NYC Climate Change - Analysis of Temperature Changes

This independent study analyzes daily temperature data for NYC over a 120+ year period (1901-Oct2022) to assess whether daily temperatures have been changing over time. This study has been motivated by the World Meteorological Organization's (WMO) briefing on the Provisional State of the Global Climate in 2022 and by this NYTimes article. The NYTimes article is in turn based on this paper by Dr. Hansen et al., and many of the research techniques within this project are guided by the methods used in the paper.

The data for this study has been obtained from the National Oceanic and Atmospheric Administration (NOAA) website here. The dataset contains ~150 years of daily temperature data (Mar 1,1869 - Oct 27, 2022) from between 1-5 weather stations that NOAA classifies under the New York County. For each day, the data contains a minimum (TMIN) and maximum (TMAX) temperature readings; these readings have been averaged across the stations to generate one daily minimum and maximum reading. These daily minimum and maximum temperatures are further averaged to generate a daily average temperature (TAVG).

Comparison Period vs Baseline Period - Analysis of Daily Temperature Differences

In order to analyze temperature changes, daily observed temperatures are compared to a baseline daily temperatures. As explained here, generally, a 30year or longer period would suffice as the "baseline period" as it is sufficiently long for meaningful statistical analysis.

For this study, a ~30year "baseline period" has been chosen from 1869 to 1900[1]. This is close to the 1850-1900 pre-industrial baseline used in the Paris Agreement Assessment report, per here. The baseline daily temperature is calculated by averaging the temperatures for that day for each year in the baseline period. For example, the June 1 baseline temperature would be calculated as, the average of the June 1 temperatures seen for every year in the base period, i.e., one for each year from 1869 through 1900 or 32 readings. This process is repeated for each day and in this way we will have 366 baseline temperatures, one for each day of the year.

The comparison dataset for this study is from 1901 through Oct 2022. The daily temperature for each day in the comparison period are compared to the baseline temperature for that day. For example, the June 1 2022 average temperature (TAVG) is compared to the June 1 baseline average temperature (Base_Avg). If there has not been a material change in temperatures over the years, the expectation is that for a majority of days, this difference should be close to zero. The distribution of these differences is analyzed by decade[2] in Figure 1.

