Spatial and Temporal Patterns of Temperature During the 2009 Heatwave in South-East Australia

DUE JUNE 9 2023

James Currie, s3901655

Spatial Information Science Analytics, Geom1057

RMIT UNIVERSITY

Executive Summary

In 2009, a significant heatwave occurred in South-East Australia. To understand the extent and severity of this heatwave, and the way that the temperature changed during this heatwave, a time-series point maximum and minimum temperature dataset was scraped from the Bureau of Meteorology. This dataset was then interpolated using Ordinary Kriging to create a time-series raster dataset, containing the maximum and minimum temperatures of each day in the study period. The time-series raster dataset was used to create derived datasets, demonstrating the length of exposure to extreme heat during the heatwave, to be used as a metric of the severity of the heatwave. These products are a good starting point for an analysis into the broader effects of the heatwaves.

Aim

Research Question: What was the spatial and temporal extent of the heatwave, and what were the most severely affected areas?

This research question is broken down into 2 objectives, and 3 datasets. The objectives are to answer the following questions:

- **What was the daily maximum and minimum temperature in South-Eastern Australia during the 2009 Heatwave?**

  To answer this question, a minimum and maximum temperature raster dataset were produced for each day in January and February, 2009.

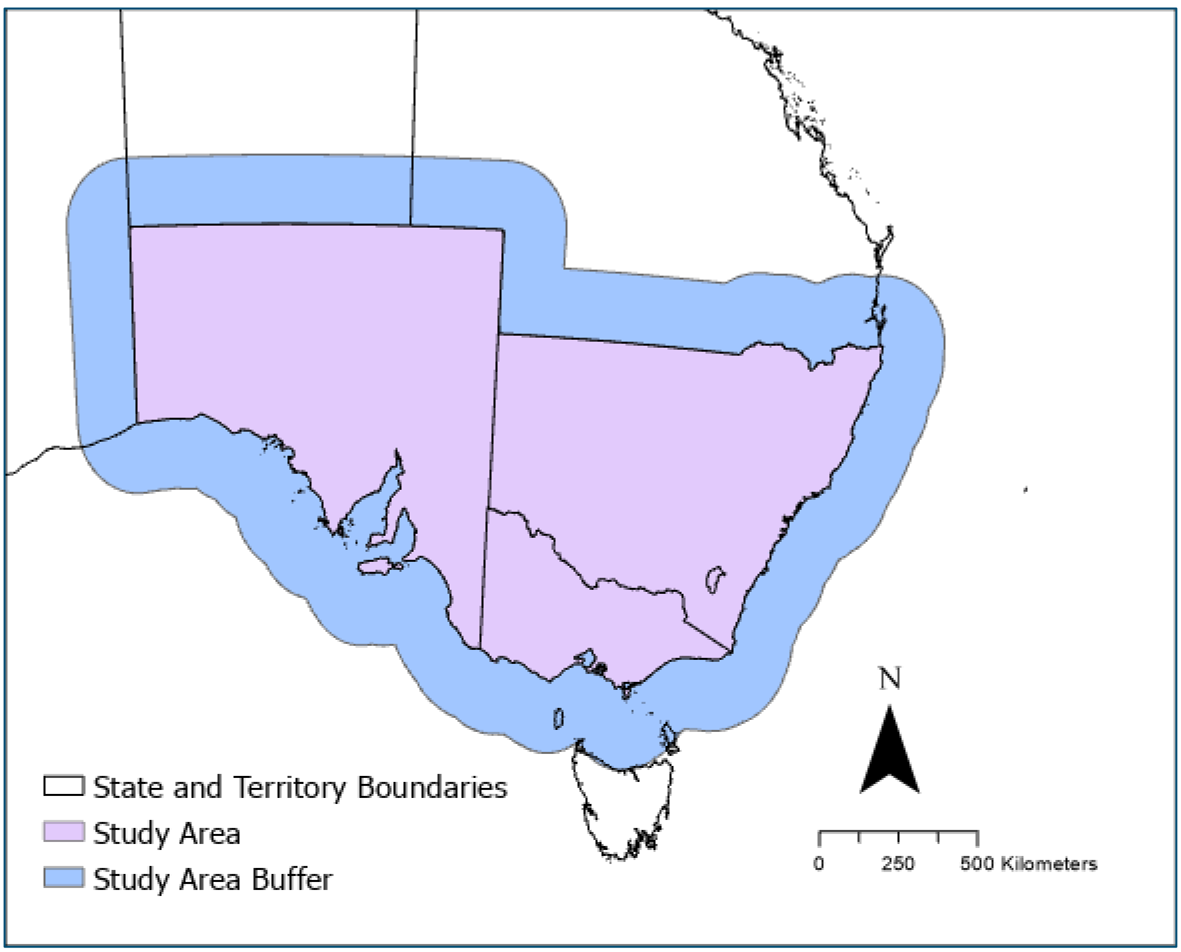
- **How long were areas above a certain temperature threshold defined as a heatwave?**

