Proposal: Unraveling the Link Between Upper-Tropospheric Patterns and Extreme NYC Weather

DATA 450 Capstone

David Lopez

February 5, 2025

1 Introduction

The upper troposphere plays a crucial role in shaping weather patterns across the Eastern United States. Large-scale atmospheric features, such as ridges and troughs, influence temperature extremes, storm development, and precipitation patterns. Changes in these circulation patterns can have profound effects on regional climates, including an increase in the frequency and intensity of extreme weather events.

In recent decades, there has been growing interest in understanding how upper-tropospheric variability is evolving in response to broader climate trends. Shifts in the positioning and amplitude of ridges and troughs can alter storm tracks, contributing to prolonged heatwaves, heavy rainfall, and even disruptions in seasonal weather patterns. However, the extent to which these changes are occurring and their direct connection to extreme weather events in urban centers like New York City remains an open question.

This study will utilize daily geopotential height data at the 500 mb level from the NOAA-CIRES 20th Century Reanalysis (V3) dataset. By identifying the steepest south-to-north gradients in geopotential height, we will track the location of the jet stream and use peak-finding methods to determine the daily positions of ridges and troughs. A long-term analysis of these features will help assess whether their spatial variability has changed over time and whether these shifts correlate with the frequency and intensity of extreme weather events in NYC. The findings from this research could improve our understanding of atmospheric dynamics and help refine future climate and weather prediction models.

2 Dataset

This study utilizes data from the NCEP/NCAR Reanalysis Dataset, a comprehensive atmospheric dataset produced by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR). This dataset provides a global reanalysis of atmospheric conditions, incorporating observational data and model forecasts to reconstruct past weather patterns.

The dataset originates from NCEP's Global Forecast System (GFS), which has been operational since July 1976. The data is continuously updated and processed at NCAR, with additional quality control and error-checking performed to ensure consistency. The dataset is provided as a blend of observational and model data, where missing values are interpolated, and a daily average is created from twice-daily analyses (00Z and 12Z).

For this study, the geopotential height at the 500 mb level is the primary variable of interest. Geopotential height represents the altitude at which a given pressure level is found and is crucial for identifying ridges and troughs in the upper troposphere. These features influence large-scale atmospheric circulation and play a key role in extreme weather events.

The dataset is provided in NetCDF format and includes the following variables:

- Geopotential Height (hgt) – The height (in meters) of the 500 mb pressure surface, used to track atmospheric circulation patterns.

The data is publicly available through the NOAA Physical Sciences Laboratory:
NCEP/NCAR Reanalysis Dataset
Citation:
NCEP/NCAR Reanalysis. NOAA Physical Sciences Laboratory. Available at: https://psl.noaa.gov/data/gridde

3 Data Acquisition and Processing

The dataset used in this study is obtained from the NCEP/NCAR Reanalysis dataset, which is publicly available through the NOAA Physical Sciences Laboratory. The data is downloaded in NetCDF format, a common format for storing multidimensional climate and weather data. Since I have never worked with NetCDF files in Python before, the first step will be learning how to properly read, extract, and manipulate these files using Python libraries such as xarray and netCDF4.

Once I have developed proficiency in handling NetCDF files, the processing of the dataset will involve the following steps:

- **Extracting Geopotential Height Data** – The NetCDF files contain multiple atmospheric variables, but only the geopotential height field at the 500 mb level will be retained for analysis.

- **Identifying the Jet Stream** – The jet stream is located by determining the regions with the steepest south-to-north gradients in geopotential height. This step helps establish the large-scale flow structure of the upper troposphere.
- **Detecting Ridges and Troughs** – A peak-finding algorithm will be applied to the geopotential height field to identify daily ridge and trough locations. Ridges correspond to local maxima, while troughs correspond to local minima.
- **Temporal and Spatial Analysis** – The detected ridge and trough locations will be analyzed over time to assess long-term trends and variability. Any shifts in their frequency, amplitude, or geographic positioning will be evaluated.
- **Data Cleaning and Smoothing** – To reduce noise and ensure robust trend analysis, spatial and temporal filtering techniques may be applied where necessary.

These processing steps will provide a foundation for understanding how upper-tropospheric circulation patterns have evolved over time and their potential connection to extreme weather events in New York City.

