COMP1204: Data Management
Coursework One: Hurricane Monitoring

Xiaoke Li
31951473

March 13, 2022

1 Introduction

This work aims to extract useful storm data from a set of hurricane reports in KML format. The structure of tags in these reports resemble that of XML files. Instead of using an XML parser, the script uses regular expressions and the Unix utility sed extensively to filter and group data of individual storms.

The extracted data is exported in CSV format. Storm plots are generated as PNG files from the CSV output using the provided create_map_plot.sh script.

2 Create CSV Script

```
#!/bin/bash

# 1 Set 2 varialbes to extract data from KML,and save as CSV

kml_input=$1
csv_output=$2

#   Set the first line of output file
echo "Timestamp,Latitude,Longitude,MinSeaLevelPressure,MaxIntensity" > $2

# 2 Collect the 5 required data
#         Cut off useless parts then add unit at the end of each line

#   grep timestamp
grep "<dtg>" $1 | cut -d ">" -f 2 | cut -d "<" -f 1 > ts.txt

#   grep latitude
grep "<lat>" $1 | cut -d ">" -f 2 | cut -d "<" -f 1 \
    |  sed "s/$/& N/g" >lat.txt

#   grep longitude
grep "<lon>" $1 | cut -d ">" -f 2 | cut -d "<" -f 1 \
    | sed "s/$/& W/g" > lon.txt

#   grep max intensity
grep "<intensity>" $1 | cut -d ">" -f 2 | cut -d "<" -f 1 \
    | sed "s/$/& knots/g" > inten.txt

#   grep min sealevel pressure
```

```
grep "<minSea.*>" $1 | cut -d ">" -f 2 | cut -d "<" -f 1 \
    | sed "s/$/& mb/g" > press.txt

# 3 Paste data in order, output final data as a CSV file
paste -d ',' ts.txt lat.txt lon.txt press.txt inten.txt >>$2
```