  To answer this question, two datasets were derived from the daily minimum and maximum temperature datasets. These datasets show the days in a row each area had a maximum temperature over 40 degrees Celsius, and the days in a row each area had a minimum temperature over 25 degrees Celsius.

Introduction

Exposure to hot and excessive ambient temperatures can cause pain, suffering and death(Crimmins et al., 2016). In Australia, heatwaves have been responsible for more human deaths than any other natural hazard, including bushfires, storms, tropical cyclones and floods (Argüeso D et al., 2009). As such, it is critical to understand the severity of heatwaves, and this report acts as a case study of how to understand the intensity and severity of heatwaves. Fatigue is a significant factor in the severity of heatwaves, and is highly influenced by the length of time that the heatwave occurs, hence the length of the heatwave is measured.

Study Area



The study area includes the mainland area of the Australian Capital Territory, South Australia, New South Wales, and Victoria. Weather stations within 2 degrees of the study area were considered in analysis.

The study area of Southern Australia, New South Wales, the Australian Capital Territory and Victoria was chosen as they are spatially contiguous and highlighted as being specifically effected by the heatwave (National Climate Centre, 2009).

Methodology

Input Datasets

The input datasets are listed below.

Dataset	Source	Use
Daily Maximum Temperature Records	**Bureau of Meteorology (Received by python scraping script)**	**Source maximum temperature point data**
Daily Minimum Temperature Records	**Bureau of Meteorology (Received by python scraping script)**	**Source minimum temperature point data**
Administrative State Boundaries	**Australian Bureau of Statistics**	**For delineating the study area**

Process

1. Temperature Data Aggregation

A python script was built to import BOM maximum and minimum temperature data from individual stations. This data was aggregated into one dataset and refined data to the study time period. Station data labelled as 'unreliable' or with null values were removed.

```
In [13]:
#Scraping the data

#Define headers (So that the BOM website treats this script as a browser)
headers={'User-Agent': 'Mozilla/5.0'}

#Download the data for each station
for stat in StatList:

    #Go to the corresponding access code page and record the access code. (I'm not sure what this code is or means but you n
    link1 = 'http://www.bom.gov.au/jsp/ncc/cdio/wData/wdata?p_stn_num=' + str(stat) +'&p_display_type=holdingGraph&p_nccObsC
    response = requests.get(link1, headers = headers)
    AccessCode = response.text.rpartition(':')[2]

    #Use the station number and the access code to download the stations data as a zip file
    link = 'http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dailyZippedDataFile&p_stn_num=' + str(stat) +'&
    response = requests.get(link, headers = headers)
    print('Station number', stat, 'Downloaded as a .zip from the link: ' + link)
    with open(zipst+'/'+ str(stat) +'.zip', 'wb') as f:
        f.write(response.content)
print('Done')
```

```
Station number 36143 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=36143&p_c=-261275049&p_nccObsCode=122&p_startYear=2023
Station number 63281 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=63281&p_c=-800908752&p_nccObsCode=122&p_startYear=2023
Station number 35134 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=35134&p_c=-246891351&p_nccObsCode=122&p_startYear=2023
Station number 35290 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=35290&p_c=-249088579&p_nccObsCode=122&p_startYear=2023
Station number 87159 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=87159&p_c=-1519350016&p_nccObsCode=122&p_startYear=2023
Station number 63010 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=63010&p_c=-794063779&p_nccObsCode=122&p_startYear=2023
Station number 83004 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=83004&p_c=-1377944563&p_nccObsCode=122&p_startYear=2023
Station number 44010 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=44010&p_c=-387387779&p_nccObsCode=122&p_startYear=2023
Station number 70005 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
lyZippedDataFile&p_stn_num=70005&p_c=-980151764&p_nccObsCode=122&p_startYear=2023
Station number 70328 Downloaded as a .zip from the link: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_display_type=dai
```

