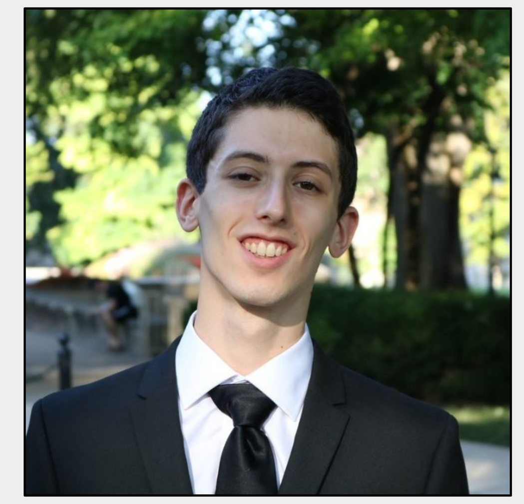
Heat Wave Identification Using an Operational Weather Model and Analog Ensemble

Alon Sidel[1,2], Weiming Hu[2], Martina Calovi[2], Guido Cervone[2]

[1] The Pennsylvania State University, College of Earth and Mineral Sciences, Department of Meteorology and Atmospheric Science
[2] The Pennsylvania State University, Institute for Computational and Data Sciences, Geoinformatics and Earth Observation Laboratory

Alon Sidel - ays10@psu.edu



Background

Heat waves, a prolonged period of abnormally hot weather, are the deadliest natural disaster in the United States[1]. Despite the phenomenon's ubiquity, there is no scientific consensus on how a heat wave is defined. This is primarily due to the subjectivity of a heat wave. No single meteorological variable represents the interaction between the human body and its thermal environment, what constitutes as abnormally hot weather varies by both climate and urban density, and there are many lengths of time where heat exposure can be considered dangerous[2]. Currently, heat waves are almost exclusively predicted based on the expectations of low-frequency variability patterns[3]. There has yet to be an accurate prediction of heat waves from an operational weather model. This project uses the Analog Ensemble (AnEn) technique, which generates probability information from a deterministic predictive model, to more accurately identify heat waves across the continental United States.

Generating AnEn

Generating predictions using AnEn requires both forecast (NAM) and observation (ASOS) data over a specified area. Six cities were chosen across the contiguous United States for their uniform spatial distribution. From the NAM, historical model analysis and forecasts were collected for the time period between January 1, 2009 and July 31, 2018 as training data for AnEn. For verification, data from 23 ASOS stations of surface temperature, relative humidity, hourly cumulative precipitation, cloud cover at different levels, and wind vectors were collected at hourly intervals from October 1, 2008, to September 1, 2018.