Fig1: Comparison Period v/s Baseline Period - Distribution of in Daily Temperature Differences
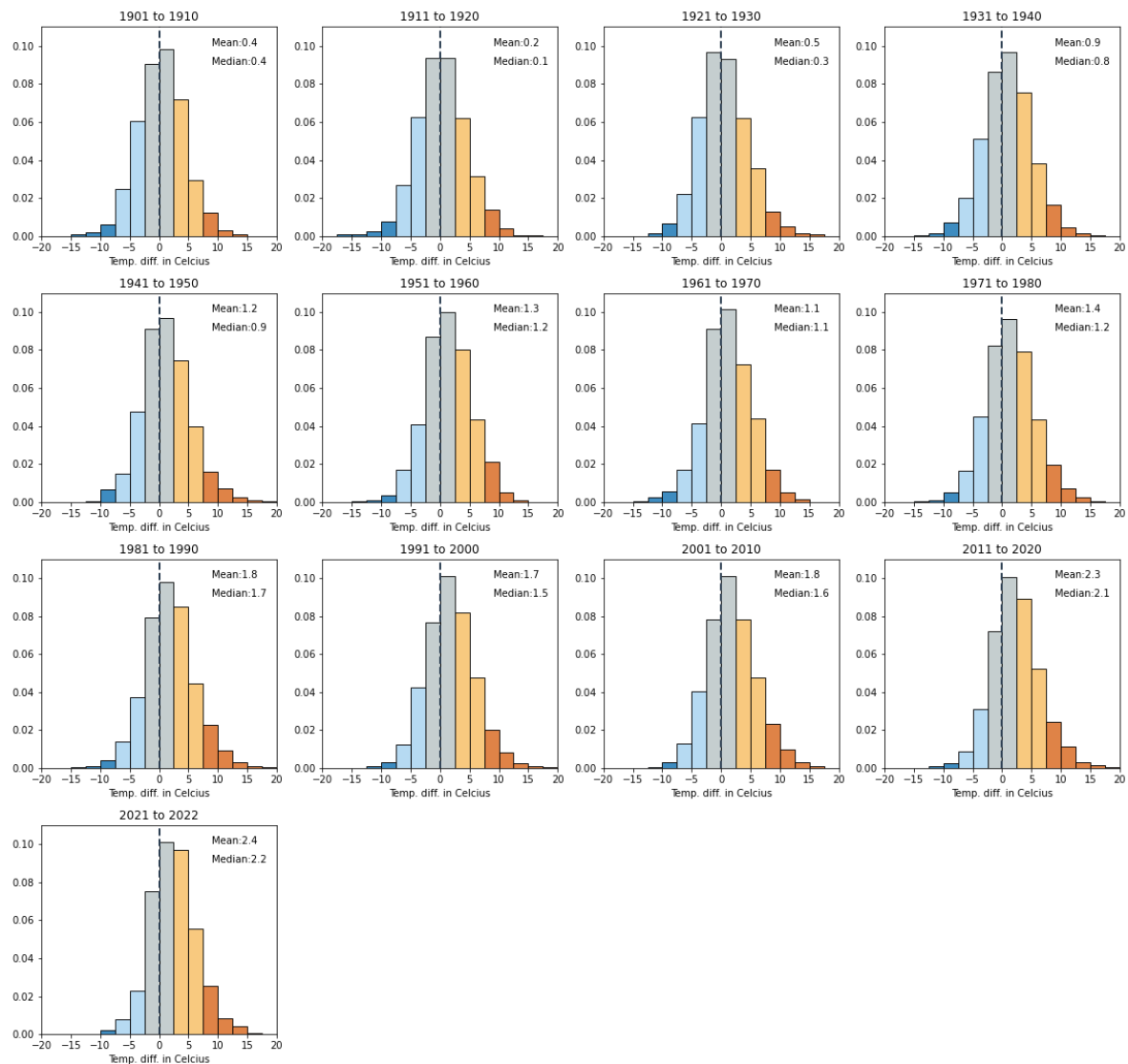
Each subplot within Figure 1 above, shows the **distribution of the differences** between the daily temperatures observed in that decade and the corresponding baseline temperatures expected for that day. Bars on the right of 0 indicate positive difference, i.e., warmer daily temperature in that decade when compared to the baseline. Similarly, the left of 0 would indicate cooler daily temperatures in that decade when compared to the baseline. In this study, differences of +/-2.5C to +/- 7.5C against the baseline, are considered large differences (amber for warmer temperatures and light blue for cooler temperatures; and differences greater than +/- 7.5C against the baseline, are considered very large differences (red for warmer temperatures and dark blue for cooler temperatures)

The key observations:

- For each decade the daily temperature differences against baseline appear to be normally distributed. In the initial decades of the 20th century, this normal distribution appears to be centered at 0, i.e, mean 0 implying on average, no differences with baseline. But as we move through the decades, we see the normal distribution moving more to the right, i.e., warmer temperatures or positive differences; and 2011 onwards, the distribution mean is

closer to ~2.4 (i.e., on average, temperatures are ~2 times higher than the baseline period). This movement can also been seen in Figure 2 below.