Python script scraping data from every station in Australia, which was then refined to the study area.

2. Interpolation Model Identification

Although many interpolation methods have been used on temperature data, Thin Plate Spline, Kriging, and Natural Neighbour were selected to be trialled due to preferable results in previous studies (Hofstra, N et Al., 2009). Through manual testing of these interpolation methods in ArcGIS Pro using a random sample of dates, Ordinary Kriging was selected as it produced the lowest Root Mean Square error (RMSe).

3. Interpolation Model Application

The interpolation model template was applied to the point dataset for each date using a python script, which ran within ArcGIS Pro using Arcpy and the Geostatistical Analyst toolkit. This template model is an xml document, where some parameters must be set statically, and therefore apply to all dates in the dataset the same, and others are optimised by the software to minimise RMSe on each individual date. Examples of parameters which had to be standardised for the whole dataset are trend removal and the search neighbourhood. These were optimised by average RMSe for the whole dataset.


```xml
<?xml version="1.0"?>
<model name="Kriging" optimize = "ByCrossvalidation">
  <dataset Label="Dataset" ParamName="dataset1" dataset-type="DVA" />
  <enum name="KrigingMethodType">Ordinary</enum>
  <enum name="KrigingResultType">Prediction</enum>
  <items name="Datasets">
    <item name="Dataset">
      <enum name="TrendType">None</enum>
      <model name="NeighbourSearch" options="CopyFromVariogram">
        <enum name="Type">Standard</enum>
        <value name="NeighboursMax" auto="false">50</value>
        <value name="NeighboursMin" auto="false">2</value>
        <enum name="SectorType">Four45</enum>
        <value name="MajorSemiaxis" auto="false">6</value>
        <value name="MinorSemiaxis" auto="false">6</value>
        <value name="Angle">0</value>
      </model>
    </item>
  </items>
  <model name="Variogram">
    <value name="DatalayerCount">1</value>
    <value name="NumberOfLags" auto="false">12</value>
    <value name="LagSize" auto="false">0.8</value>
    <enum name="PairsType" auto="false">Semivariogram</enum>
    <bool name="NuggetOn">true</bool>
    <value name="Nugget" auto="false">0.54</value>
    <value name="MeasurementError100">100</value>
    <bool name="VariogramModelAuto">false</bool>
    <model name="VariogramModel">
      <enum name="ModelType">Stable</enum>
      <value name="Parameter" auto="false">1</value>
      <value name="Range" auto="false">6</value>
      <bool name="Anisotropy">false</bool>
      <value name="Sill" auto="false">32.70353888355749</value>
    </model>
  </model>
</model>
```


Interpolation model template xml document. Variables under the 'Variogram' section were optimised individually for each date, whereas variables above were optimised broadly for the whole dataset.

