TABLE OF CONTENTS

INTRODUCTION

The client is involved in monitoring European Ozone levels and has collected new 'Cluster Based Ensembles', which they wish to display in a clear, simple and easy to interpret format. The volume of data the client has gathered spans numerous years, with the client requesting a clear way to visualize their results, to help understand the variation between their models. Due to the vast amount of data involved, it has been agreed that the code should load and display the CBE for 24 hour, 1-hour intervals, spanning a given day. The data will be displayed on-top of a European map, showing basic geographical rivers, lakes, cities and coastlines, with the CBE results being visualized by a fixed colour scale, representing the volume of ozone recorded at each geolocation. Each CBE hour will be displayed separately, starting from the first 1-hour interval, it then has been requested to implement a simulation animation, as well as the ability to individually step through the results. This has been agreed to be implemented via two methods which allow for a choice between, a "step mode" or a "movie mode", the first being a method to increment and decrement through all the 1-hour intervals manually, and the latter method being an automatic loop, displaying each 1-hour interval incrementally. A further requirement requested by the client is for the code to include an option for altering the display of the data to be suitable for colour blindness, it has been stressed that this is a very important feature, due to their own client's agenda.

SPECIFICATION

The features requested by the ECMWF include, the ability to zoom on the map to gain a more accurate ozone visualisation of the local area, and two methods which allow for a choice between, a "step mode" or a "movie mode", the first being a method to increment and decrement through all the 1-hour intervals manually, and the latter method being an automatic loop, displaying each 1-hour interval incrementally. A colour-blind mode has also been specifically asked for, due to ECMWF's own target audience.

SMART Targets:
1. All ozone 1-hour intervals will be displayed on a map of the entire European continent, showing all major geographical rivers, lakes, cities and coastlines.
2. Each ozone value will be visualized by a fixed standard colour or a colour-blind scale, representing the volume of the ozone recorded at each geolocation.
3. Each 1-hour ozone interval should be loaded and displayed onto the map within 0.5 seconds.
4. The figure plot should allow for zoom. This will be a zoom resolution of up to the maximum, 1 data plot pixel (individual ozone value at any given geolocation).
5. A "step mode" should load and plot the first ozone 1-hour interval onto the map, and then allow for a manual keyboard press to re-plot the ozone layer, to either the next or previous 1-hour interval. When displaying the first 1-hour interval, it should loop back to the last 1-hour interval when the previous key is pressed. When displaying the last 1-hour interval, it should loop back to the first 1-hour interval when the next key is pressed.
6. A "movie mode" should load and plot the first ozone 1-hour interval onto the map, then re-plot the ozone layer to the next 1-hour interval, after a 0.2 second pause. This is to create an animation effect, with all the 1-hour intervals being displayed incrementally, looping back to the first 1-hour interval when the last interval is reached.
7. The program should display two questions upon start-up. The first being a choice of colourblind mode or standard mode, and the latter being a choice of "step mode" or "movie mode". These questions should only allow for 'Yes' or 'No' answers, and the entirety of the question should be displayed within a separate window upon start-up of the program.

CODE

FLOWCHARTS

Before each feature was developed, a flowchart was created, to help understand the process flow. (Appendix B, C, D, E, F).

PROCESSES

There are 4 functions within the script; main, settings, stepMode and movieMode (Appendix G).

CHOICES

COLOUR SCHEME

The project requirements state, a standard colour scheme and a colorblind scheme is needed. Therefore, the colour scheme 'jet' was chosen for the standard scheme, due to its clear colour ranges. However, the 'jet' colour scheme includes green and red which, red-green coluorblindness cannot perceive. Therefore, the 'bone' colour scheme was chosen for the alternate colour scheme, due to it containing no colours and thus, will be suitable for all types of colourblindness.

For further choices… (Appendix I).

TARGETS

The code archives the SMART targets set due to… (Appendix H)

STRUCTURE

The entire code is located within in one file which is broken down into functions, allowing for easy adaptation of variables and features, due to each feature being encapsulated within its own function. Thus, it is simple re-use and adapt functions if a similar feature is needed, such as a adding a third mode of displaying the data would just require either copying the 'stepMode' or 'movieMode' function and modifying it slightly. This is exactly how the 'movieMode' function was created, 'stepMode' was copied and modified to create 'moveiMode', as they are very similar. Another example of easy adaptation would be, if a new colour scheme was needed, it would just require a new colourmap to be passed to the 'settings' function.

OUTSTANDING ISSUES

The product has one outstanding issue. This issue being, when the zoom feature is used, the map is scaled up, which ignores the map's boundaries and fills the entire screen upon high magnification. The latitude and longitude are also not updated to match the new coordinate range. This issue was noted early on in development, and a fix was attempted however, due to time constraints, the fix could not be fully implemented in the final product.

RESULTS & CONCLUSIONS

TARGETS ACHIEVED?

Each number corresponds with each SMART target number (Page 3)

1. Every ozone 1-hour interval is displayed on a map of the entire European continent, showing all major geographical rivers, lakes, cities and coastlines. Therefore, this task was achieved. (Figure 1).
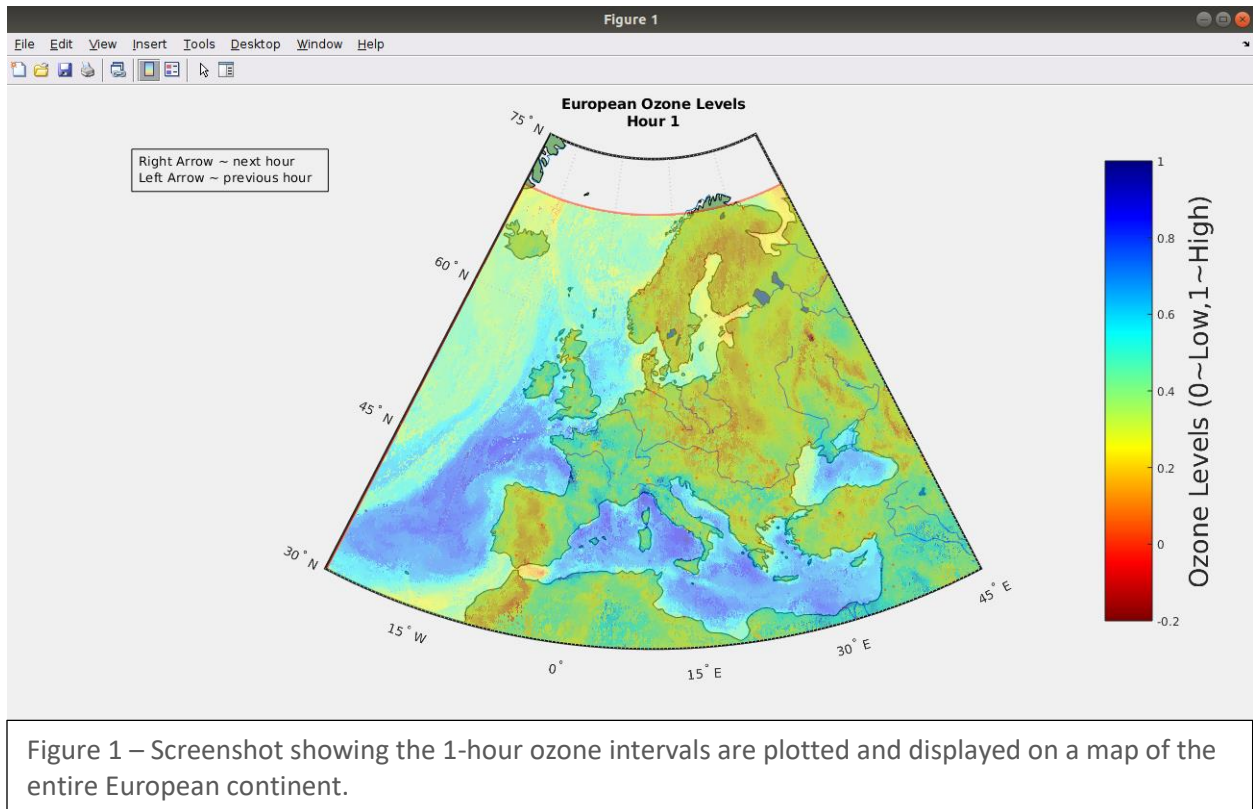

Figure 1 – Screenshot showing the 1-hour ozone intervals are plotted and displayed on a map of the entire European continent.