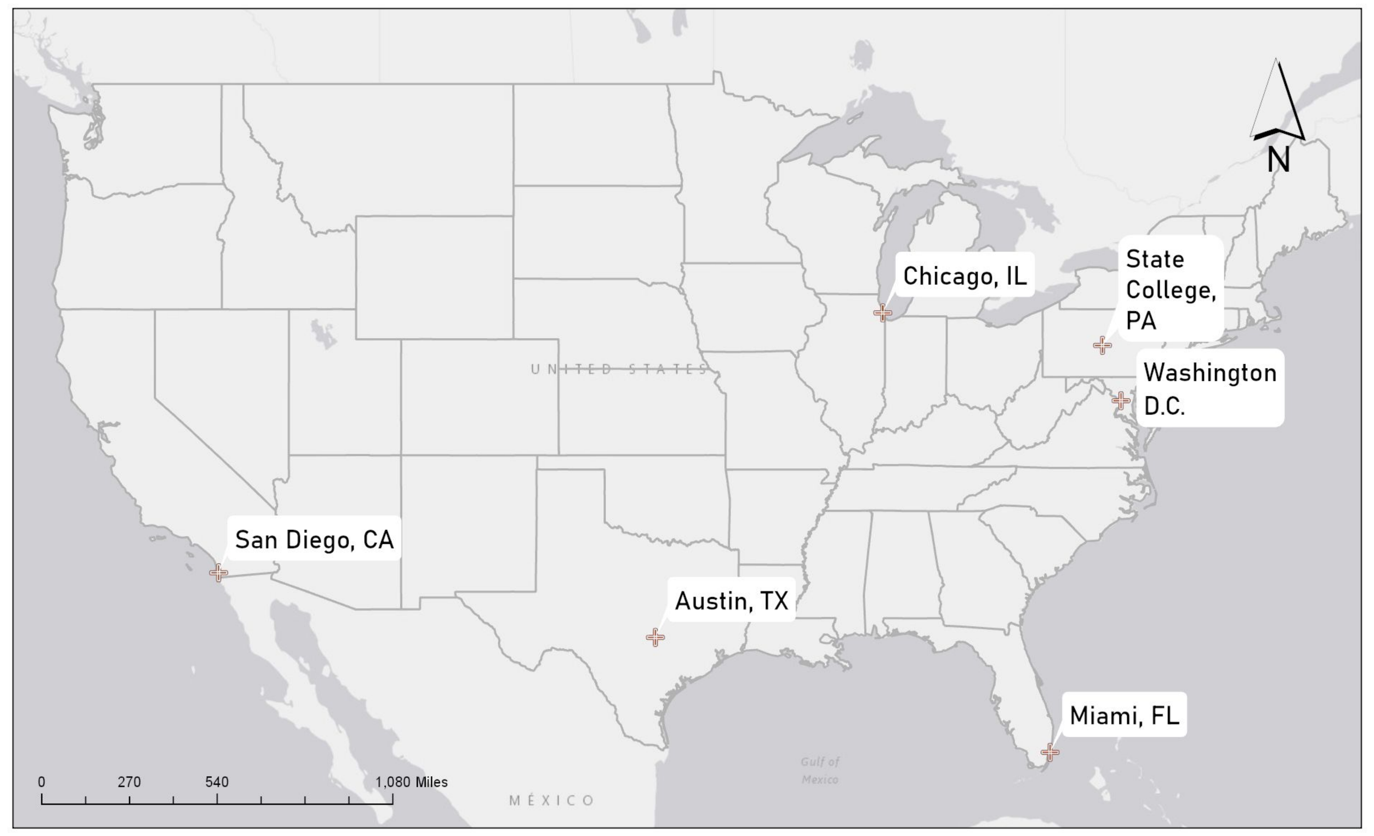


Figure 1: Geographic location map for the six sampled cities distributed over the contiguous United States. Each city contains multiple ASOS stations that belong to a corresponding NAM grid point.

AnEn selects past forecasts similar to a current model prediction for a forecast window and grid point, known as analogs. The corresponding observations to the analog forecasts are then used to construct ensemble members. The assumption of AnEn is that, given similar weather regimes, numerical weather models tend to make similar mistakes, which can then be quantified using the historical model errors. Since AnEn uses these observations as ensemble members, the errors are corrected in AnEn forecasts. Both the variables used to construct ensemble members and the ensembles themselves can be weighted.

Heat Index

Heat index is one of the most widely-used methods of assessing human discomfort in the outdoors in the United States. Its two independent variables are dry-bulb temperature and relative humidity, with variables such as vapor pressure, core body temperature, and wind speed being constants approximated for average conditions[4]. Due to heat index's use as a measure of comfort, heat index is defined to have meaningful value only at temperatures greater or equal to 27°C and relative humidity greater or equal to 40%.

Result

AnEn consistently outperforms the NAM for almost all of the 84 hours that constitute the Forecast Lead Time (FLT), with a weakening reliability farther out in the forecast. Note that the NAM has a tendency to underpredict extreme heat, potentially leading health officials to underestimate the severity of a heat wave.

