Statistical Validation of a High Resolution Temperature Prediction Model for New York City

Louis Waxman[1,2], Dr. Brian Vant-Hull[2,3], and Dr. Reza Khanbilvardi[2,3]

[1]Ardsley High School, [2]CUNY CREST Institute, [3]The City College of New York

Abstract

A high resolution urban temperature model with a resolution of 100 meters was developed by Dr. Brian Vant-Hull and his team at the NOAA CREST Center at The City College of New York. To create the model, field campaigns recorded temperature data and other information. A multivariable linear regression was performed based on collected temperature data during field campaigns against surface characteristics of elevation, water fraction, vegetation, albedo, building height, and building area fraction by location to create a model of standard deviations from the mean temperature. This model imports a uniform weather prediction daily for New York City as a whole and then produces standard deviations from the mean based on the model created by the multivariable linear regression of temperature anomaly amplitudes versus weather variables. These standard deviations are then converted to a temperature with reference to the uniform predicted temperature. Creating a high resolution temperature prediction model allows city planners to more accurately and precisely understand the differences in temperature across a municipality; this is especially important in the summer due to heat-related

night. The objective of this summer internship research was to validate the high resolution temperature prediction model (described above) for New York City.

Data Collection

During the summers of 2012 and 2013, field agents collected temperature, relative humidity, and illumination during a series of 19 walking campaigns along 8 fixed routes in Manhattan at a uniform 1.5 meters above ground. Additionally, between June 24 and September 20, 2013, ten sets of temperature, relative humidity, and illumination sensors were placed on lampposts around Manhattan between 3.1 – 3.9 meters above street level with the approval of the New York City Department o… on recorded data … es.

Sensor Locations:

145th St.
120th St. West
120th St. East
81st St.
57th St. West
57th St. East
35th St.
14th St.
Prince St.
Reade St.


Figure 1


Figure 2
A typical sensor installation

Temperature Anomaly Model Development

The data collected during the summers of 2012 and 2013 were converted to standard deviations from the average by geographic location. Second, the standard deviations from each day were averaged together. Next, the average standard deviations were regressed against surface features such as elevation, vegetation, and building fraction. Based on these surface features, a map of the difference of temperature from the average scaled by standard deviation is produced, not temperature itself. A uniform temperature prediction is imported as the mean and then the model generates the high resolution temperature prediction based on the standard deviation map.

Analysis Method

The analysis was done using Matlab and Excel. Data for the observed temperature at each of the 10 fixed stations during the observation period at 14:00 were extracted for each day from June 24 to September 21 2013. Next, the standard deviation from the mean temperature of all predicted temperatures for each day was extracted. Following this, the predicted standard deviation from the mean temperature for each of the ten locations was extracted from the model based on the latitude and longitude of each station. To calculate the predicted temperature at each location, the following formula was used: T(x) = Tavg + Tsd*map(x), where T(x) is the predicted temperature at location x, Tavg is the uniform observed temperature, Tsd is the calculated dispersion of the data from the mean, and map(x) is the predicted standard deviation of the temperature from the mean at the given location (x).

Second, the standard deviation of the observed temperature versus the uniform average forecast temperature across all ten locations for each day over the 90 day period as calculated. This was done by setting T(x) equal to Tavg across all ten locations for each day and repeated for each day over the 90 day period. The T(x) was then subtracted from the observed T and the standard deviation of these values across all points and dates was calculated to find the uniform temperature standard deviation.

The above process was then repeated but T(x) was set equal to Tavg + Tsd*map(x) to produce the standard deviation of the temperature produced by the high resolution model. It was found that the high resolution model is indeed more accurate with an average standard deviation of 0.551, which is lower than the uniform temperature standard deviation of 0.590. Therefore, the high resolution model is more accurate, on average, than the

Analysis Method cont.

