

|  |
| --- |
| Coventry University  5011CEM Big Data Programming Project Specification Document  THE VISUALIZATION OF ATMOSPHERIC OZONE INTENSITY |

**Author: Evaldas Vainavicius**

**Document Version 1.0**

**06/02/2020**

**Table of Contents**

[1. Introduction 3](#_Toc38209703)

[2. Project Requirements 3](#_Toc38209704)

[2.1 Related documents 3](#_Toc38209705)

[2.2 Terms/Acronyms and Definitions 4](#_Toc38209706)

[3. Risks and Assumptions 5](#_Toc38209707)

[4. Out of Scope 5](#_Toc38209708)

[5. System/ Solution Overview 6](#_Toc38209709)

[5.1 Context Diagram/ Interface Diagram/ Data Flow Diagram, Application Screen Flow, Sitemap, Process Flow 6](#_Toc38209710)

[6. Project Management 6](#_Toc38209711)

[7. References 6](#_Toc38209712)

[8. Open Issues 7](#_Toc38209713)

[Appendices 7](#_Toc38209714)

# Introduction

This document aims to improve the European climate models by providing a visualisation of 7 ozone models, which data was given from The European Centre for Medium Range Weather Forecasts (*ECMWF*).

There are 14 major climate models used worldwide however, in this document there are going to be used only 7 of them which can produce more than 250 chemical species and pollutants. In addition, the grid that will be used is going to be 0.1° latitude and longitude. The output of the models are provided hourly.

The project would benefit many institutions which require to visualize the data in a GUI form to help analyse problems dependent on the Ozone level in atmosphere at a specific location. This could help institutions which are interested in climate changes to predict the processes and draw a plan of actions or potentially make conclusions out of the results got.

# Project Requirements

* The final program will be able to read only two specific file types, which are *NetCDF* (Provided by the *ECMWF*) and Excel file type.
* The final program will be able to visualise the data which is given in the file and produce a map of Europe in two ways: for colour blind people and for those who can see all colours.
* The final program will provide the user with 24 hour *CBE* (Cluster Based Ensemble) data, from which user can use a slider or next/previous buttons to go through it.
* The final program will also let the user to change the colour of a map in 5 different variations. This includes full colour or grey or red and so on, which are made for colour blind people.

## Related documents

|  |  |  |
| --- | --- | --- |
| **Component** | **Description of The Document** | **Article Title** |
| Article 1 | This article concludes why data visualization is a highly valuable point to have in order to communicate with executives, functional leads and data scientists. It also shows how to adapt the data presented and visualize it in an appropriate manner so different people from different jobs would understand what is presented.  <https://www.researchgate.net/publication/299391071_Data_Visualization_Enhancing_Big_Data_More_Adaptable_and_Valuable> | Data Visualization Enhancing Big Data More Adaptable And Valuable |
| Article 2 | This article presents the challenges of Big Data visualization. It includes new methods, applications, and technology progress of Big Data visualization.  [https://ieeexplore.ieee.org/abstract/document/7918044/authors#authors](https://ieeexplore.ieee.org/abstract/document/7918044/authors%23authors) | Big data visualization Tools and challenges |
| Article 3 | This article gives a brief summary of some of the multi-dimensional visualization techniques that are used in data mining. It also gives the right solutions of how data should be presented in businesses and why some of them fail to get the correct results they want.  <https://scholar.googleusercontent.com/scholar?q=cache:XNvvy14egmMJ:scholar.google.com/+Ajibade+and+Adediran+-+2016+-+An+Overview+of+Big+Data+Visualization+Techniques+&hl=lt&as_sdt=0,5> | An Overview of Big Data Visualization Techniques |

## Terms/Acronyms and Definitions

|  |  |  |
| --- | --- | --- |
| **Term/Acronym** | **Definition** | **Description** |
| ECMWF | The European Centre for Medium Range Weather Forecasts. | The data provider for the project. |
| CBE | Cluster Based Ensemble data. | Data which has been parallel processed and from which were created matrices for 24 hour forecast of Ozone. |
| NetCDF | Network Common Data Form | NetCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access and sharing of array-oriented scientific data. |
| GUI | Graphical User Interface | Graphical user interface is a form of user interface that allows users to interact with electronic devices using graphical icons and audio indicator such as primary notation, instead of text-based user interfaces. |

# Risks and Assumptions

**Assumptions:**

* The program is capable of reading the *NetCDF* format files, which contain all the data with longitude and latitude included. Using different file types rather than *NetCDF* will not work! In addition, modifying the original *NetCDF* file, by providing different file of the same type may cause problems with the program.
* The program itself has been created using a 1920x1080 resolution, therefore other resolution options are not included. In addition, the program is capable of running on a laptop computer and not on mobile device.
* The loading time for each map plot, while using the program should be less than 2 minutes. The computer specifications that were used during the project can be found in the **Appendix** section **Appendix A**, also with the minimum computer specifications for running Matlab software.