2. Each ozone value is represented by a colour from a fixed colour scale at each geolocation result. Therefore, this task was achieved. (Figure 2)
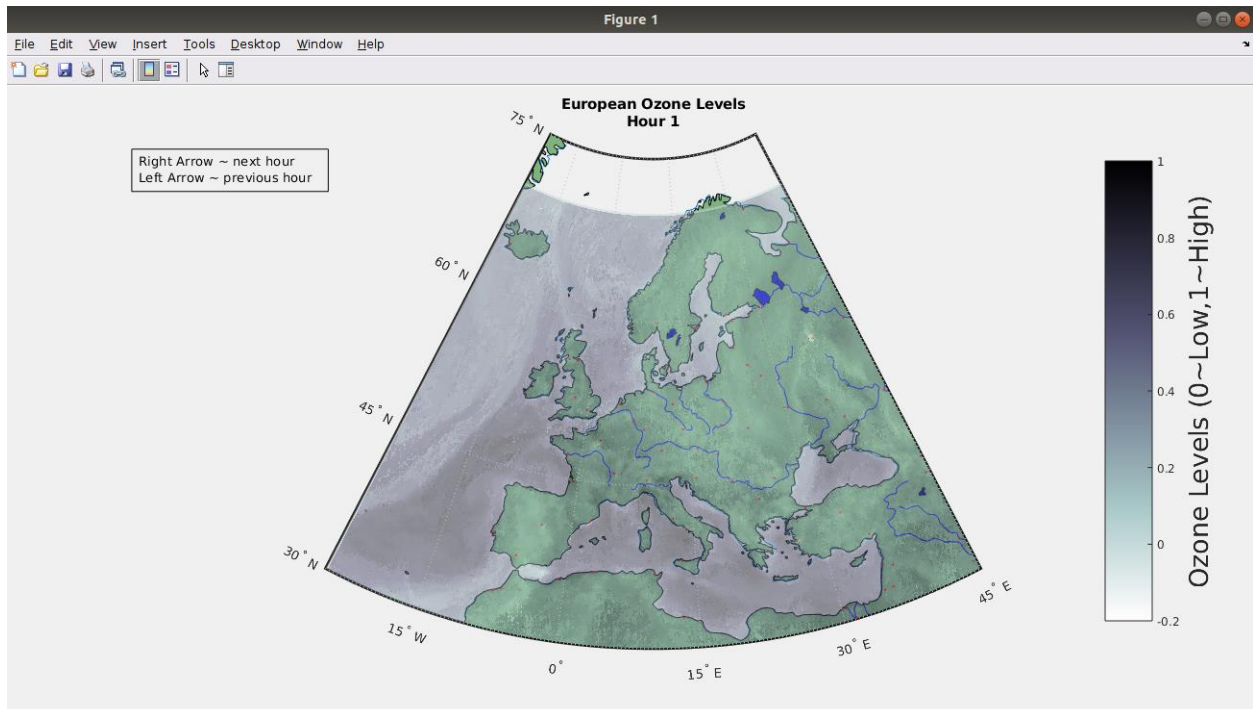

Figure 2 – Screenshot showing the colourblind mode option.

3. Each 1-hour ozone interval has been tested to load and display to the screen within 0.5 seconds, for the stepMode and movieMode. These tests have been done 3 times, with each test resulting in an execution time within 0.5 seconds. Therefore, this task was achieved. Cannot be shown via screenshots.

4. The figure zoom feature is enabled, allowing a magnification to 1 data result plot pixel. However, the issue regarding the map boundaries and updating the latitude and longitude was not fixed. Therefore, this target was achieved but not to a high standard. (Figure 3)
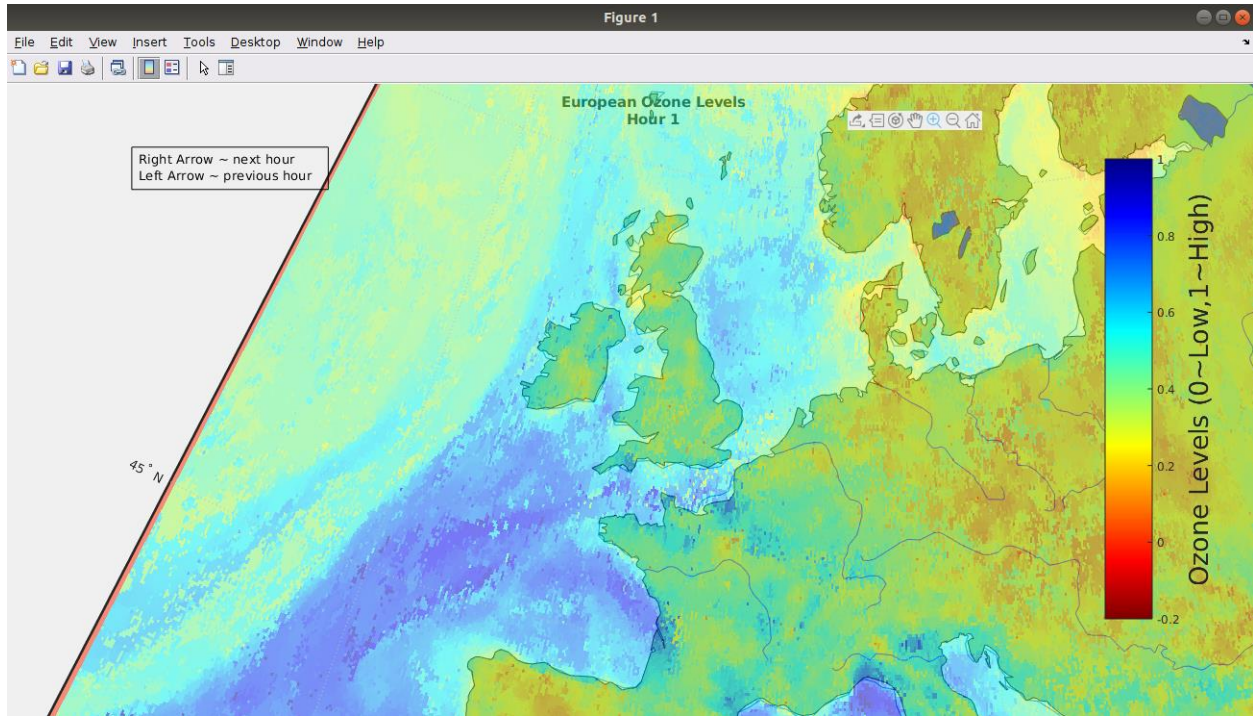

Figure 3 – Screenshot showing the zoom feature. However, as shown the feature expands the map outside of the defined boundaries, which should not happen. The colour bar, title and key should not be covered by the map.

