Using OSDC to Advance Public Understanding of Temperature Extremes

Race Clark*, Christopher Natoli†, Jill Hardy*, William Matthews‡
*University of Oklahoma, Norman, OK 73019
†University of Chicago, Chicago, IL 60637
‡Fisk University, Nashville, TN 37208

***Abstract*—The Open Science Data Cloud (OSDC) provides great utility for accessing archived meteorological data from the NOAA National Climatic Data Center (NCDC). Using this data, we propose a product that will make relaying temperature extremes more meaningful and visual to the public. This is achieved by creating histograms of daily minima and maxima temperatures from land-based instrumentation, at different time scales.**

I. Introduction

Television and radio broadcasts are two common ways in which meteorological information is relayed to the public. The majority of Americans consider local television and radio their primary source for weather. In recent years, the field has grown to include several national cable and satellite networks, as well as online methods (e.g. video broadcasts). In fact, The Weather Channel is the most widely distributed cable channel in the United States [1]. With broadcast meteorology having such an expansive role across the country, it is important for on-air meteorologists to provide context and expertise to the viewing public.

A common way of adding context to weather information is the inclusion of daily temperature maxima and minima, especially if either of these values is close to setting a record. But more often than not, these values are not supported by any information about the frequency or magnitude of such events. For example, in the month of July in Oklahoma, how unusual is a daily maximum of 50°F if the typical daily high is 95°F? Providing this kind of information can be useful for the public and experts to properly interpret temperature anomalies and to place them into an appropriate historical context. However, it can be difficult for meteorologists to not only find these data, but more importantly, to display them in a way that is informative and readily understandable.

The Open Science Data Cloud (OSDC) is a very useful infrastructure that is capable of finding and displaying temperature extremes data easier for use by meteorologists and climatologists [2]. Since OSDC is a publically-available data cloud for managing large datasets, the historical temperature data of all global first-order surface weather observation points can be easily accessed and visualized using the OSDC's infrastructure. The rest of this paper will describe the datasets and methodology necessary to achieve this aim, as well as preliminary results from the research.

II. Data and Methods

A. Temperature Data

The National Oceanic and Atmospheric Administration's (NOAA) National Climatic Data Center (NCDC) maintains archives of daily meteorological observations from thousands of in-situ station instruments across the globe. These datasets are available on the OSDC console for public use. The dataset utilized in this study is the Global Historical Climatological Network (GHCN), which compiles information from over 20 different sources of ground-based instrumentation [3]. The length of record changes for every station, with some dating back to the 1800's. The dataset is reconstructed once a week to ensure the most updated information, and is also quality controlled at this time. For our purposes, the only variables we collected were the daily minimum and maximum temperatures, in tenths of degrees Celsius.

B. Methodology

To explain the concept of this project, seven GHCN stations were chosen that represent a varierty of climatological regimes from around the United States. The stations, as well as the start and end dates of their records of daily temperature extremes, are shown in Table I.

Station Name	Record Start Date (MM/DD/YYYY)	Record End Date (MM/DD/YYYY)
Chula Vista, CA	09/01/1918	06/12/2014
Monterey Peninsula Airport	12/01/1968	06/14/2014
Oklahoma City, OK	01/01/1948	06/14/2014
Miami Beach, FL	01/05/1927	06/15/2014
Vienna, VA	10/04/1948	06/14/2014
Chicago Midway Airport 3 SW	02/29/1928	06/15/2014
Memphis International Airport	01/01/1940	06/14/2014

TABLE I: Global Historical Climatological Network (GHCN) stations used in this study, with the start and end dates of the temperature records.

Once the data were collected from the OSDC's mirror of the NCDC's public-facing FTP server, the max temperature data were converted back to their original units of degrees Fahrenheit. The investigators used `matplotlib`, a commonly available Python library, to create histograms of maximum temperatures at the observing sites of interest. It is a relatively trivial matter to extend this process to minimum temperatures, as well as all U.S.-based GHCN sites and, indeed, those across the entire world. Additionally, histograms of all maximum temperature observations in a particular station record can be plotted, or the data can be segregated into useful subunits, including seasons, months, or even individual days. The Python code to create these plots is available at: https://github.com/chrisnatoli/max_temperatures.