4. Derived Dataset Creation

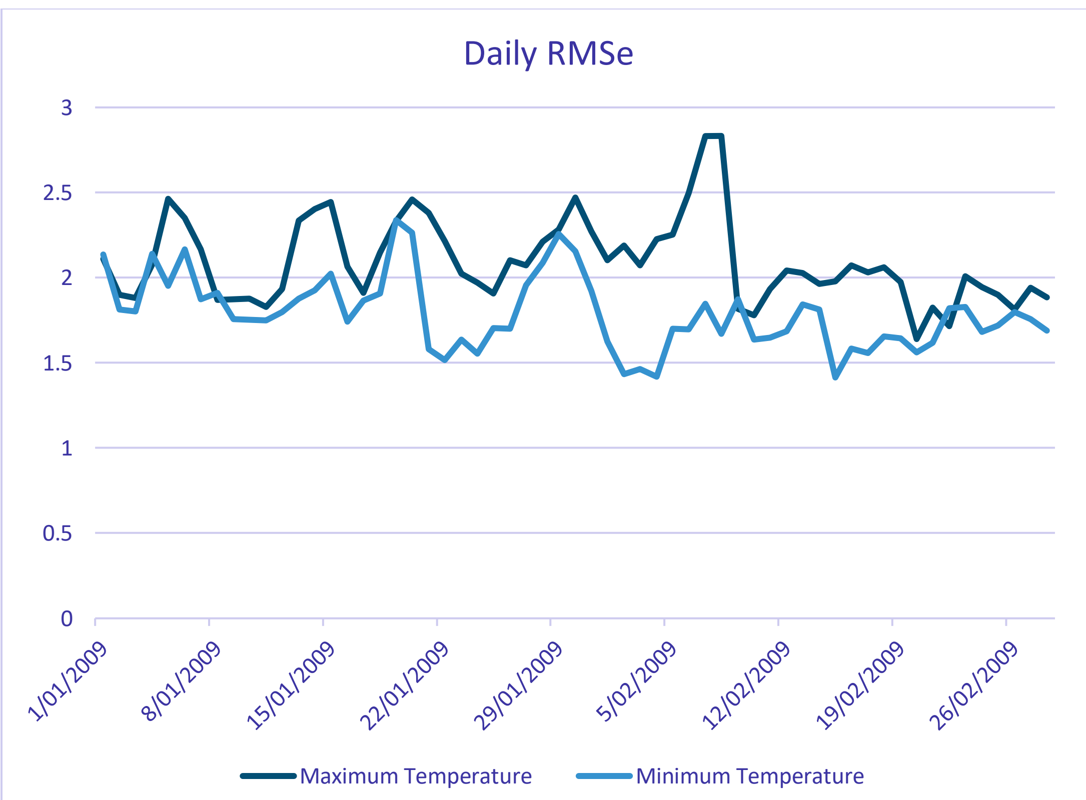
Derived datatsets were created using a python script, iterating a raster calculation over each day in a row, where the temperature was over the identified threshold. If the temperature dropped below this threshold, the count was restarted. The highest count of days in a row was recorded and saved to the derived datasets.

Results

The datasets produced and included in this submission, to be viewed alongside this report are listed below:

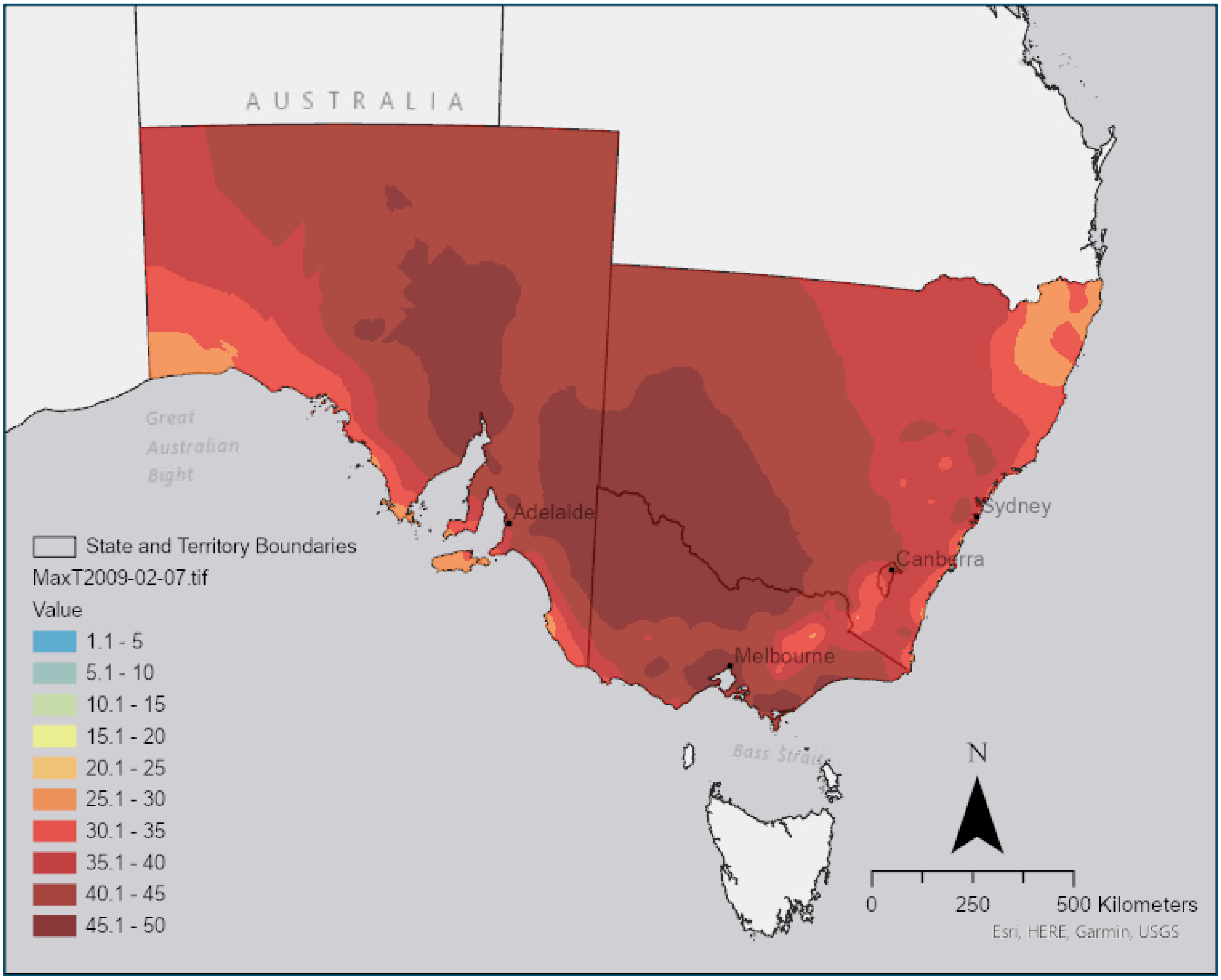
- The ZIP folder 'Rasters' contains the daily maximum and minimum temperature raster datasets.
- Animations have been produced displaying the daily temperature variation over time. This have been included in the submission in the 'Animations' folder. The individual maps that make up these animations are also included under the 'PNG maps' folder.
- Maps of the derived datasets are included below in this document, and they are also included in the folder 'Derived'.
- Note that no file geodatabase is supplied as per the assessment as no file geodatabase was used for storing the inputs or outputs of this project.
- The scripts used and other information in the above methodology can be found under the 'Programs' folder