3 Storm Plots

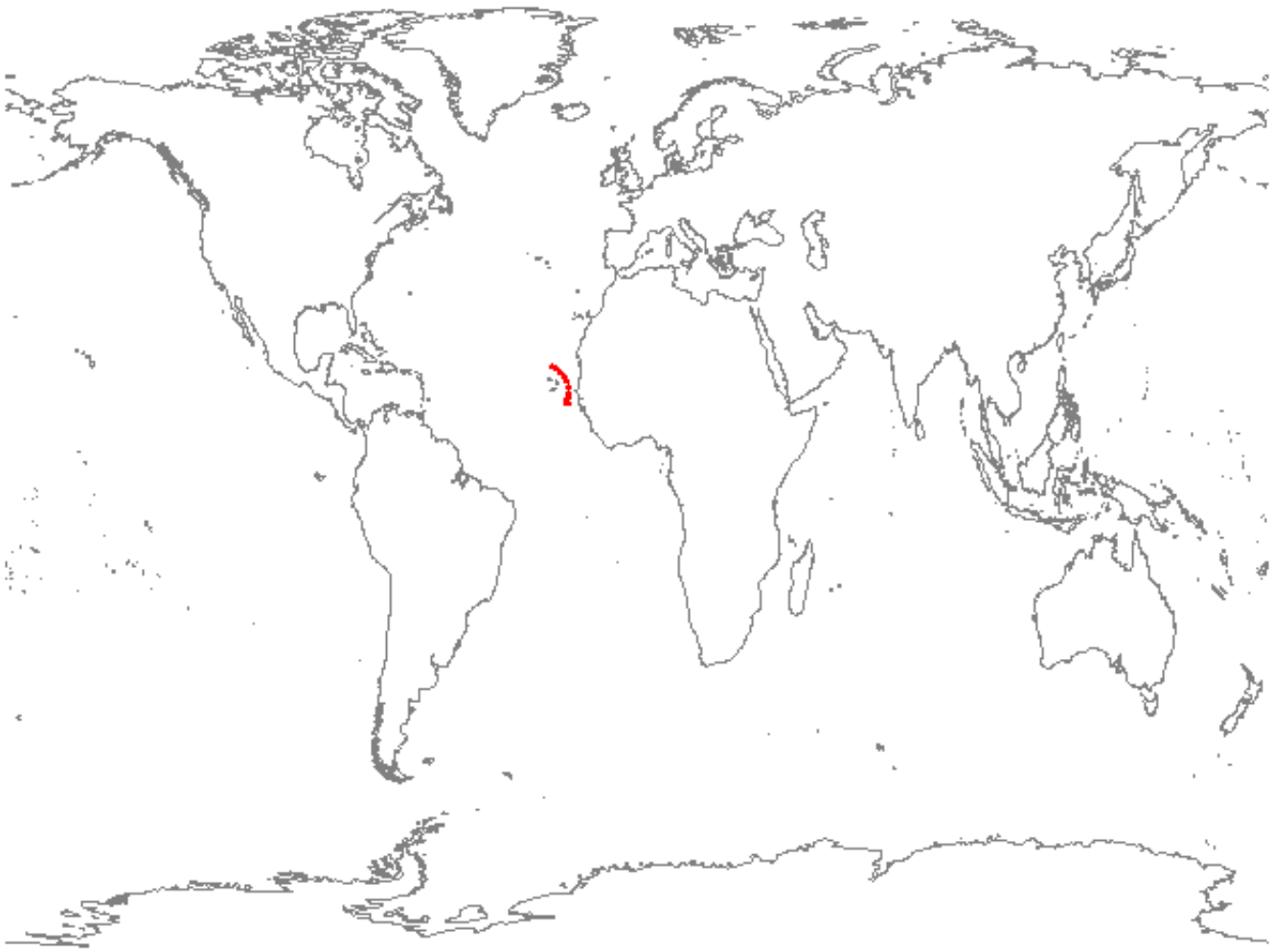

Figure 1: Generated from al102020.kml

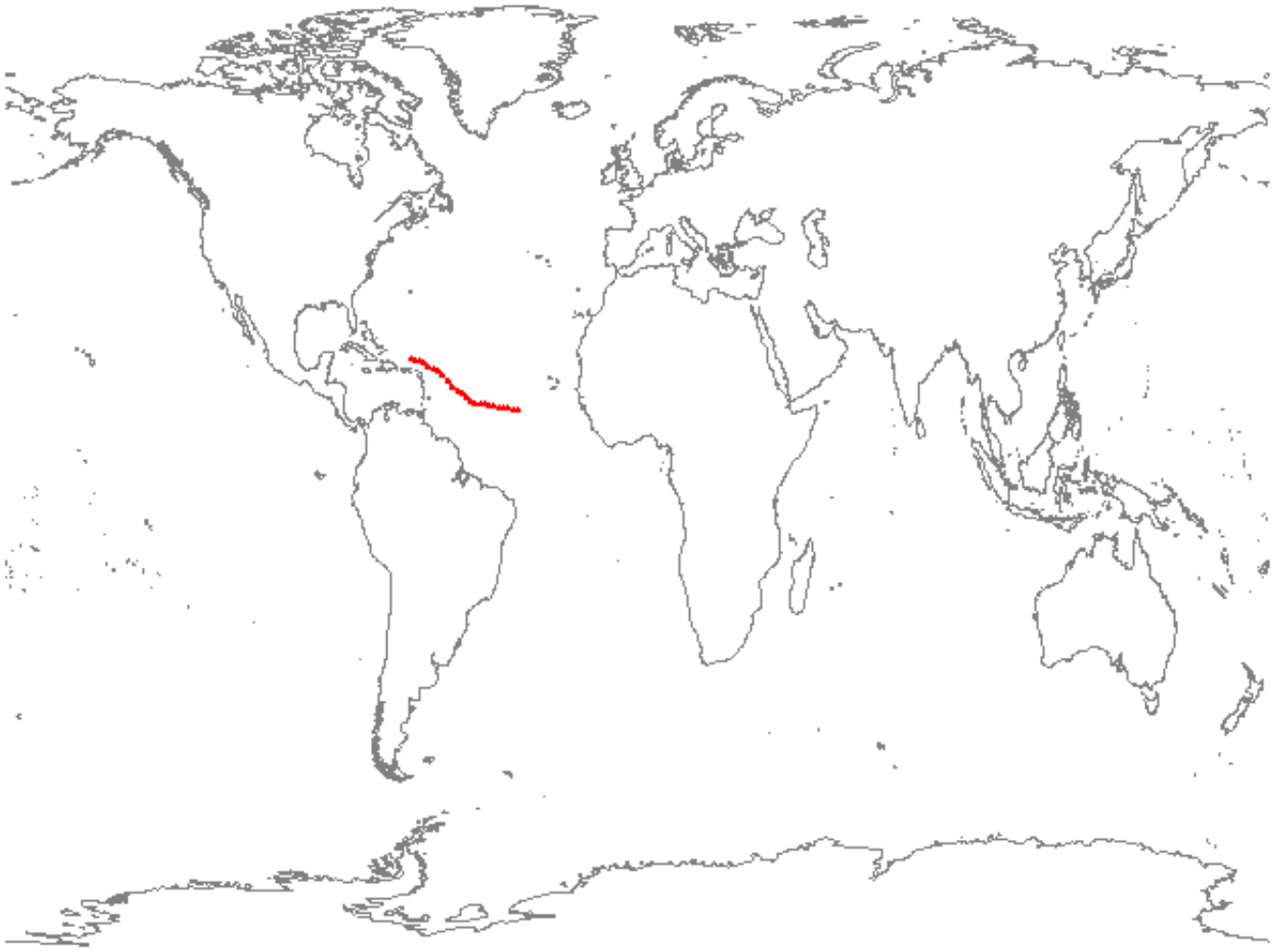