III. Results

The Chicago heatwave of March 2012 attracted worldwide attention for producing never-before-observed temperature extremes, including the area's earliest ever occurrence of +80°F maximum temperatures (Figure 1). Although the Chicago site has only been operational for 86 years, it is clear that the March 18, 2012 temperature of 81°F is a significant outlier. This is even more apparent in Figure 2, where all of the 80°F and higher temperatures ever recorded in the month of March in Chicago occurred in March 2012.

The yearly maximum temperature distribution in Chicago is bimodal and shows high variation (Figure 3). This is particularly evident in late winter and early spring, as seen in Figure 2. Seasonal variability is quite typical of a continental climate and was seen amongst other continental sites. On the other hand, the Mediterranean climate of Monterey, CA (Figure 4) and the subtropical climate of Miami Beach, FL (Figure 5) are unimodal with much less variability. These characteristics can be attributed to their close proximity to large bodies of water.

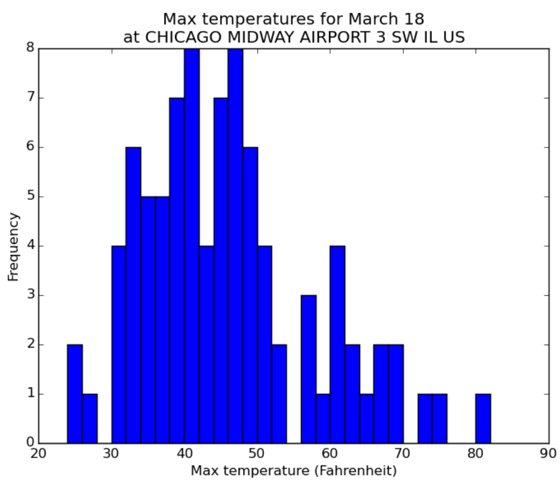

Fig. 1: Histogram of all March 18 maximum daily temperatures from 1928-2014 at the Chicago Midway Airport 3 SW station site.

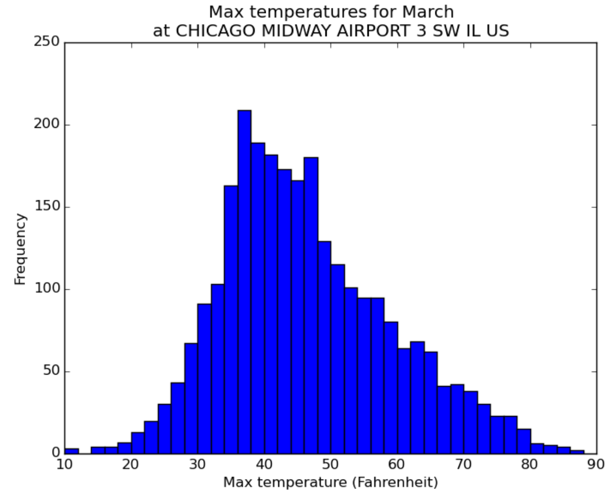

Fig. 2: Histogram of all March maximum daily temperatures from 1928-2014 at the Chicago Midway Airport 3 SW station site.

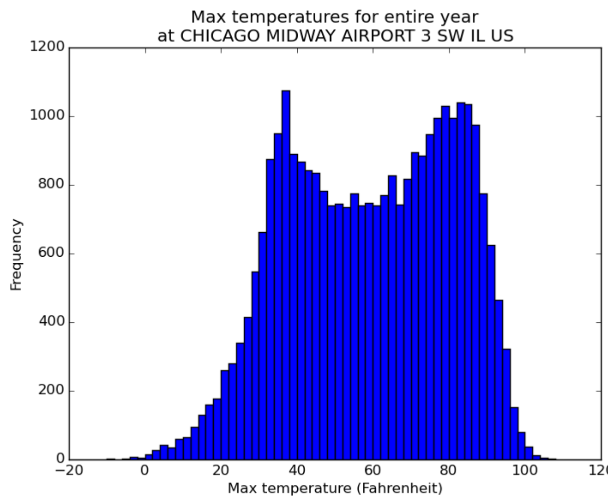

Fig. 3: Histogram of every maximum daily temperature from 1928-2014 at the Chicago Midway Airport 3 SW station site.