RMSe Results

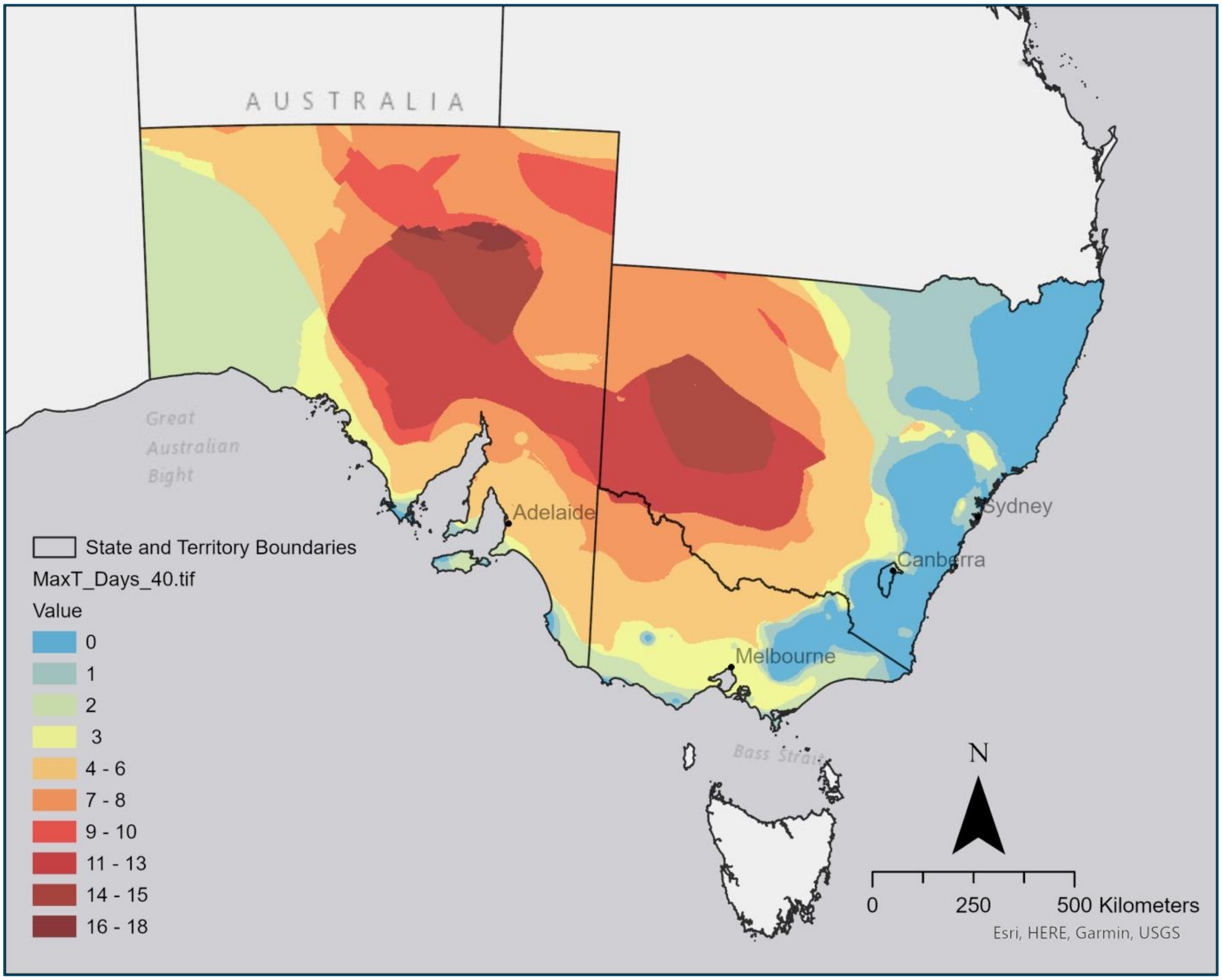


The Root Mean Square error is plotted above, and was generally acceptable, although there were some noticeable peaks, particularly towards the centre of the study period.