**Risks:**

* The data file can be changed but at your own risk. The program is created to read data from NetCDF format file where it is capable of storing data in a matrices of size: 400x700. Therefore, using another data file can cause problems for the program and will require overwriting of the program!
* The program have been written with the versions of standard libraries available at that time, without any bugs in them. Any changes that have been made to the program by the user, considering the use of different libraries may cause bugs in the standard libraries. Change program at your own risk.

# Out of Scope

* The cloud compilation is not going to be used in this project.
* The final program will be able to read only files which are readable by Matlab software. However, this may require to do some changes in the program itself.
* The final product will be written only in a Matlab programming language and will be understandable in the English language. The final product will not include other languages such as German, Russian or any other.
* The program is using the CBE data which has been provided. Therefore, no algorithm for the CBE is provided.
* The final program is only going to visualise the 24 hour data using a graphical user interface (GUI).

# System/ Solution Overview

The purpose of this software is to provide a user with a 25 hour GUI map, which is user friendly and easily interactive. When running the software provides one map per one hour, which concludes in 25 maps where each map is different hour, which is shown at the top of the map. The maps are different and show different data respectively. The user can switch between the maps using the allocated buttons or a slider provided at the bottom of the map. In addition, maps can be presented in different colours which are specifically designed for people with colour blindness.

# Context Diagram/ Interface Diagram/ Data Flow Diagram, Application Screen Flow, Sitemap, Process Flow

No graphical representations are provided. The reason for this is because everything can be seen when the program is run.

# Project Management

The project management is going to take place every week on Mondays, Wednesdays, Sundays.

Everything is going to be recorded in the Excel file provided, which will include the following:

* The overall aims that the project supposed to achieve.
* What changes have been made, if no changes have been made then nothing will be written.
* What problems have occurred and how it was dealt with.
* When did the changes have been applied (The date of the changes).

For a full project management please see **Appendix B** in the **Appendix** section of the specification document. In addition, to see the how the version control of the project was working here is a [GitHub link](https://github.coventry.ac.uk/vainavie/Visualization-of-Atmospheric-Ozone-Intensity-Project.git).

# References

Ajibade, S.S. and Adediran, A. (2016) *An Overview of Big Data Visualization Techniques in Data Mining* [online] available from <<https://scholar.googleusercontent.com/scholar?q=cache:XNvvy14egmMJ:scholar.google.com/+Ajibade+and+Adediran+-+2016+-+An+Overview+of+Big+Data+Visualization+Techniques+&hl=lt&as_sdt=0,5>> [12 04 2020]

Ali, S.M., Gupta, N., Nayak, G.K., and Lenka, R.K. (2016) *Big Data Visualization: Tools and Challenges* [online]available from <http://ieeexplore.ieee.org/document/7918044/> [13 04 2020]

Fiaz, A.S.S., Asha, N., Sumathi, D., and Navaz, A.S.S. (2016) *Data Visualization: Enhancing Big Data More Adaptable and Valuable* [online] available from <<https://www.researchgate.net/publication/299391071_Data_Visualization_Enhancing_Big_Data_More_Adaptable_and_Valuable>> [14 04 2020]

Wikipedia (2020) *Graphical User Interface* [online] available from <<https://en.wikipedia.org/wiki/Graphical_user_interface>> [27 03 2020]

Wikipedia (2020) *NetCDF* [online] available from < <https://en.wikipedia.org/wiki/NetCDF> > [27 03 2020]

# Open Issues

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Issue ID** | **Issue** | **Raised By** | **Raised On** | **Solution/ Decision** | **Resolved By** | **Resolved On** | **Status** |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

# Appendices

**Appendix A**

Computer specifications that were used to create the program:

* OS (Operating System): Windows 10
* CPU: Intel Core i5-7300HQ
* RAM: 8GB
* SSD Hard Drive
* GPU: Nvidia Geforce GTX 1050