5. Two separate pop-up windows are displayed upon start-up. The first being a choice of a colourblind option and the second being the option of 'stepMode' or 'movieMode'. Therefore, this task was achieved. (Figures 4 and 5)
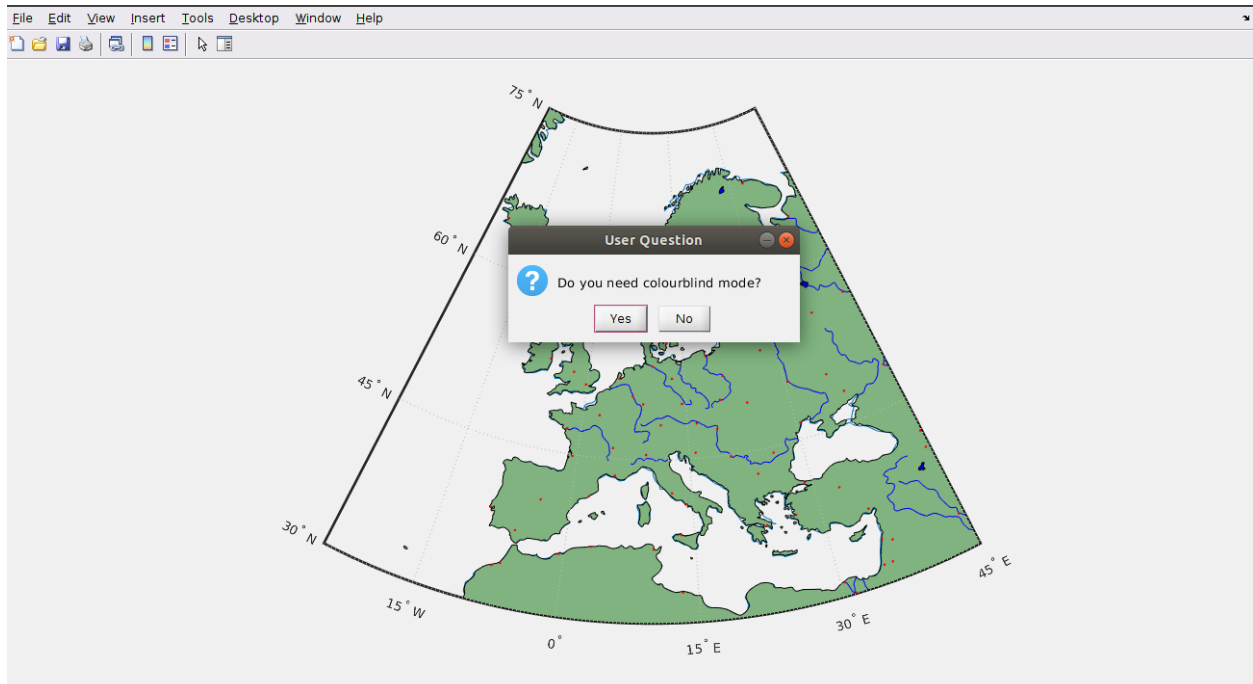

Figure 4 – Screenshot showing the colourblind pop-up question window.
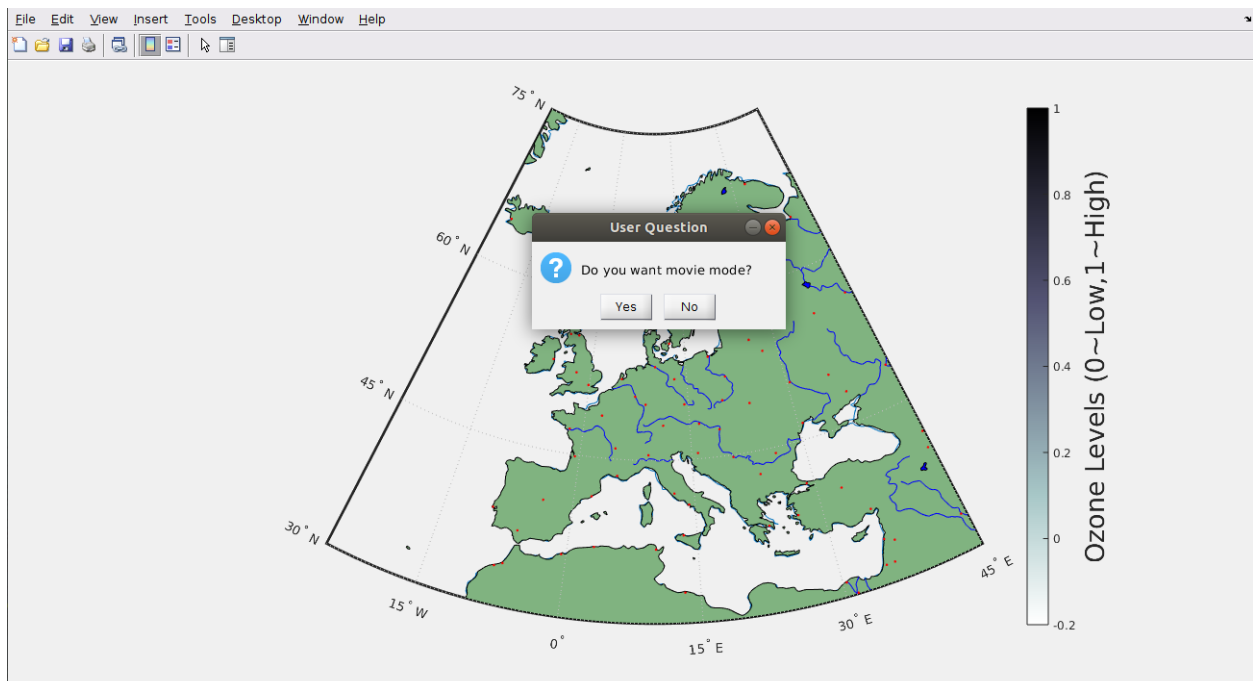

Figure 5 – Screenshot showing the movie mode pop-up question window.