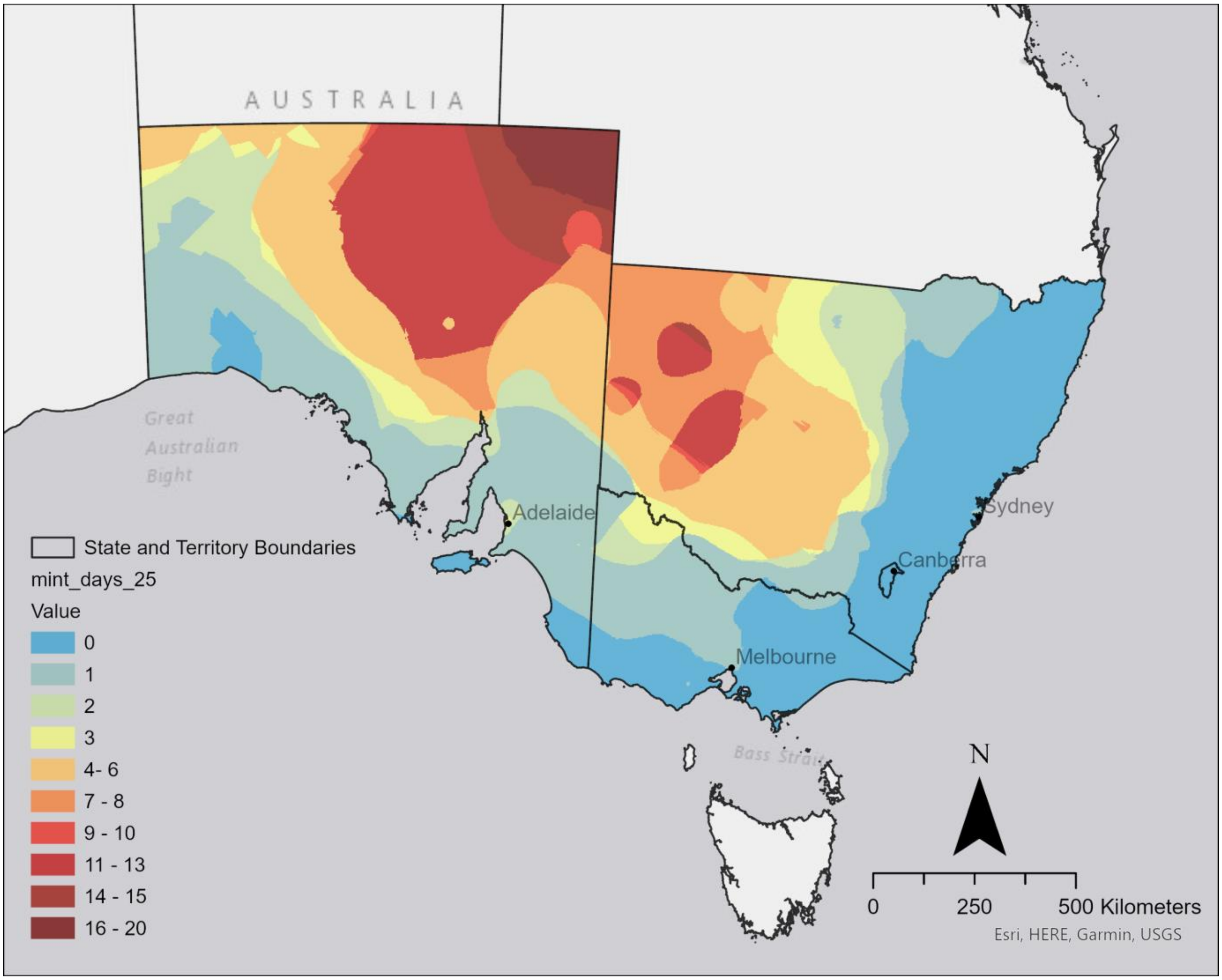
Maps



Interpolated map of maximum temperature on 07/02/2009. Similar maps for each day are included in this submission. Huge areas of South Eastern Australia experienced temperatures over 45 degrees Celsius on 07/02/2009.



Derived Dataset 1 – Number of days in a row where the maximum temperature was over 40 degrees Celsius. Some areas were over 40 degrees Celsius for 18 days in a row, whereas some areas near the ocean never reached 40 degrees Celsius.



Derived Dataset 2 – Number of days in a row where the minimum temperature was over 25 degrees Celsius. A huge portion of eastern NSW and Victoria never reached a minimum temperature over 25 degrees Celcius, however some areas had a minimum temperature over the threshold for 20 days in a row.

Discussion and Recommendation

The output daily raster datasets clearly demonstrate the movement of the extreme temperature over space and time with visually evident patterns. Two major peaks of the heatwave intensity in maximum temperature can be seen on 28/01/2009 and 07/02/2009, with extreme temperatures inbetween. Similar patterns emerge with the minimum temperature dataset, however the first peak appears to be more severe (with a minimum temperature over 30 degrees Celsius in large parts of South Australia). This output aligns with the existing understanding of the heatwave (National Climate Centre, 2009).

The derived datasets demonstrated that significant areas of South Eastern Australia never reached above the temperature thresholds, particularly along the eastern coast of Victoria and NSW. Areas that were most severely effected were towards central Australia.

A notable limitation of this project is that the measured variable to define the severity of the heatwave is the maximum and minimum temperature, however a preferable metric would be the deviation of the maximum and minimum temperature from the normal maximum and minimum temperature at each place, which is how heatwaves are actually defined (John Nairn Et Al., 2013). This, as-well as including other environmental factors such as humidity and wind speed in the analysis, would make for a more complete measure of severity of the heatwave.