Full specifications for running Matlab software can be found in the link here: Minimum computer specification

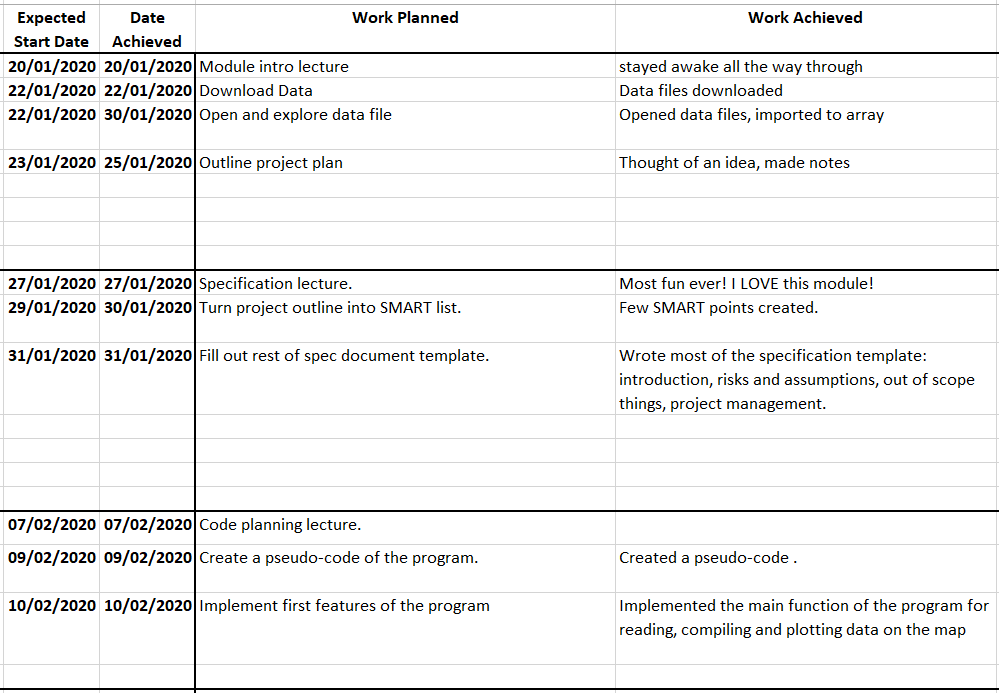
**Appendix B**

Full project management can be seen below.

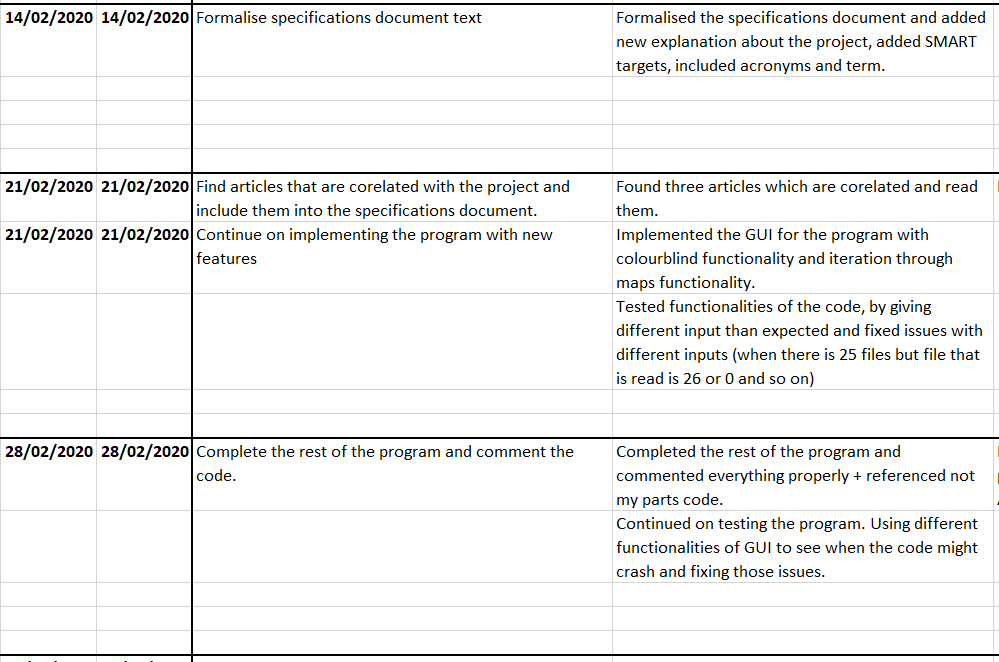
|  |  |
| --- | --- |
| **Week** | **Progress** |
| **1** | * Outline project plan. * Fetch and extract the data given for further use in program |
| **2** | * Create SMART target list * Modify the specification document |
| **3** | * Write a pseudo-code for the program * Start implementing the first features of the program |
| **4** | * Formalise specification document to match the program |
| **5** | * Find articles correlated to the project and include them into specification document * Continue on implementing the program with new features |
| **6** | * Complete implementing the features of the program * Comment the program in an appropriate manner |
| **7** | * Fix issues, bugs that may occur while running the program * Continue on commenting the program |
| **8** | * Start writing the report for the project * Modify the specification document |
| **9** | * Write the rest of the report * Post the program on GitHub |
| **10** | * Complete writing the report * Complete writing the specification document * Create a viva video presentation |
| **11** | * Complete a viva video presentation * Finish the project |