4 Research Questions and Methodology

1. **Has the variability of upper-tropospheric ridges and troughs changed over time?**
    - **Method:** Using the extracted 500 mb geopotential height data, I will apply a peak-finding algorithm to detect daily ridge and trough positions. A time series analysis will then be conducted to assess trends and variability in their locations and amplitudes over the dataset's time span.
    - **Expected Output:** Time series plots showing long-term changes in ridge and trough frequency, location, and intensity. Additional statistical metrics will quantify trends.
    - **Estimated Time:** 8-10 hours
2. **Can changes in ridges and troughs be linked to extreme weather events in New York City?**
    - **Method:** Historical extreme weather events in NYC (e.g., heatwaves, heavy precipitation, cold outbreaks) will be identified from meteorological records. Ridge and trough patterns during these events will be analyzed to determine correlations. Composite maps will be created to compare geopotential height anomalies during extreme events versus climatological averages.
    - **Expected Output:** Correlation analysis results, composite maps of geopotential height patterns, and case studies of specific extreme events.
    - **Estimated Time:** 10-12 hours
3. **Have the positions of the jet stream, ridges, and troughs shifted latitudinally over time?**

- **Method:** The jet stream's location will be estimated using the steepest geopotential height gradients. Ridge and trough latitudes will be analyzed over time to identify possible poleward or equatorward shifts. Statistical tests will be applied to determine significance.
- **Expected Output:** Time series and spatial trend analyses showing shifts in ridge, trough, and jet stream positions.
- **Estimated Time:** 8-10 hours

4. **How do changes in the location and intensity of ridges and troughs relate to specific types of extreme weather events (e.g., heatwaves, heavy precipitation) in NYC?**

   - **Method:** This question will involve examining the relationship between observed shifts in ridges and troughs and the frequency and intensity of specific types of extreme weather events in NYC. Statistical analysis and correlation testing will be conducted to evaluate these relationships, and composite maps will be created for visual comparison.
   - **Expected Output:** Correlation analysis results for different extreme weather types, along with detailed maps and case study comparisons.
   - **Estimated Time:** 10-12 hours

These analyses will provide insight into how upper-tropospheric circulation patterns have evolved and their potential link to extreme weather events in New York City.

5 Work plan

Week 4 (2/10 - 2/16):
- Learn how to work with NetCDF files in Python using libraries like xarray or netCDF4 (4 hours)
- Extract 500 mb geopotential height data from the dataset and begin basic exploration (3 hours)

Week 5 (2/17 - 2/23):
- Apply peak-finding algorithm to identify ridges and troughs (4 hours)
- Visualize and analyze initial ridge and trough data (3 hours)

Week 6 (2/24 - 3/2):
- Conduct time series analysis on the variability of ridges and troughs (4 hours)
- Begin gathering historical weather data for extreme events in NYC (3 hours)

Week 7 (3/3 - 3/9):
- Analyze the relationship between ridges/troughs and extreme weather events in NYC (5 hours)

- Begin working on visualizations and composite maps (2 hours)
- Presentation prep and practice (4 hours)

Week 8 (3/10 - 3/16):
- Continue visualizing and analyzing the relationship between ridges/troughs and extreme events (5 hours)
- Prepare and practice presentation of findings (2 hours)
- Poster prep (4 hours)
- Presentation peer review (1.5 hours)

Week 9 (3/17 - 3/23):
- Begin writing the final report, focusing on methodology and results (5 hours)
- Peer feedback on the analysis and interpretation of results (2 hours)
- Poster Draft 1 due Monday morning 3/24 at 9am. Poster Draft 2 due Sunday night 3/30.

Week 10 (3/24 - 3/30):
- Finalize report and prepare presentation materials (5 hours)
- Poster revisions (1.5 hours)
- Peer feedback (1.5 hours)
- Final Poster due Sunday 4/6.

Week 11 (3/31 - 4/6):
- Prepare and practice final presentation (5 hours)
- Revise report based on peer feedback (2 hours)

Week 12 (4/7 - 4/13):
- Final report revisions (5 hours)
- Final presentation preparations and practice (2 hours)

Week 13 (4/14 - 4/20):
- Draft blog post summarizing research findings (4 hours)
- Peer feedback and blog post revisions (3 hours)
- Blog post draft 1 due Sunday night 4/28.

Week 14 (4/21 - 4/27):
- Peer feedback (3 hours)
- Blog post revisions (4 hours)

Week 15 (4/28 - 5/4):
- Blog post revisions (2 hours)
- Peer feedback (2 hours)

6 References

NCEP/NCAR Reanalysis. NOAA Physical Sciences Laboratory. Available at: https://psl.noaa.gov/data/gridde