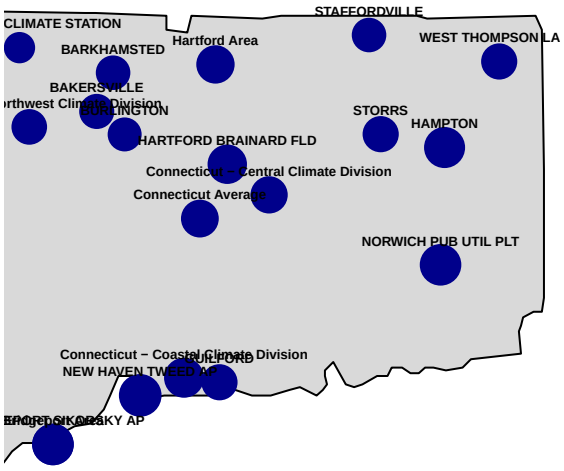
Homework 2 - Replicating Connecticut Temperature Graphics

Jack Bienvenue

2025-08-25

Please note that "CT3456" was not available when I was downloading data. Also note that I chose to calculate the daily mean temperature as the average of the high and low temperature for a station each day.

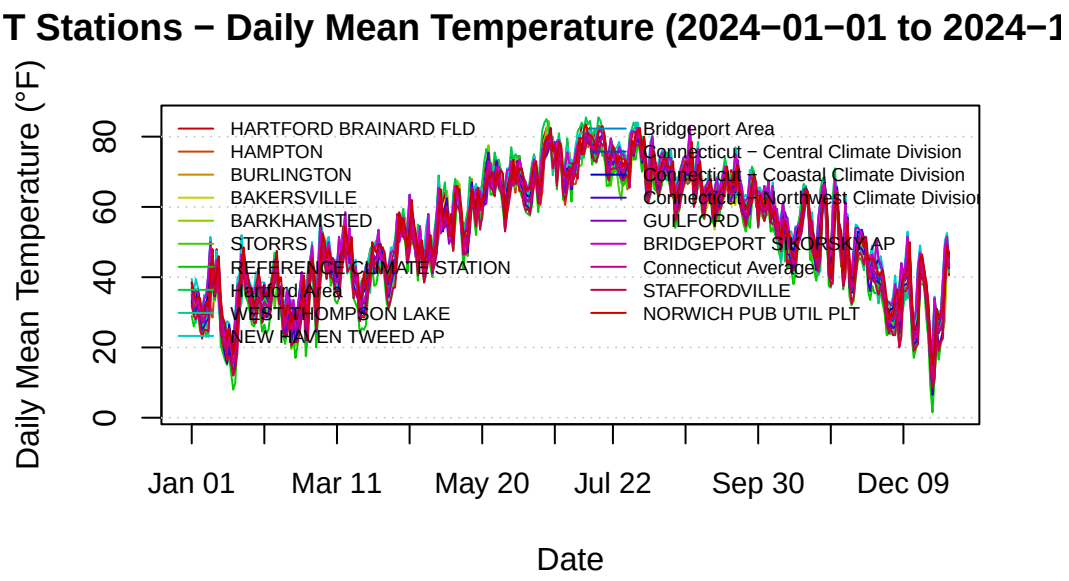
Question 1: Replication of CT Temperature Map



Question: What can you say about the spatial distribution of temperature over CT?

It is very difficult to make out quite what the example is expressing as there is neither a legend available nor a title to describe the data. I will assume that the size of the dots are meant to represent temperature at different weather stations. We observe higher temperatures towards the Connecticut River Valley (roughly following a line of latitude through Hartford) and along the coast. We observe lower temperatures in the higher elevation, continental data points in the Northwest and Northeast in this map. In my replicate of the plot, I mapped temperature on January 1, 2024. We observe low temperature heterogeneity for that day.

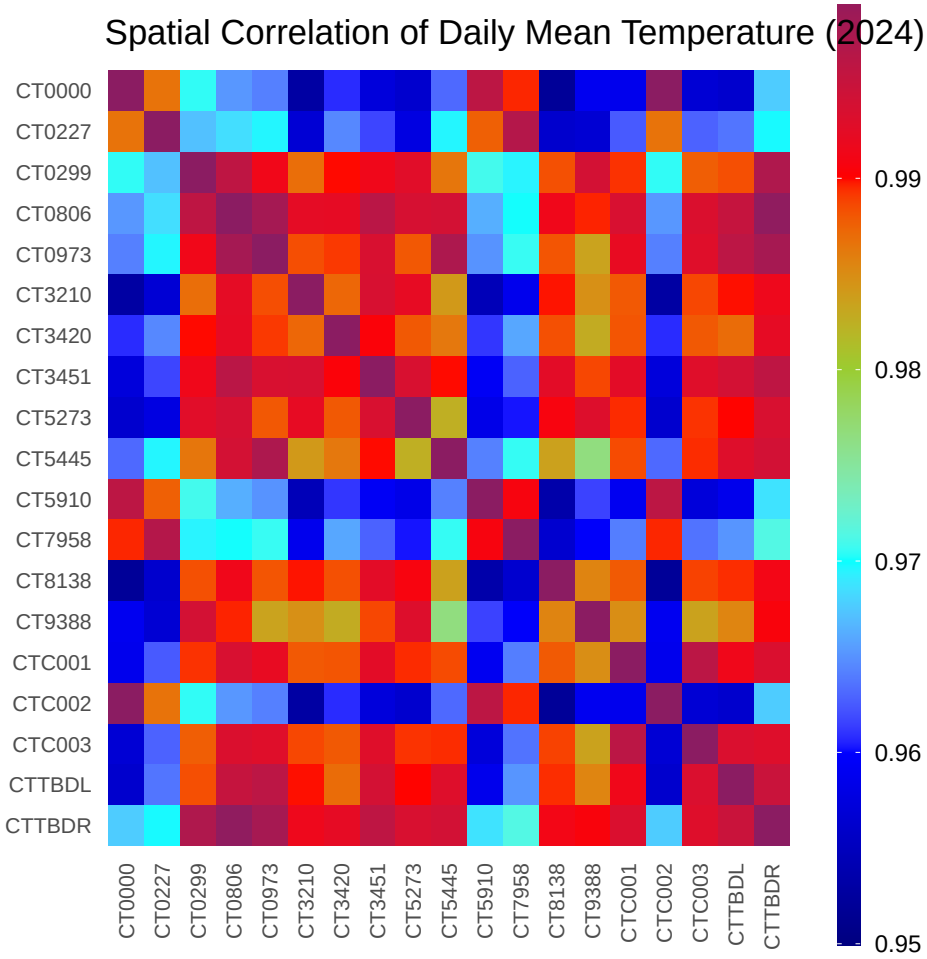
Question 2: Replication of Daily Mean Temp. Chart