To further validate the high resolution temperature prediction model for different locations across New York City, data were downloaded from the New York City MetNet database that are within the geographic area covered by the model. These data were collected from fixed sensors recording temperature and other parameters. However, once these data were processed using the same procedure as before, it was found that these data produced very high standard deviation values such as 2. Upon inspection of the data, it was found that the temperature data was rounded to the nearest degree thus further decreasing its accuracy. Based on this discovery and the fact that the MetNet data had standard deviations significantly greater then the … Heat Island Pr… York, these val…

Figure 3

$$SD = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

Results and Discussion

It was found that with a fixed correction factor of 2.0, the high resolution temperature prediction model produces the most accurate results with respect to the actual observed temperature. Scatterplots in degrees Celsius with a best fit linear trend line were made of observed T vs. uniform T, observed T vs. predicted T, and observed T vs predicted T with the correction factor of 2.0. The trend lines had a slope of about 1, indicating that the predicted and observed temperature track each other. However the most accurate model prediction, with a correction factor of 2.0, had a linear trend line with a slope of about 1 but with a y-intercept of 1.436. The y-intercept of 1.436 for the best fit line indicates that the entire model is off by +1.436 degrees Celsius for locations ~3.5m above ground; this is likely due to the uniform temperature prediction that the model relies on. Therefore, to correct the model for locations ~3.5m above ground, the uniform prediction will be decreased by about 1.5 degrees Celsius to correct the model.

Acknowledgements

This research was funded by NOAA CREST (NOAA CREST– Cooperative Agreement No: NA11SEC4810004) and supported by The Pinkerton Foundation. I would also like to thank my mentors Dr. Emiko Morimoto, Dr. Brian Vant-Hull, Dr. Shakila Merchant, and Dr. Reza Khanbilvardi for assisting me with my research.