Figure 2: Generated from al112020.kml

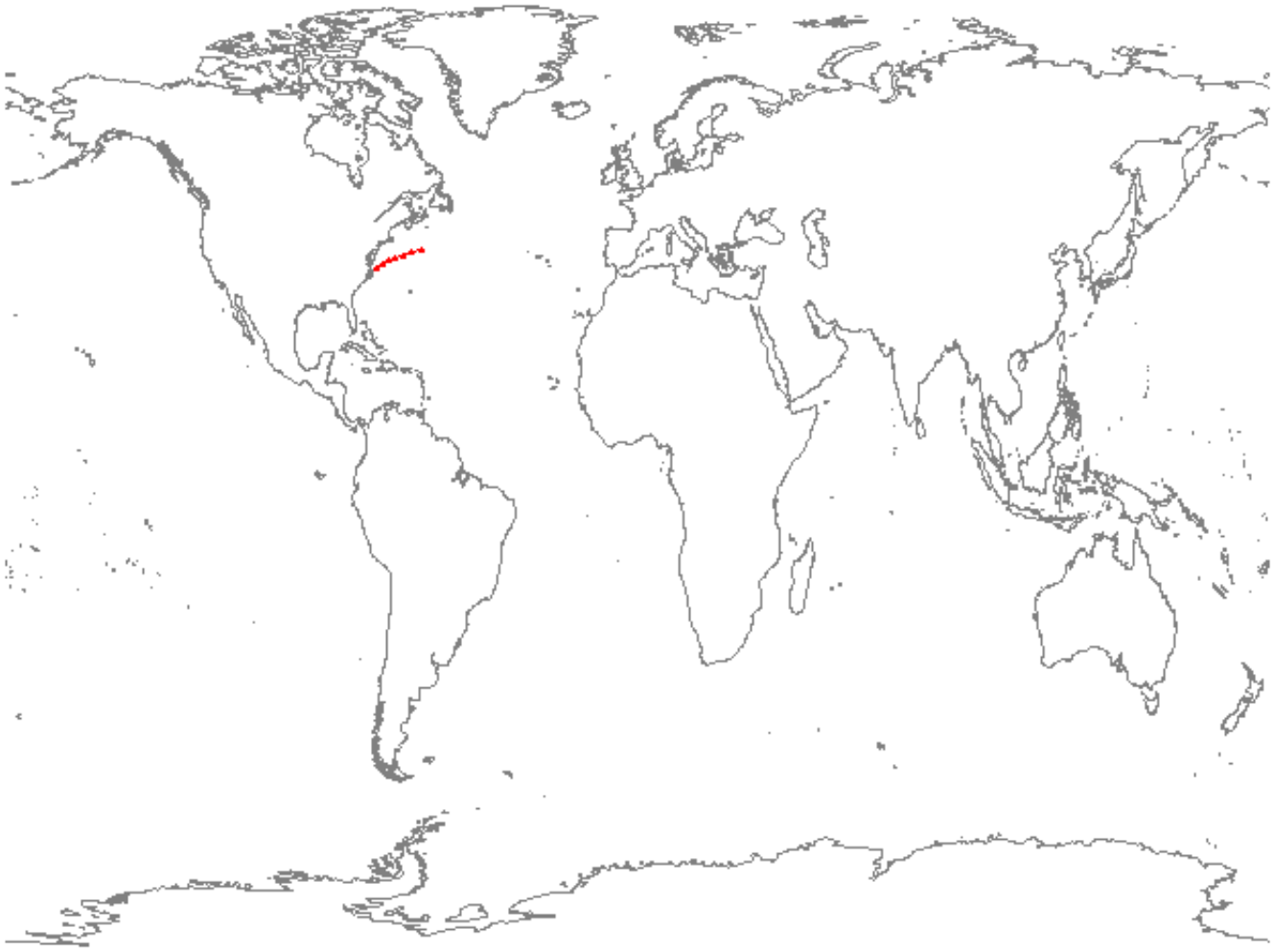

Figure 3: Generated from al122020.kml