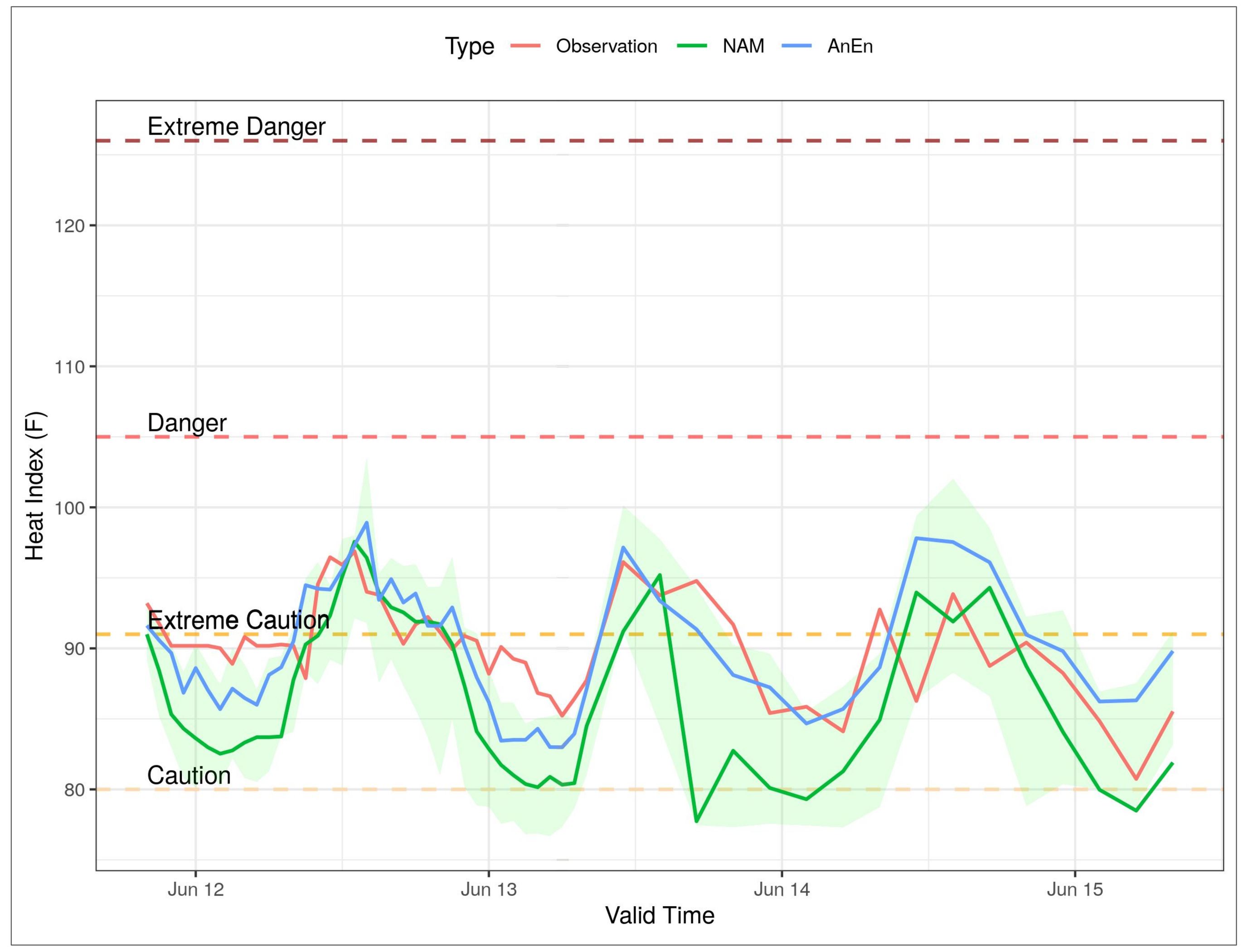


Figure 2: A sample heat index forecast outputted by AnEn in blue. The green area represents the range of ensemble members' individual forecasts. NAM (green) and observations (orange) are included for comparison. Dotted lines represent heat stroke risk[5]. Data from Miami, Florida 20:00 EDT on June 11th, 2017.

As mentioned before, AnEn can be optimized by determining optimal weights for each predictor in the generation of the heat index analog. Else, an assumption is being made that each predictor affects heat index equally. A Monte Carlo algorithm was created and run 10,000 times to produce the optimal weights. For each run, a randomly-generated set of weights was inputted in AnEn. Its forecasted outputs were then extracted and compared to the previous iteration of the algorithm. If the mean difference between the forecast and observed values were smaller than the previous iteration, the weights of that iteration would be kept. Else, the weights were discarded.

Verification

AnEn can also be optimized by weighting ensemble members. The most common weighting strategy is to average all the forecasts together at each time. This is effective for variables with a natural Gaussian distribution, but extreme heat values associated with heat index do not have a normal distribution. Thus, only temperatures below 80°F are considered valid for this particular optimization. Alternatively, the ensemble can be interpreted through its third quartile (75%). Figure 3a is optimized with the third quartile, while Figure 3b is optimized with the 2nd quartile (mean).