6. Step mode loads and plots the first 1-hour interval onto the map, and then the left arrow key and right arrow key are used to step through each 1-hour ozone interval, loading and plotting the previous or next 1-hour interval onto the map, depending on which key is pressed. When the first 1-hour ozone interval is displayed and the left arrow is pressed, the last 1-hour ozone interval is loaded and displayed. When the last 1-hour ozone interval is displayed and the right arrow is pressed, the first 1-hour ozone interval is loaded and displayed. Therefore, this task was achieved. (Figure 6)
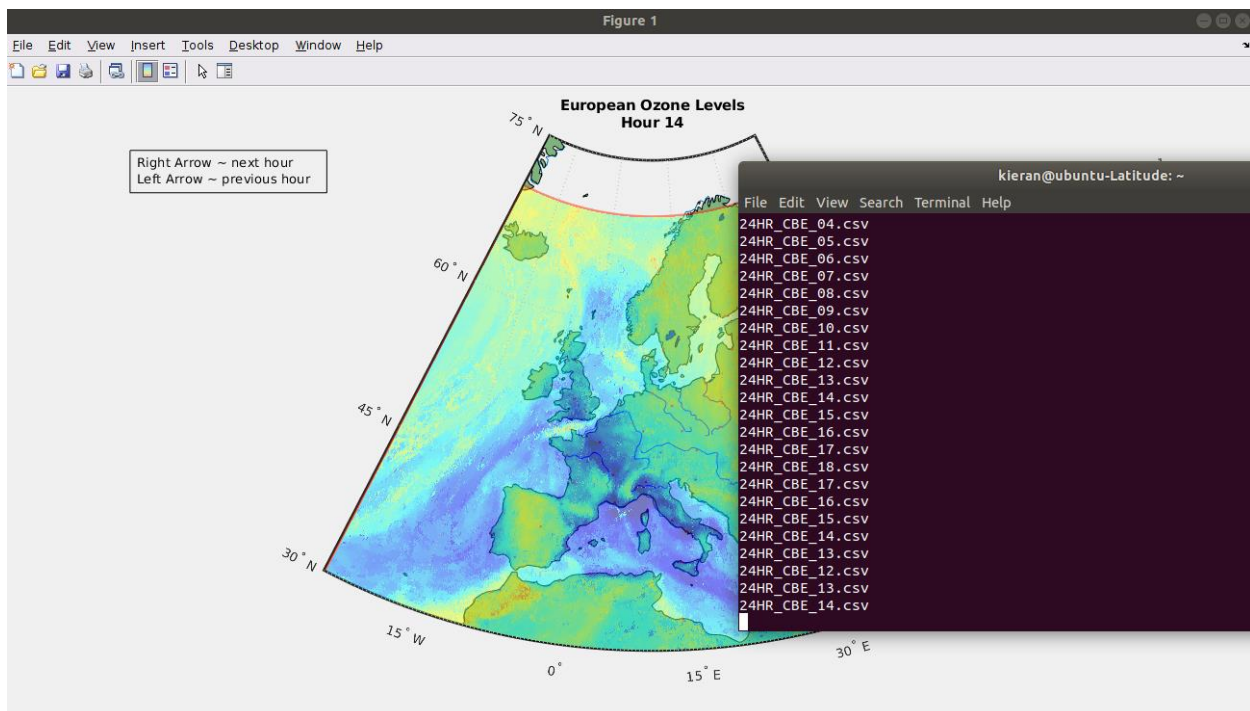

Figure 6 – Screenshot showing the step mode feature. The terminal of the left shows the 1-hour ozone intervals being stepped through.

7. Movie mode loads and plots the first 1-hour ozone interval onto the map, then after a 0.2 second pause, the ozone layer is re-plotted to equal the next 1-hour ozone interval, looping back to the first 1-hour ozone interval when the last interval is reached. Therefore, this task was achieved. (Figure 7)
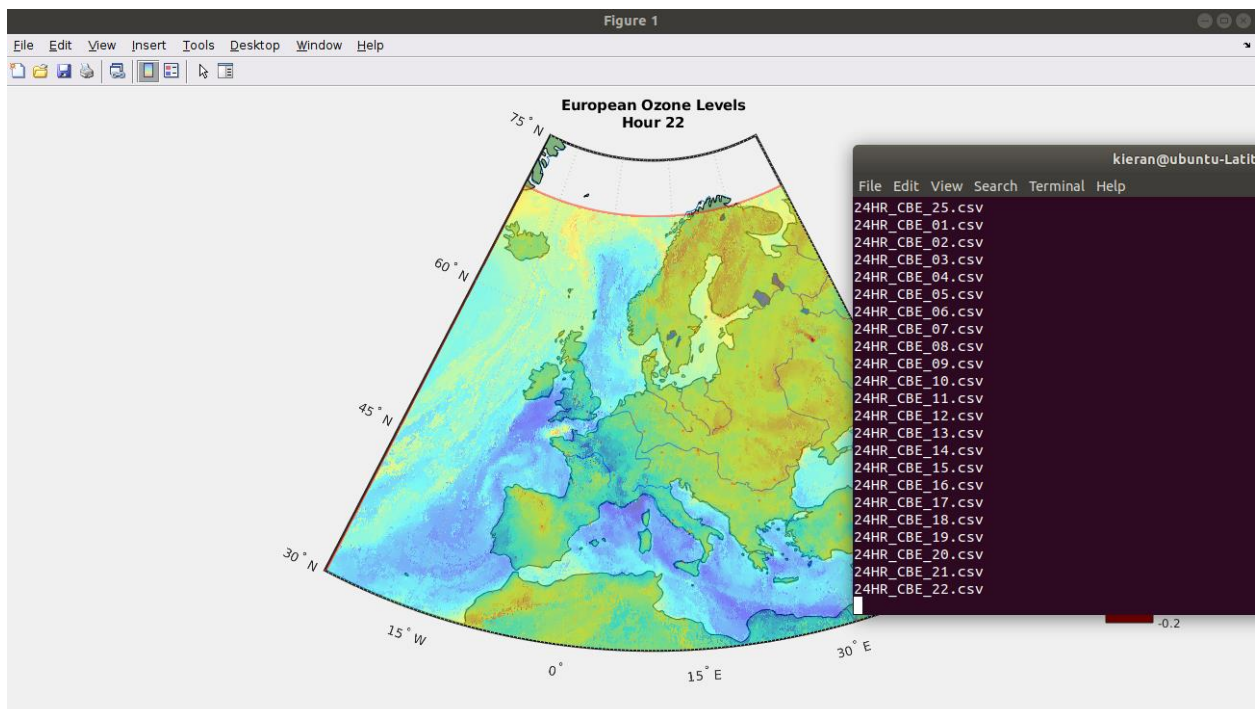

Figure 7 – Screenshot showing the movie mode feature. The left terminal showing the 1-hour intervals being displayed incrementally, in a loop.

CRITICAL ANALYSIS

Many of the SMART targets were achieved to a high standard, however some could be improved.

The step mode feature, (SMART target 6), is suitable for displaying the 1-hour ozone intervals for a single day. However, if the results to be displayed were for every day, for more than 2 weeks, this feature would become very complicated and cumbersome. This is due, the vast amount of data files that would be needed to be stepped through to reach a desired 1-hour interval. A solution to this would be, create a search bar that would require a date to be entered, once a valid date was entered, it would display the given day's first 1-hour ozone interval, and then allow the left and right arrows to step through the given day.

The movie mode feature, (SMART target 7), is suitable for displaying the 1-hour ozone intervals for a single day. However, it would not suitable to be display results for every day, for more than a few months. This is due to the execution time to display every day's results would be significantly large. To reduce the execution time, the same solution for the step mode could be implemented. This could be to have two input boxes; the first box would be the start date and the second input box would be for the end date. Once two valid dates are inputted, the program would display all 1-hour intervals between the two dates.

The overall product is suitable for a large screen such as a laptop or monitor however, this would be not work for smaller screen such as a mobile phone screen. This is due to the scaling of the map and text is

fixed. A solution would be to use the current figure's window resolution and use that to set the scale of the map and text. This would allow any reasonable screen size to display the map and text correctly.

The final critique of this product would be the zoom feature. As mentioned before, the zoom feature works however, it is to a poor standard, due to the issue regarding the map boundaries and updating the latitude and longitude. A solution would be, use the Geoaxes library and plot the ozone data as a geodensity plot. This would correct the issue regarding the latitude and longitude not updating and fix the issue with the map expanding out of the designated boundaries.