The pictures below proves how the project management document was updating every week.

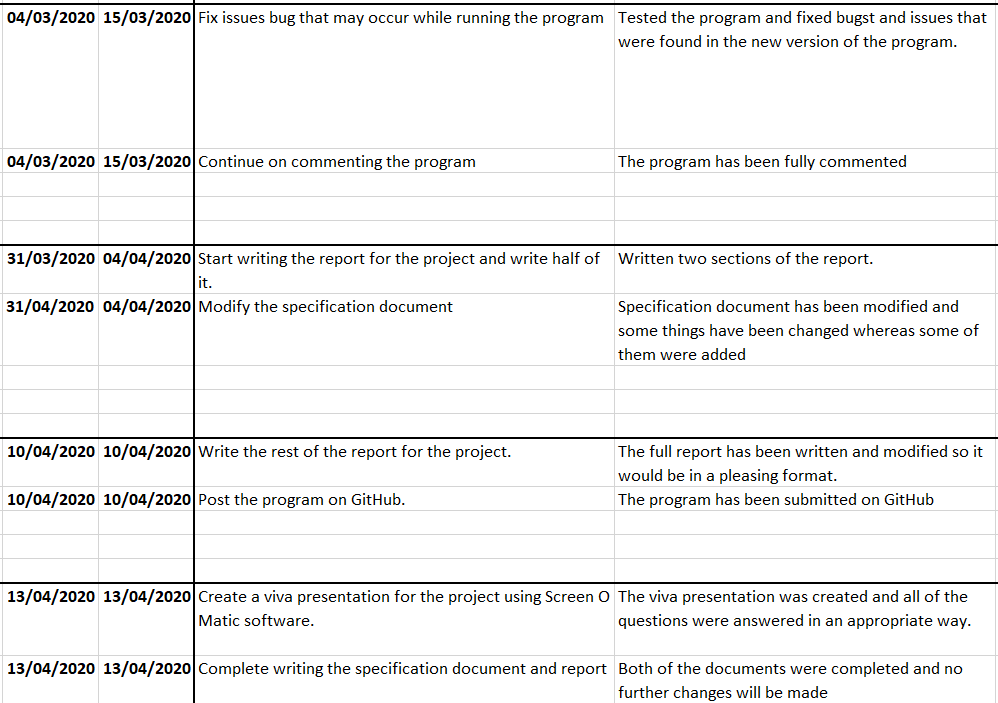
**Table 1. Showing the first, second, third and fourth week of how the project was managed.**



**Table 2. Showing from weeks 5-8 how the project was managed**



**Table 3. Showing 8-11 weeks, of how project was managed.**



# List of pictures

[Table 1. Showing the first, second, third and fourth week of how the project was managed. 9](file:///C:\Users\171207\Downloads\Desktop\Big%20data%20Project\Project%20files\Specification%20Document%20Template%20ver02.docx#_Toc38210654)

[Table 2. Showing from weeks 5-8 how the project was managed. 10](file:///C:\Users\171207\Downloads\Desktop\Big%20data%20Project\Project%20files\Specification%20Document%20Template%20ver02.docx#_Toc38210655)

[Table 3. Showing 8-11 weeks, of how project was managed. 10](file:///C:\Users\171207\Downloads\Desktop\Big%20data%20Project\Project%20files\Specification%20Document%20Template%20ver02.docx#_Toc38210656)