- Over the decades we can also see an increase in "extremely large" positive differences, implying significantly warmer temperatures. Further analysis on extreme temperatures is shown in Figure 3.
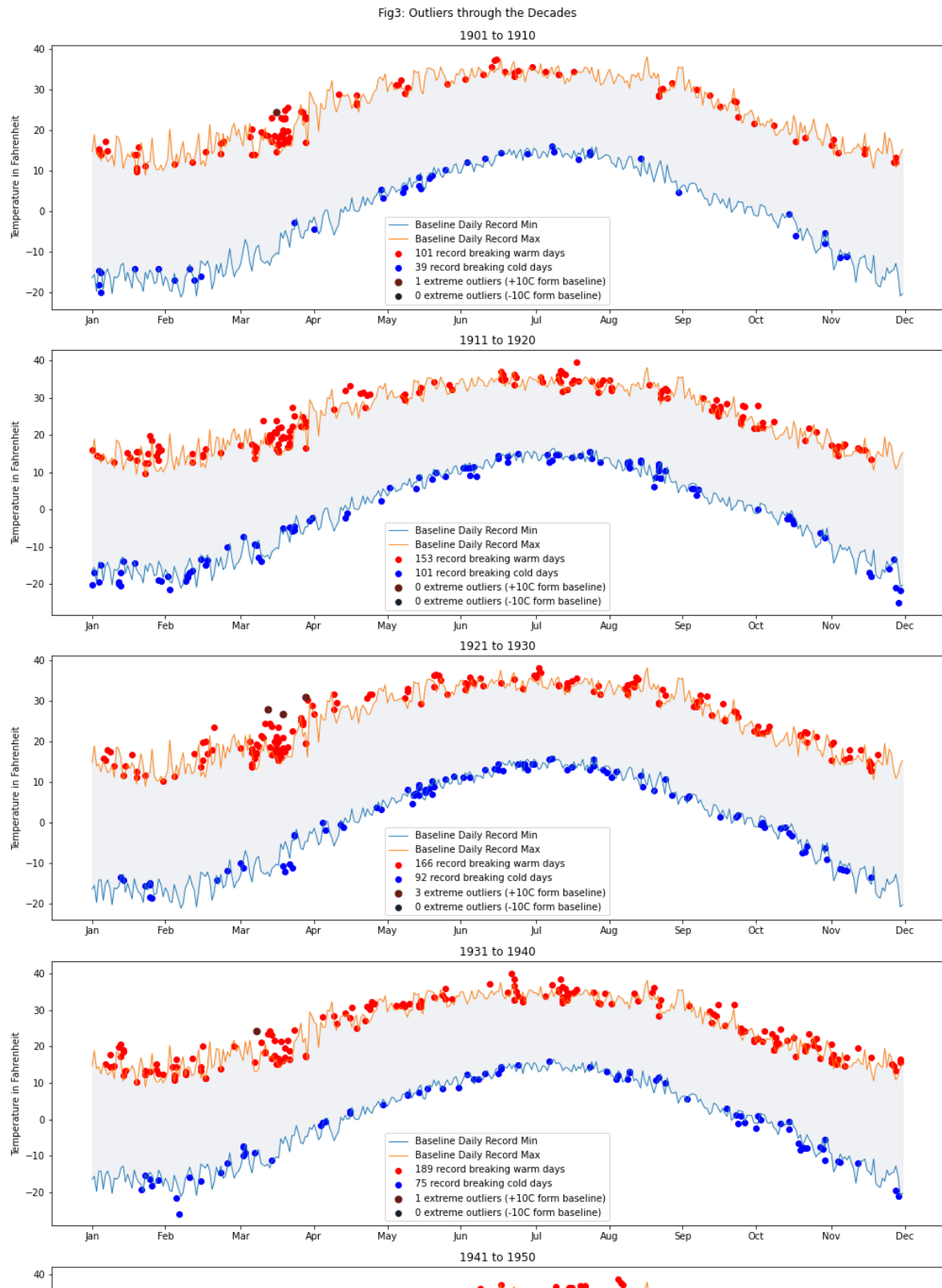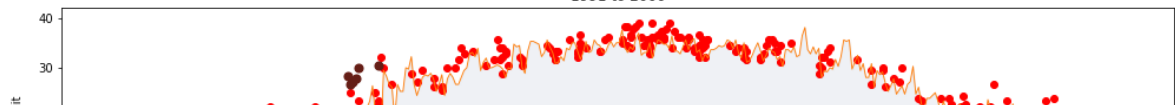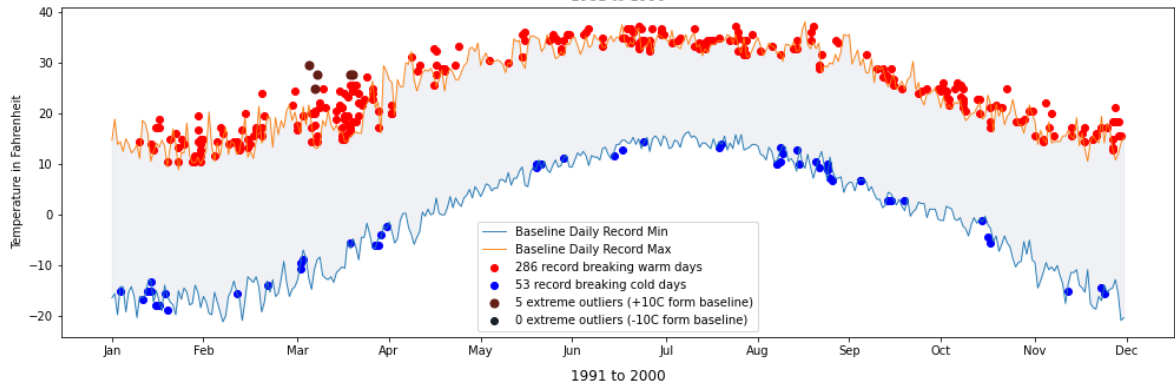