FUTURE WORK

If more time was allocated to this project, the four solutions mentioned in the Critical analysis would have been implemented into the final product, improving the product's quality significantly. However, adding more features to the program would be unnecessary, the client requested a program to display their results in a clear, simple and easy to interpret format, to help understand the variation between their models, and that is what the product offers.

If the client wished for more features to be later implemented, then these features would be welcomed and implemented accordingly. However, the program meets the specifications agreed upon and so no extra future work would be needed, other than the improvements mentioned above.

SUMMARY

The project has been coded in MatLab, utilizing the Mapping Toolbox library to display 24 hour, 1-hour intervals, spanning a given day. This was achieved by loading and re-plotting each 1-hour interval on-top of a European map, showing basic geographical rivers, lakes, cities and coastlines, with the CBE results being visualized by a fixed colour scale, representing the volume of ozone recording at each geolocation. An option for altering the display of the data to be suitable for colour blindness has been implemented, alongside the option to choose to step through each 1-hour interval or to have the 1-hour intervals automatically displayed as an animation.

Overall, the project meets the specification agreed upon by the client however, some areas mentioned could be improved upon, if more time and resources are allocated to this project in the future.

LIST OF APPENDICES

APPENDIX A
INTRODUCTION
The client, European Centre for Medium Range Weather Forecasts, has collected new 'Cluster Based Ensemble', which they wish to display in a clear, simple and easy to interpret format. They have specifically requested for a program that can visualise their ozone results, to help understand the variation between their models, with data being displayed onto a map of the European continent.

This document will explain in further detail, what the client requires, including an in-depth feature list, explaining what they desire. It will also clearly highlight the legal requirements and procedures that will be followed to develop this product, with any risks and assumptions clearly presented beforehand. A brief overview of the product will also be stated.

PROJECT REQUIREMENTS
The features requested by the ECMWF include, the ability to zoom on the map to gain a more accurate ozone visualisation of the local area, and two methods which allow for a choice between, a "step mode" or a "movie mode", the first being a method to increment and decrement through all the 1-hour intervals manually, and the latter method being an automatic loop, displaying each 1-hour interval incrementally. A colour-blind mode has also been specifically asked for, due to ECMWF's own target audience.

SMART Targets:
1. All ozone 1-hour intervals will be displayed on a map of the entire European continent, showing all major geographical rivers, lakes, cities and coastlines.
2. Each ozone value will be visualized by a fixed standard colour or a colour-blind scale, representing the volume of the ozone recording at each geolocation.
3. Each 1-hour ozone interval should be loaded and displayed onto the map within 0.5 seconds.
4. The figure plot should allow for zoom. This will be a zoom resolution of up to the maximum, 1 data plot pixel (individual ozone value at any given geolocation).
5. The program should display two questions upon start-up. The first being a choice of colourblind mode or standard mode, and the latter being a choice of "step mode" or "movie mode". These questions should only allow for 'Yes' or 'No' answers, and the entirety of the question should be displayed within a separate window upon start-up of the program.
6. A "step mode" should load and plot the first ozone 1-hour interval onto the map, and then allow for a manual keyboard press to re-plot the ozone layer, to either the next or previous 1-hour interval. When displaying the first 1-hour interval, it should loop back to the last 1-hour interval when the previous key is pressed. When displaying the last 1-hour interval, it should loop back to the first 1-hour interval when the next key is pressed.
7. A "movie mode" should load and plot the first ozone 1-hour interval onto the map, then after a 0.2 second pause, re-plot the ozone layer to the next 1-hour interval. This is to create an animation effect, with all the 1-hour intervals being displayed incrementally, looping back to the first 1-hour interval when the last interval is reached.

TERMS/ACRONYMS AND DEFINITIONS

Term/Acronym	Definition	Description
ECMWF	European Centre for Medium Range Weather Forecasts	The client's organisation
CBE	Clustered Based Ensemble	The mean of the ozone models as our ensemble value

RISKS AND ASSUMPTIONS

All code developed is not open source, and all permissions legally lie with the license holder, the sole developer of the program. The code cannot be altered or used without the license holder's permission.

Assumptions

Hardware:

- Minimum – Dell Latitude 3460 Intel Core i5

Software:

- Linux Ubuntu 18.04.4 LTS – Fully supported Operating System
- Microsoft Windows 10 – No Support but works upon release of product
- MatLab version R2019b – Fully Supported

CBE Files:

- All CBE data files will be provided by the client or a third-party upon request.
- CBE files no larger than 15mb
- The 25 unique CBE files will be for the 24 hour, 1-hour intervals, spanning a given day
- CBE file format either .fig or .csv only
- The CBE files must be named in-order, the first hour containing '1' and the last containing '25'

Risks

- Faulty or incorrect CBE files provided by the client or third-party may alter the visualisation result and is legally the client's duty to fix these issues, may lead to extra costs if it delays development.

- If the CBE files used change format during or after development, then it is the client's duty to ask permission from the license holder if they wish to alter the program. Extra costs may be involved in doing so. Any alteration of the code without the license holder's permission will lead to an illegal activity.

- If the program stops working due to MatLab or third-party intervention, then it is not the license holder's duty to develop a fix the program. Additional payment will be required if the license holder agrees to develop a fix.

- If the software contains bugs due to the license holder's mistakes or does not meet the specifications upon release, then the license holder will modify the software free of charge to fix these issues, so the product meets the standards agreed upon, if requested by the client.

- Any issues with standard libraries used within the product will legally not be the license holder's fault and will require additional payment if modifications are required.

- Using and loading more 25 1-hour intervals will not be supported and may result in issues with the program.

- Renaming the CBE files will result in program failure.

OUT OF SCOPE

- All clustered based ensemble data files given will be post-clustered and processed, therefore the program is not involved in processing the data and is only involved in visualizing CBE files.

- Not every city, river and lake will be displayed on the European map, the ones displayed will depend on the MatLab Mapping Toolbox decision to include them.

- Any features not mentioned within the specification requirements will not be implemented unless the client requests additional features during or after development at an additional cost.

SYSTEM/ SOLUTION OVERVIEW

The program will be developed within MatLab, it will require 25 1-hour interval CBE files that represent each ozone hour's result from a given single day. Once the program is executed, a figure window will load, alongside 2 pop-up question windows.

The first question will ask the user if they want to use the colour-blind mode, if 'Yes' is selected then a colour scheme of white to grey will be loaded, if 'No' is selected then the standard colour scheme will be loaded. This is to meet the requirement requested by the client for the code to include an option for altering the display of the data to be suitable for colour blindness. The second question will ask the use if they want 'step mode' or 'movie mode' to be loaded. 'Step mode' will allow for manual loading of each 1-hour interval via keyboard key presses whereas, 'movie mode' will load each 1-hour interval automatically in a loop, starting from hour 1 interval. This is to meet the client's need to have a simulation animation, as well as the ability to individually step through the results.

Then final product will visualise the newly collected 'Cluster Based Ensembles', matching the client's wishes to display their results in a clear, simple and easy to interpret format. Offering the client, the ability to visualise their ozone results, to help understand the variation between their models, with data being displayed onto a map of the European continent.