Question: What can you say about the spatio-temporal distribution of temperature over CT?

In this plot, it appears as though there is a high level of correlation between different temperature stations in Connecticut. In nearly all instances, temperatures move together between the set of twenty weather monitoring station. In such a small region of interest, we would expect this kind of homogeneity. Upon looking closer, one can observe that there is some recurrence in the highest and lowest temperature regions. Essentially, this means that some locations in Connecticut are relatively consistently warmer or cooler than others. I attempted to use base R to generate a replication of this plot–as it appeared to be constructed in base R–but had some troubles with the finer detials. I hope that this suffices.

Question 3: Replication of Spatial Correlation of Daily Mean Temps.

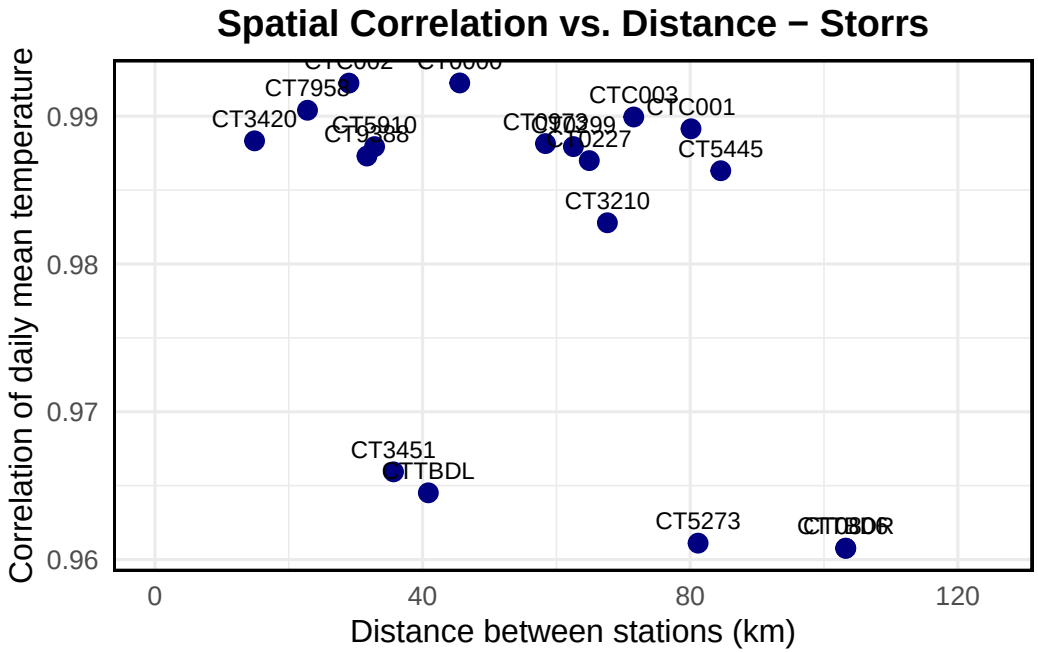


Question: If Storrs data is missing, which station could give the most reasonable prediction for Storrs?

The Storrs weather station is encoded as "CT8138." Looking in the row or column referencing "CT8138," we can identify the highest-correlation weather stations. These appear to be "CTBDR," "CT3451," "CT3456," and "CT0806". These represent Bridgeport Area, Hartford Brainard FLD, Hartford, and Bridgeport's Sikorsky airport, respectively. This is an interesting result, considering that there are some more proximate options for weather station pairs. However, we find that over the course of 2024, Bridgeport and Hartford provided the highest correlation match for temperatures, and therefore they would be very reasonable temperature predictions for Storrs. Note that correlation is very high among all stations, so it would be difficult to make a poor prediction using another CT station. Also, please note slightly different appearnace in correlation matrix due to attempting to match color scaling in extremely narrow range of values.

Question 4: Replication of Spatial Correlation v. Distance Chart

```
Warning: The `size` argument of `element_rect()` is deprecated as of ggplot2 3.4.0.
i Please use the `linewidth` argument instead.
```



Question: Can we see the law of geography in effect here?

Based upon Tobler's Law of Geography, we expect distance and correlation of a variable to have an inverse relationship. This relationship is not as clear as we would expect as visualized in this plot. In fact, it would be difficult to justify the claim that this plot gives any indication of the Law of Geography. With some domain knowledge of temperature, we may realize that there could be latent factors affecting the results we see here, like elevation.