Fig2: Comparison Period v/s Baseline Period - Distribution of in Daily Temperature Differences
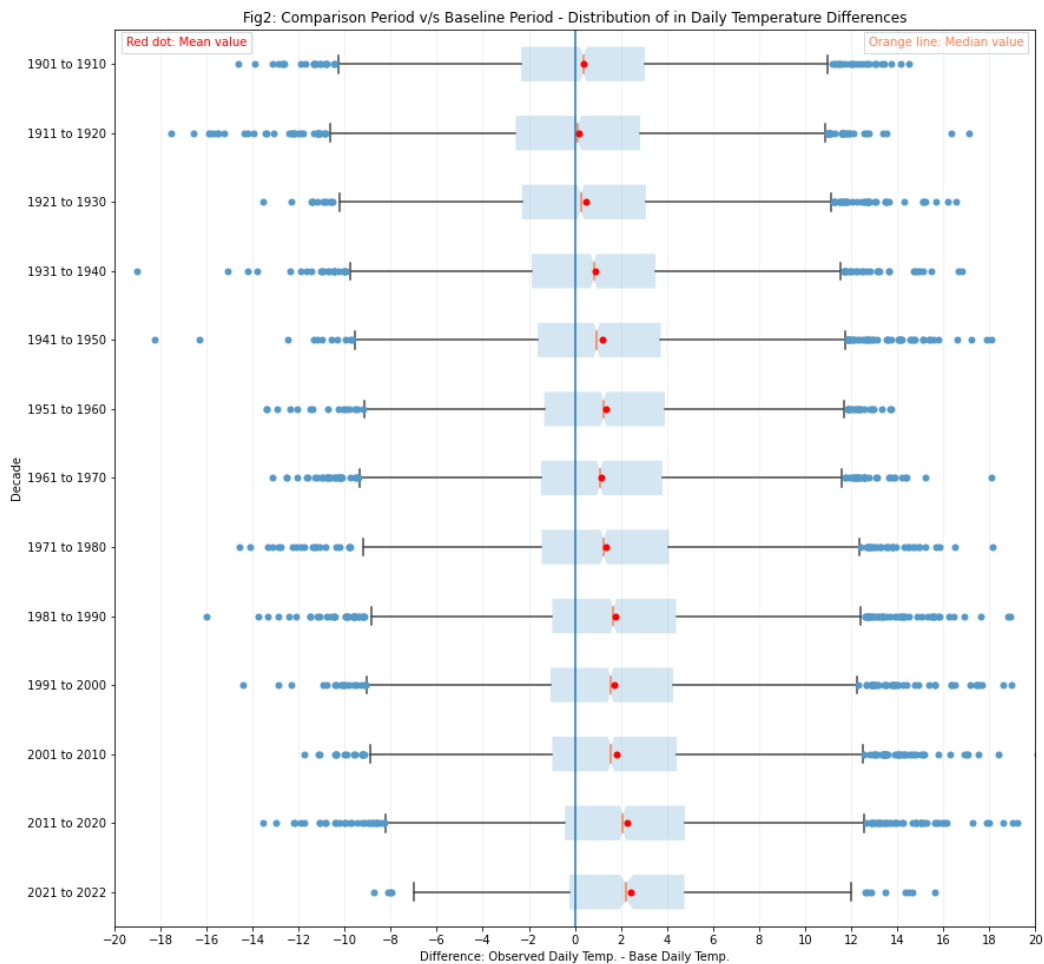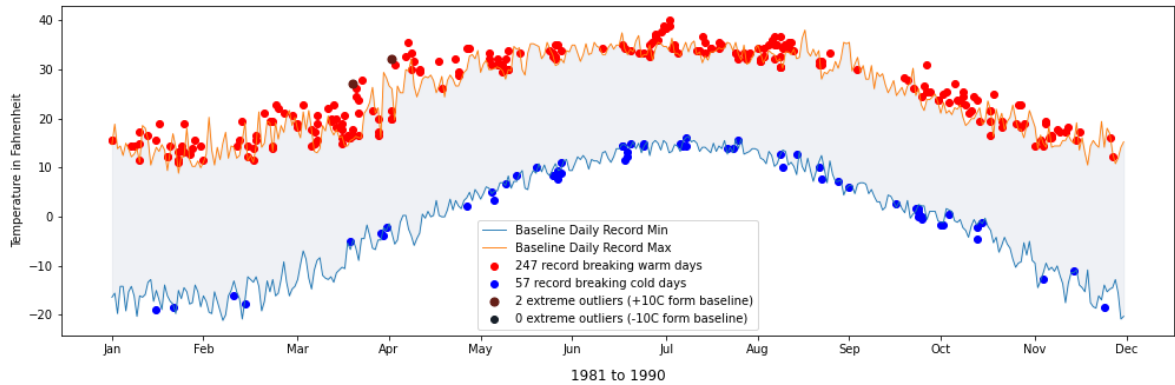