IV. Summary

The patterns presented here do not begin to scratch the surface of the powerful relationships this large and rich dataset contains. Future work should explore the overlay of statistcal information including standard deviations on the histograms. This work could be developed into a public-facing OSDC web service that would undoubtedly be of interest to hobbyists as well as professional meteorologists, climatologists, and the public at large.

Acknowledgment

This work made use of the Open Science Data Cloud (OSDC) which is an Open Cloud Consortium (OCC)-sponsored project. The OSDC is supported in part by grants from Gordon and Betty Moore Foundation and the National Science Foundation and major contributions from OCC members like the University of Chicago. The authors were supported by the National Science Foundation Partnerships for Research and Education (PIRE) Award Number 1129076. Any opinions, findings, and conclusions or recommendations expressed are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

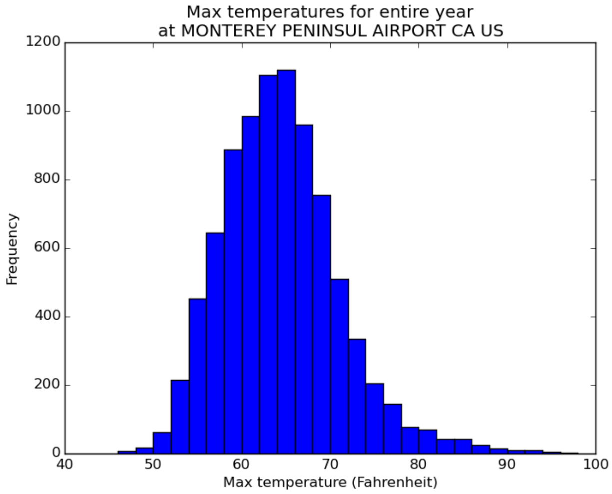

Fig. 4: Histogram of every maximum daily temperature from 1968-2014 at the Monterey Peninsula Airport station site.

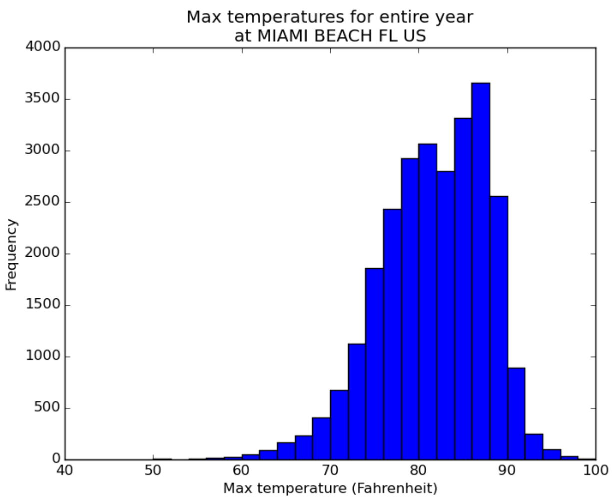

Fig. 5: Histogram of every maximum daily temperature from 1927-2014 at the Miami Beach, FL station site.

REFERENCES

[1] R. Seidman. (2013, Aug.) List of how many homes each cable networks is in - cable network coverage estimates as of august 2013. [Online]. Available: http://tvbythenumbers.zap2it.com/2013/08/23/list-of-how-many-homes-each-cable-networks-is-in-cable-network-coverage-estimates-as-of-august-2013/199072/

[2] R. L. Grossman, M. Greenway, A. P. Heath, R. Powell, R. D. Suarez, W. Wells, K. P. White, M. P. Atkinson, I. A. Klampanos, H. L. Alvarez, C. Harvey, and J. Mambretti, "The design of a community science cloud: The open science data cloud perspective," in *SC Companion*. IEEE Computer Society, 2012, pp. 1051–1057.

[3] J. Burroughs. (2009) Global historical climatology network (ghcn). [Online]. Available: http://www.ncdc.noaa.gov/oa/climate/ghcn-daily/