CONTEXT DIAGRAM/ INTERFACE DIAGRAM/ DATA FLOW DIAGRAM, APPLICATION SCREEN FLOW, SITEMAP, PROCESS FLOW

Graphical representation of product

Figure 8 - The general look of an ozone result plotted on the European map.
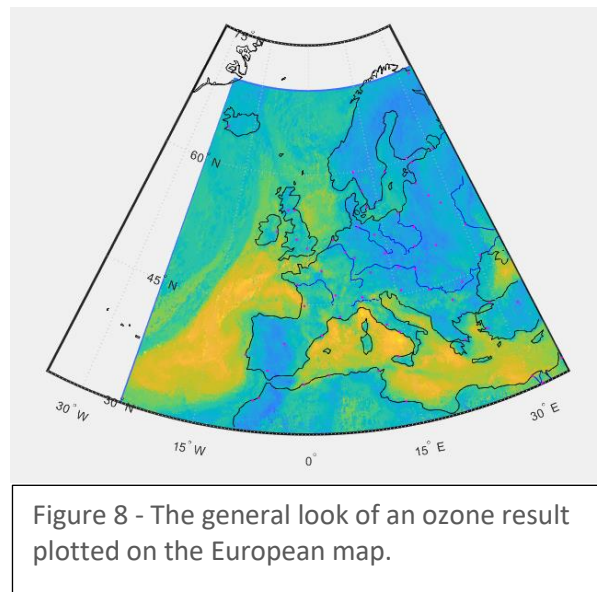
Overall Products Process Flowchart

This flowchart is located at (Appendix F).

PROJECT MANAGEMENT

The project will be managed using the Agile methodology however, as this is a sole project it will be structured slightly differently. Firstly, before any development is started, each SMART target will be converted to a user story and placed within a product backlog, with the top task being the highest priority and the bottom being the lowest priority. Each sprint will be 7 days long, with each sprint taking the top task off the product backlog and breaking it down into a sprint backlog, writing sprint backlog task down in a Product Log.

At the start of each Sprint, a Sprint Planning meeting will be held to discuss how the task from the Product Backlog is to be broken down into smaller tasks and placed with a Sprint Backlog, in priority order. Then each planned day of development, a meeting/scrum will be held before any work is done, this will be discussing what work was carried out during the previous workday, and any current standing issues regarding the project will be noted down in the meeting brief. This will then be used to re-order sprint backlog task priorities, if needed, and to plan what work is needed to be done for the current day.
After each day of development, any progress and issues will be documented in the Project Log, to be used in the next scrum.

GitHub will be used to maintain version control during this project. After each sprint feature is worked on, it will be pushed to it's given GitHub branch, and then when completed it will be merged with the master branch. This will result in the ability to rollback code if errors occur.

Due to other commitments such as, other module coursework and exams, not every day will be used to develop the project, instead 2 to 3 days a week are to be designated to this project, which may vary depending on the given week's priorities.
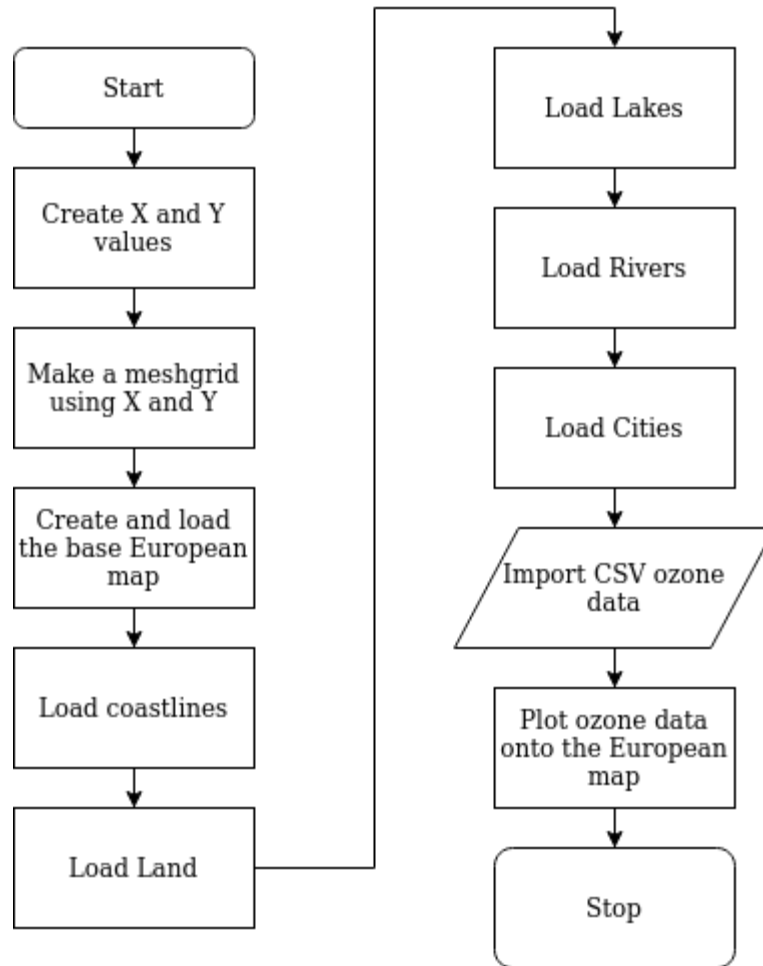
Loading basemap and Ozone data

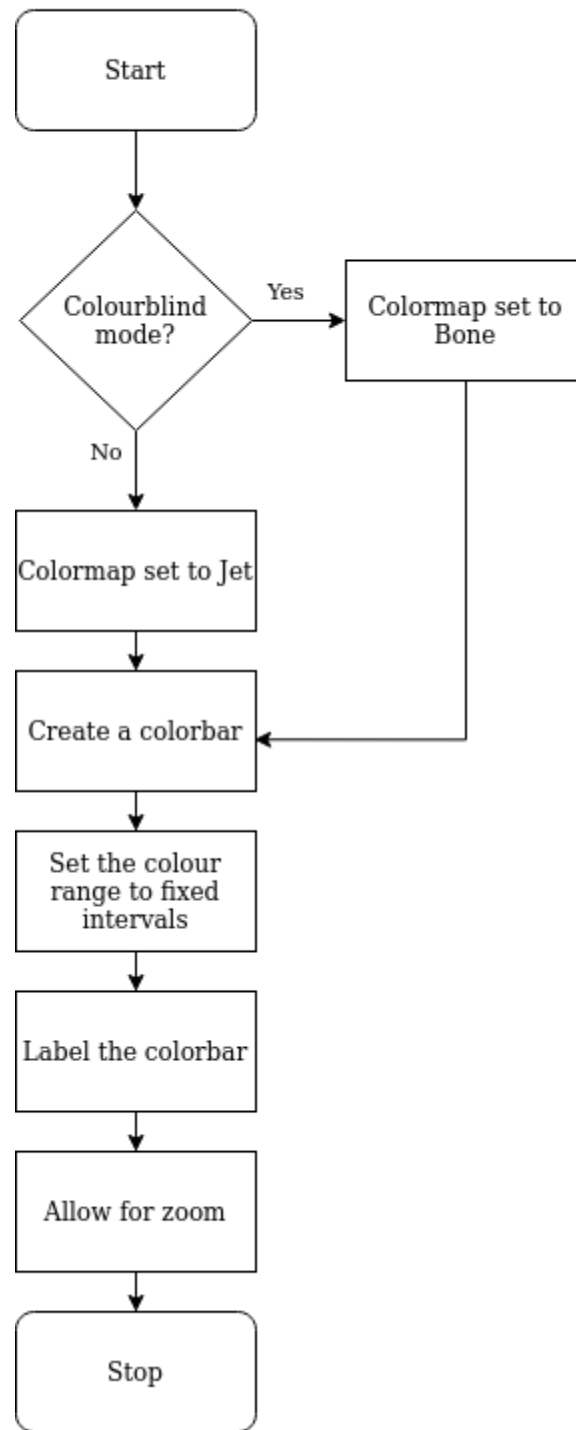Figure 9 - The flowchart for the original overall script.