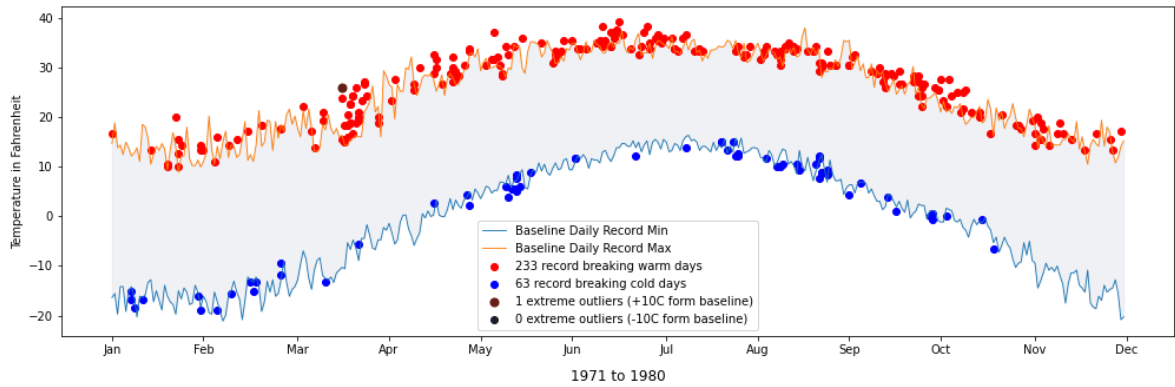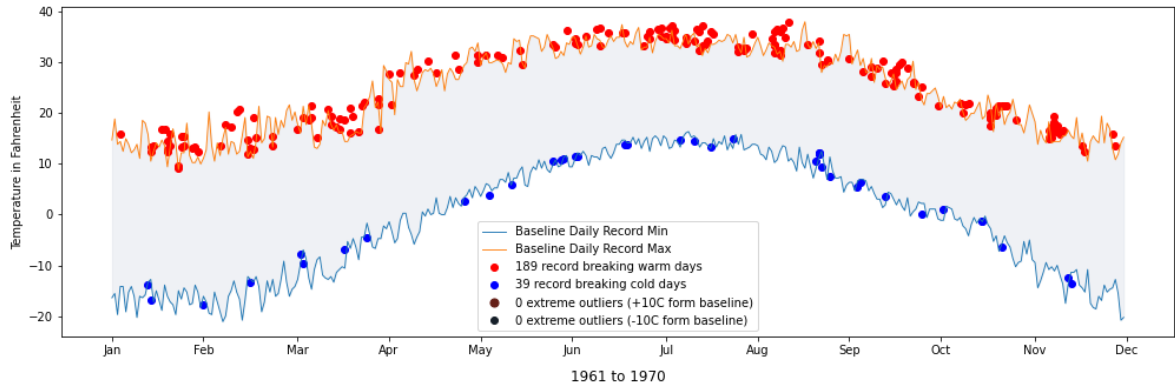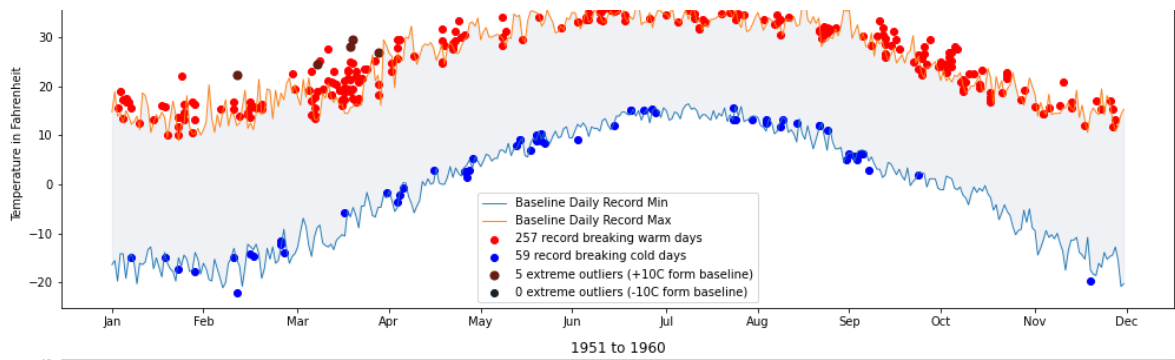
Additional Analysis: Outliers