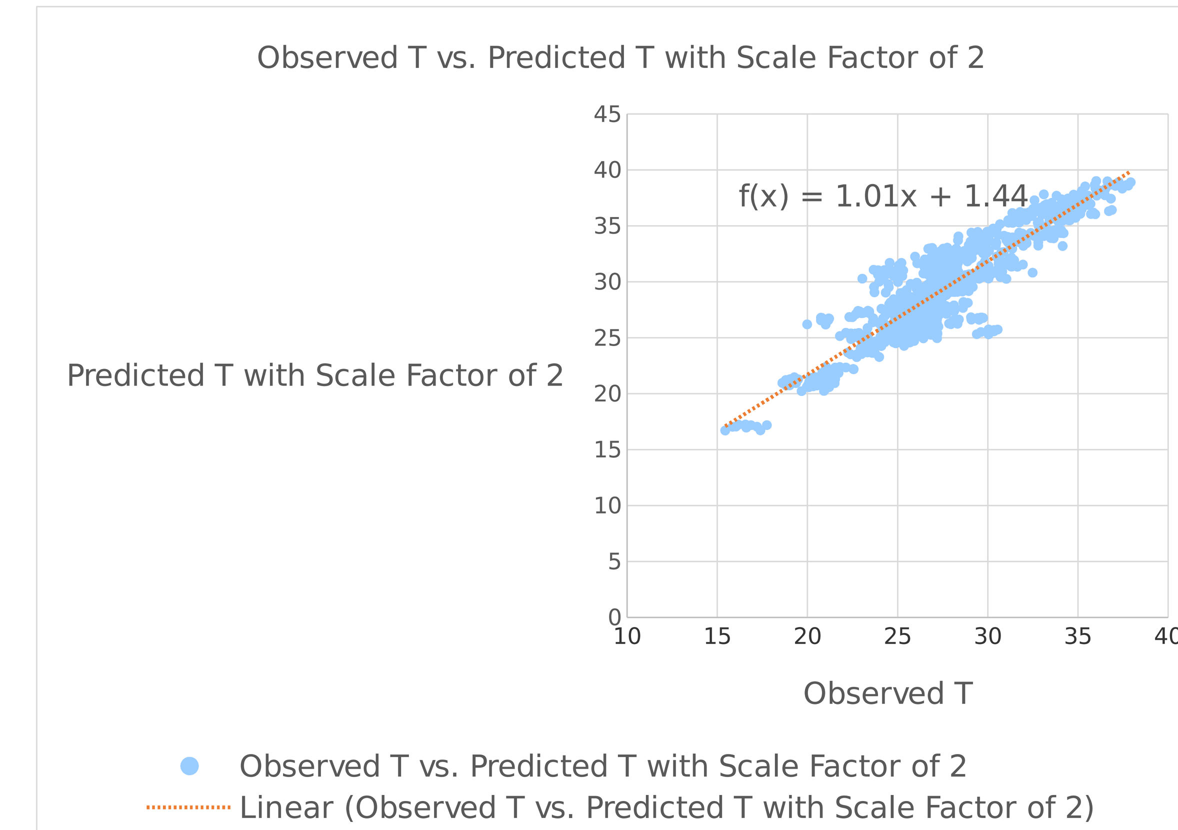

Figure 4

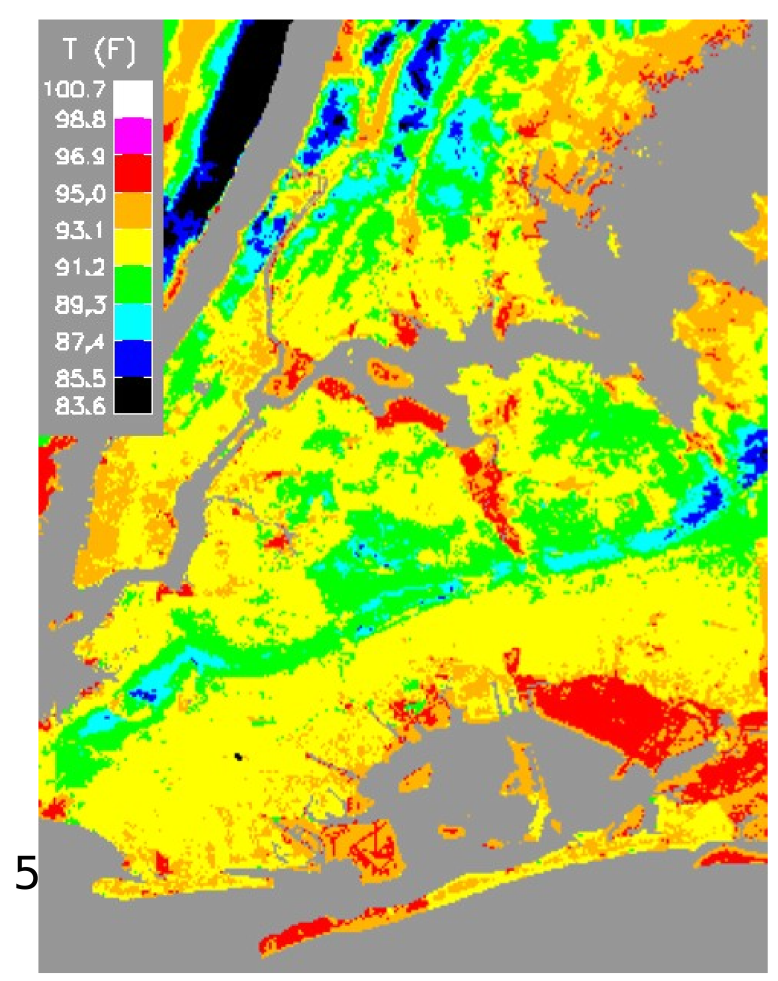

Figure 5

Sample Model Temperature Prediction Output

Conclusions and Future Work

One of the goals of this project is to develop a technique to produce high resolution temperature models in developed areas, with a particular focus on cities in the Northeastern United States. In the near future, field agents will start collecting weather data in Washington, DC to begin work on a localized high resolution weather prediction model. Based on the research conducted to validate the model for New York City it was found that for locations ~3.5m above ground, the model must be corrected by ~1.5 degrees Celsius.

References

1. The City College of New York, n.d., "Data Set Creation", Fine Scale Mapping of Manhattan's Urban Heat Island, http://glasslab.engr.ccny.cuny.edu/u/brianvh/UHI/dataset_creation.html, 28 July 2015
2. n.a., 4 February 2008, "The Law of Large Numbers As It Affects Kyle and Joseph", Team Green: Psyc261 – Statistics for Psychology, https://psyc261green.wordpress.com/2008/02/04/the-law-of-large-numbers-as-it-affects-kyle-and-joseph/, 30 July 2015