A Key limitation of this dataset and interpolation model is the inability to optimise the trend removal for each date. Kriging utilises spatial autocorrelation to produce a result, which is independent of the underlying trend in the data. As such, the underlying trend should be removed from the dataset before the spatially autocorrelated component is calculated (Liu G -J, 2023). The current model applies no trend removal because there is that is what results in the lowest overall RMSe, however this is not necessarily true for each individual date. An improved model could modify the level of trend removal on each date based on the temperature characteristics of each individual date, and what level of trend removal results in the lowest RMSe for that individual date.

These datasets provide a good starting point for further analysis into the effects of the heatwave, such as identifying correlation between the number of days in a row the temperature was over 40 degrees Celsius and excess deaths.

References

Argüeso D, D. L. A., Evans JP, Parry M, Gross M, Alexander L, Green D, Perkins S (205). Heatwaves affecting NSW and the ACT: recent trends, future projections and associated impacts on human health. NARCliM Technical Note 5. Sydney, Australia.

CENTRE, N. C. 2009. The exceptional January-February heatwave in south-eastern Australia. Bureau of Meteorology, Special Climate Statement 17.

CRIMMINS, A., BALBUS, J., GAMBLE, J. L., BEARD, C. B., BELL, J. E., DODGEN, D., EISEN, R. J., FANN, N., HAWKINS, M. D., HERRING, S. C., JANTARASAMI, L., MILLS, D. M., SAHA, S., SAROFIM, M. C., TRTANJ, J. & ZISKA, L. 2016. Executive Summary. The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment. Washington, DC: U.S. Global Change Research Program.

Hofstra, N., Haylock, M., New, M., Jones, P., and Frei, C. (2008), Comparison of six methods for the interpolation of daily, European climate data, J. Geophys. Res., 113, D21110, doi:10.1029/2008JD010100.

JOHN NAIRN, R. F. 2013. Defining heatwaves: heatwave defined as a hea timpact event servicing all community and business sectors in Australia. CAWCR Technical Report No. 060

LIU, G.-J. 2023. Lecture 5: Spatial Interpolation 4. RMIT University.

Appendix: Project Journal

This project journal outlines my progress from finding the issue I wanted to study in the project, through to finalizing the parameters and outputs of the model.

- **20/03/2023**

First look into project idea, I'm interested in weather data interpolation, however I've been struggling to find the appropriate data. Gang-Jun has suggested looking at weather in Tasmania as no other students are currently focusing there.

- **22/03/2023**

If I'm using interpolation as my method, I should use a higherarchical sorting criteria if possible (as per lecture 5 notes)
I also briefly looked into salinity interpolation of soil based on ground samples stored by DEECA.

- **25/03/2023**

Through research I found that if I'm going to study temperature I should utilize cokriging, as it takes into account elevation. Not sure if this is relevant to this subject though.

- **27/03/2023**

I came across the idea of heatwave mapping and created the script to scrape temperature data from the BOM. Gang-jun approved this project and also requested this script for use in future classes.

- **22/04/2023**

Local krigging selected and complete for the maximum and minimum temperature datasets, however upon visual interpretation of the outputs for the deviation from normal datasets, the quality is not appropriate and therefore an alternate method for the deviation datasets is identified. I'll interpolate one 'normal' minimum and one 'normal' maximum dataset and use the raster calculator to calculate the deviation from normal for each day.

- **24/04/2023**

I noticed 'glitchy' distortions in the output and discussed these with Gang-Jun. He said to go ahead with these as they has the best RMSE and aren't a problem.

- **07/05/2023**

Finalised project parameters as without trend removal - First order: 2.09 RMS AVG, Second order: 2.1 AVG, None, 2.08 AVG.

- **08/05/2023**

I realised all of the analysis so far had been using WGS84, and therefore was distorted. Reprojected inputs to Geoscience Australia Lambert. Oddly, the average RMSE is higher for the outputs.

- **09/05/2023**

Deviation from normal results are still unsatisfactory, and due to the project timeline getting close to the end I have decided to remove these from the scope of the project with Gang-Juns approval.

- **08/06/2023**

Project Report and Journal Finalised and Submitted.