Record max and min temperatures seen in the baseline period are used to create a range for comparison. The upper bound is the maximum temperature seen for a day during the baseline period (Base_RecordMax) and the lower bound is minimum temperature seen for that day during the baseline period (Base_RecordMin). Thus, this is the broadest range we can create based on the temperatures observed during the baseline period.
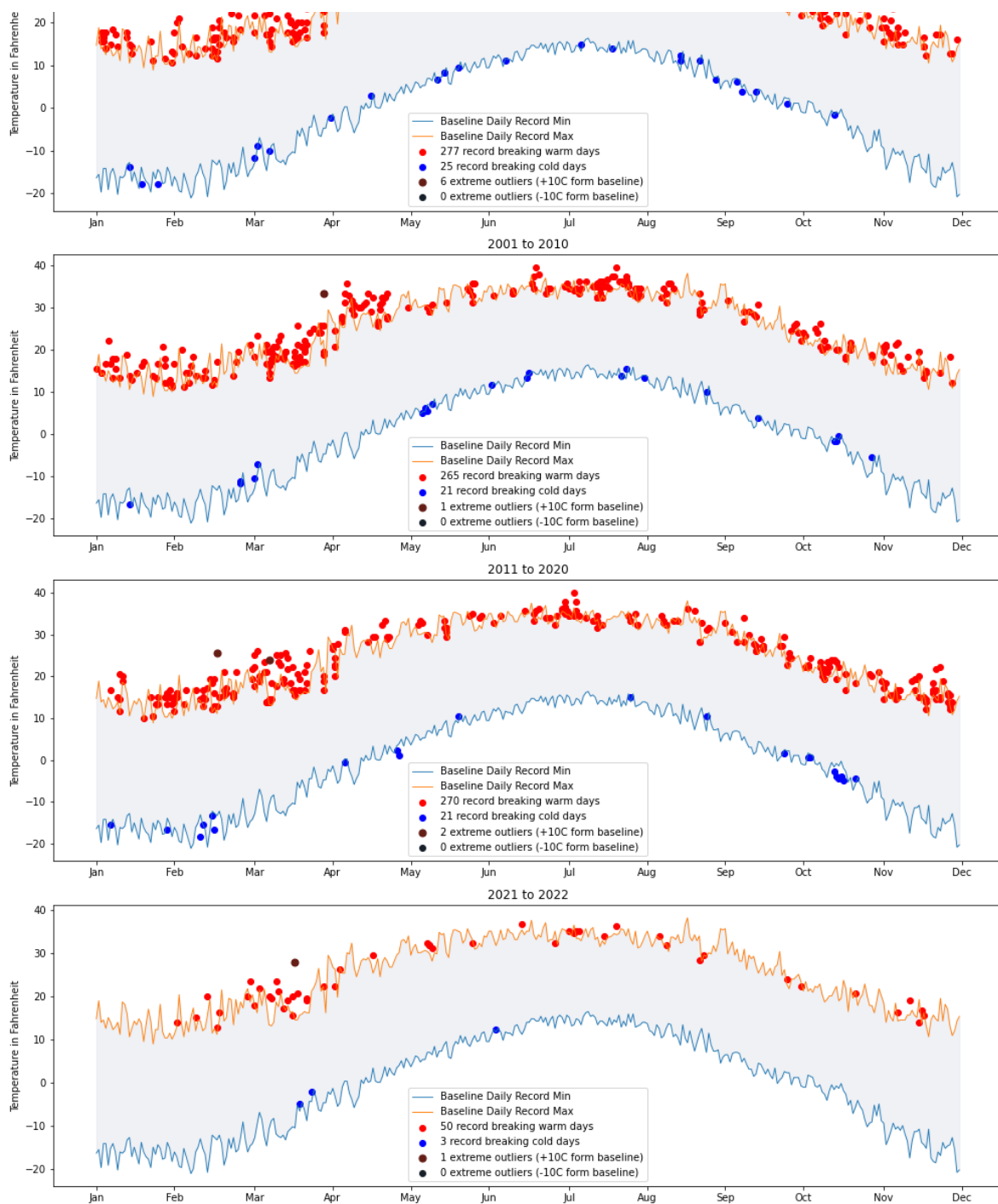
Outliers then, are any daily max/min temperatures outside of this range in the comparison period,i.e., where TMAX > Base_RecordMax or TMIN < Base_RecordTMin. For example, if the June 1 temperature for any year in a decade is greater than the record max baseline

temperature for June 1, it would show as an outlier in Figure 2. **Extreme outliers**, are any daily max(min) temperatures that are more than 10C higher(lower) than the corresponding baseline record temperature for that day.

Through the decades, we see cooler temperature outliers decreasing compared to higher temperature outliers.

Fig3: Outliers through the Decades

1961 to 1970

1971 to 1980

1981 to 1990

1991 to 2000

Reference Papers:

https://www.pnas.org/doi/epdf/10.1073/pnas.1205276109

https://www.climate-lab-book.ac.uk/2016/how-to-quantify-changes-in-climate-extremes-without-inducing-artefacts/

https://www.nytimes.com/interactive/2018/05/03/learning/08WGOITGraphLN.html

https://www.nytimes.com/interactive/2021/climate/extreme-summer-heat.html

https://public.wmo.int/en/our-mandate/climate/wmo-statement-state-of-global-climate

Footnotes:

1: Note, we do not have data for NYC prior to 1869 in NOAA

2: As expected, the last "decade" in the comparison dataset is not a true decade; we only have about ~1.5 years of data (Jan 2021 - Oct 2022)