Settings

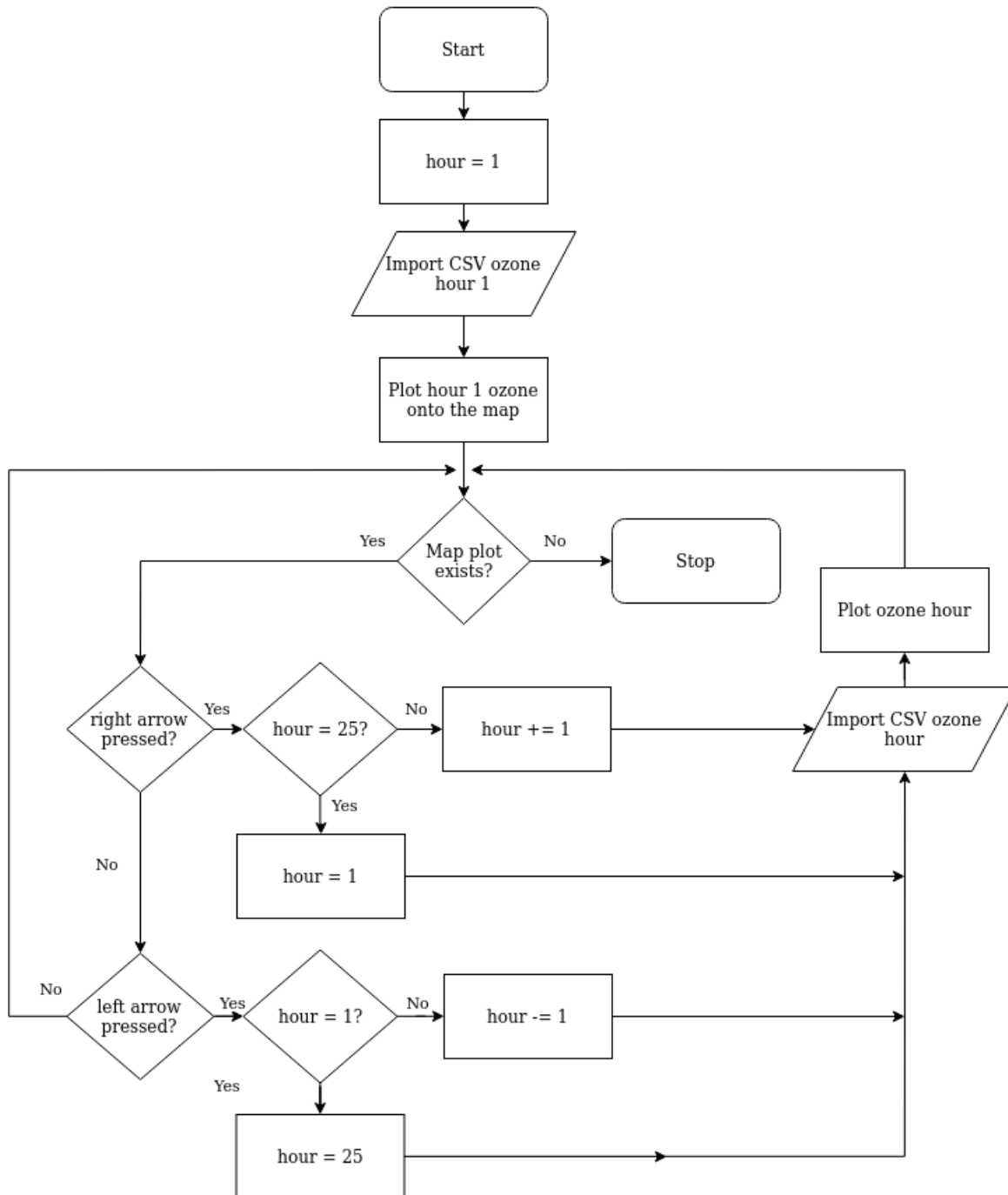Figure 10 - The flowchart for the settings function.

Figure 11 - The flowchart for the stepMode function.

Loading Each ozone hour in Movie mode

Start

hour = 1

Import CSV ozone hour 1 as Z

Plot hour 1 ozone onto the map

Map plot exists? — No → Stop

Yes

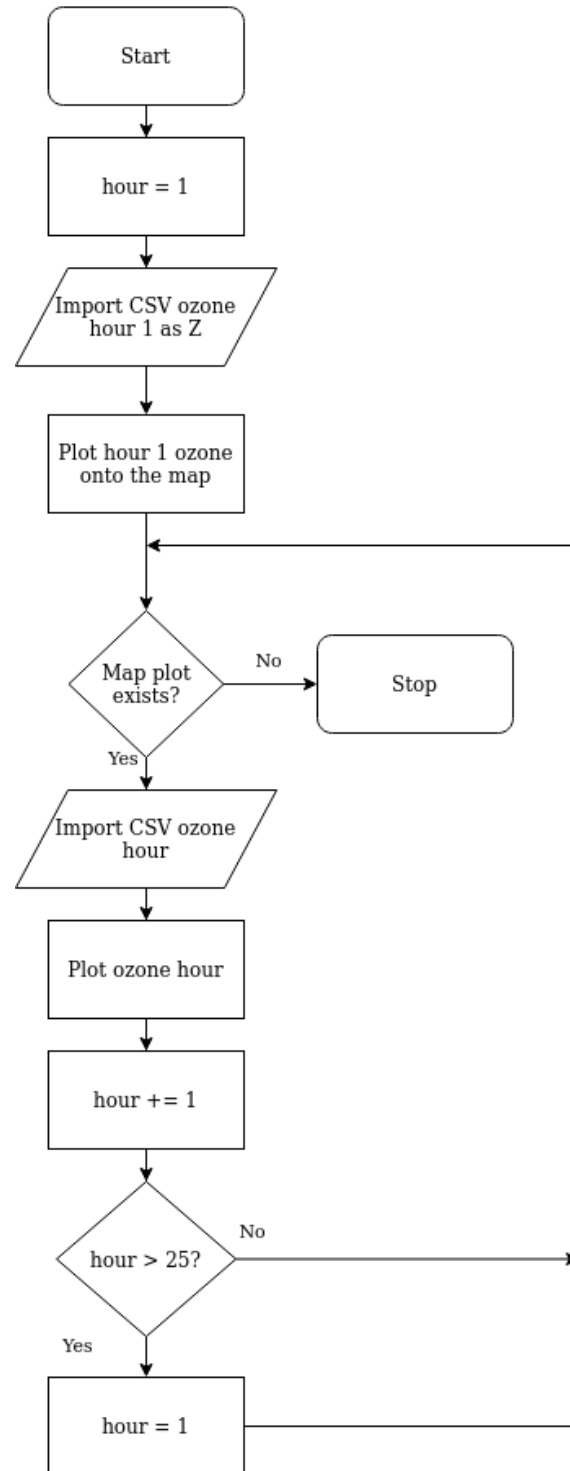Import CSV ozone hour

Plot ozone hour

hour += 1

hour > 25? — No →

Yes

hour = 1

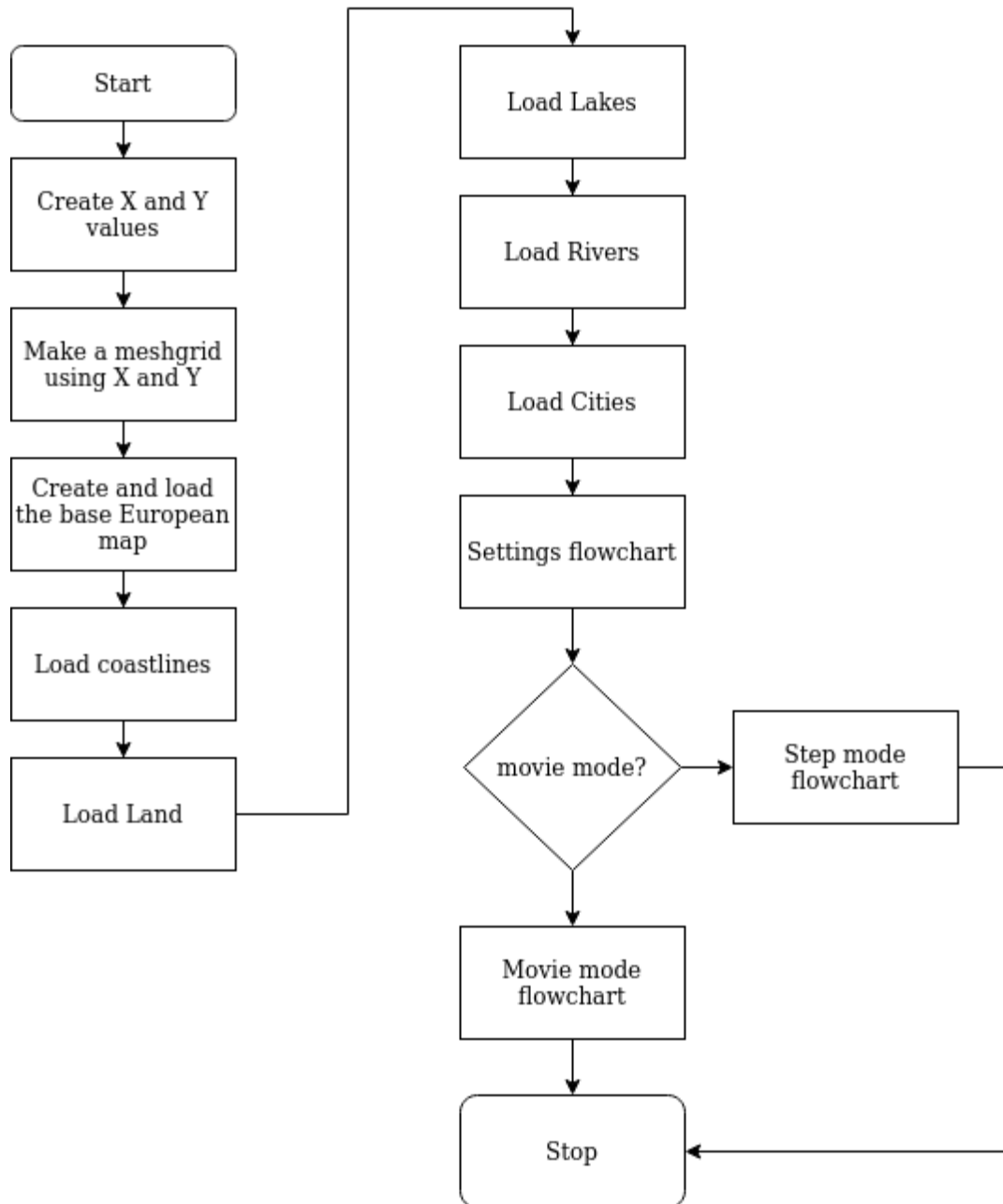Figure 12 - The flowchart for the movieMode function.

Figure 13 - The flowchart for the overall project, being the main function.

APPENDIX G

MAIN FUNCTION

(Appendix F)

- Generates Latitude and Longitude values and converts them into a meshgrid.
- Generates and loads the European base map from the Mapping Toolbox, onto a figure window.
- Loads the European coastlines, lands coordinates, lakes, rivers and cities onto the European map.
- Generates a sperate pop up window, asking the user if they desire colourblind mode. If 'Yes' then the colour scheme 'bone' will be passed, as well as the map figure, to the settings function. If 'No', then the 'jet' colour scheme will be passed instead.
- Generates another sperate pop up window, asking the user if they desire 'movie mode'. If yes` then the function movieMode will be executed, if 'No' then the function stepMode will be executed instead.

SETTINGS FUNCTION

(Appendix B)

- Sets the colour map to the value passed from the main function, 'bone' or 'jet'
- Generates a colourbar, sets its position on the figure and sets its colour range to a fixed range
- Then it enables the zoom feature, allowing the map the be zoom in and out of

STEPMODE FUNCTION

(Appendix C)

- Creates a path to the CBE files
- Imports the first 1-hour interval as Z
- Then plots the ozone layer onto the European map using X, Y, Z (Lat, Long, Ozone)
- Then a while loop is executed which stops once the map figure has been deleted
- Inside the while loop it checks for a 'leftarrow' or 'rightarrow' key press.
- If 'leftarrow' key is pressed then, the previous 1-hour ozone interval is loaded
- If 'rightarrow' key is pressed then, next 1-hour ozone interval is loaded
- This loops until the figure is closed, or deleted

MOVIEMODE FUNCTION

(Appendix D)

- Same first 4 bullet points as stepMode function
- Inside the while loop it loads the next 1-hour ozone interval
- Waits 0.2 seconds
- Then loops again, loading the next 1-hour interval

APPENDIX H

The code archives the SMART targets set due to;

- The option for a colourblind mode is meet with the pop-up colourblind question window and the settings function.
- The option to choose which mode to execute is meet with the pop-up movie mode question window in the main function.
- All 25 1-hour intervals are loaded and displayed onto a European map, either executed by the 'stepMode' or 'movieMode' function
- The European map is generated and populated with rivers, cities etc. by the main function
- The movieMode function executes the loop of loading each of the 1-hour ozone intervals incrementally, creating an animation effect
- The stepMode function allows for the end user to step through each 1-hour ozone interval, by using manual keyboard presses to re-plot the ozone layer, to either the next or previous 1-hour interval
- The settings function allows for a zoom feature

APPENDIX I

Choices continued….

KEYBOARD KEYS

Keyboard key presses where chosen in the stepMode function, instead of a button or slider due to its simplicity to develop as well as, its simplicity to operate. Using a button or slider was considered however, this would have required more lines of code, and would have been more complicated and time consuming for the end user to use. Thus, keyboard keys were chosen to be used as means to step through each 1-hour ozone interval.

POP-UP WINDOW QUESTIONS

For the project, two questions needed to be displayed to the end user, the first being the choice of colour scheme and the second being which mode to use, step or movie. Therefore, it was decided to use separate pop-up windows to ask these questions. The reasoning being, pop-up windows grab the user's attention, making it clear that they need to answer the questions before proceeding any further. Also, only 2 button options are available, making invalid input impossible, as the user can only chose between predefined answers, thus preventing errors from occurring.