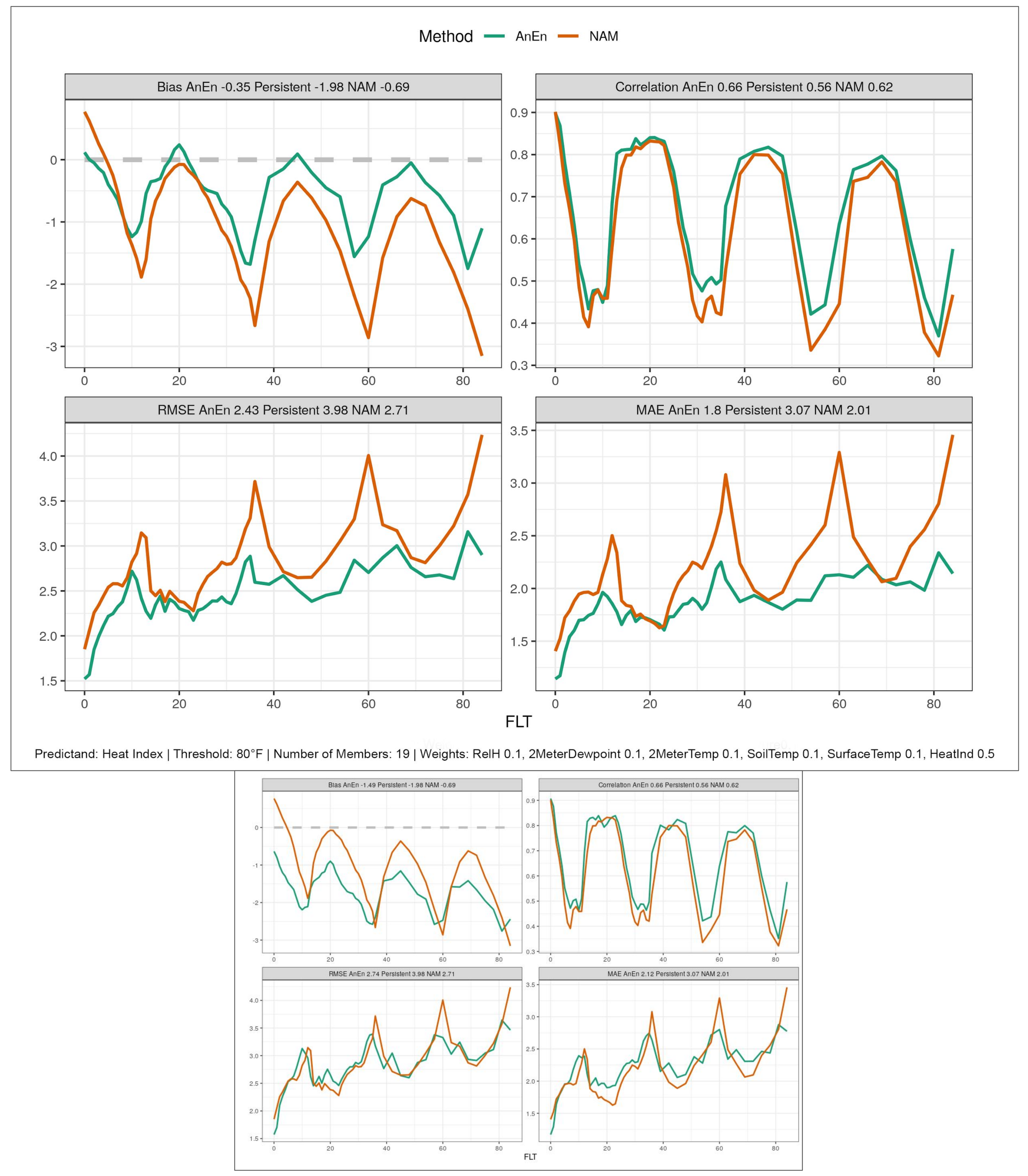


Figure 3: A comparison between the verification results of AnEn using quantiling (top, 3a) versus mean (bottom, 3b) for a single station. Results for each metric are displayed for AnEn, NAM, and the persistent forecast for 84 hours of FLTs. The AnEn forecast is in green with the NAM forecast in orange.

Verification is measured by two scalar accuracy measures, Mean Absolute Error (MAE) and Mean Squared Error (MSE). Additional metrics are bias and correlation, derived from MSE. In Figure 3a, both RMSE and MAE are lower for AnEn when quantiled suggesting a more certain forecast, but the significant improvement is the bias, which improves by a factor of four. Compared to the NAM, AnEn has a smaller bias, root-mean squared error, and mean absolute error, as well as higher correlation. The persistence forecast, not plotted, assumes future values will continue the trend of past values. AnEn consistently produces much better results than the persistent forecast, suggesting it is not strictly trend-based.

Conclusion

Heat waves are predicted to intensify as the global climate becomes warmer[6]. AnEn, a deterministic ensemble comprised of analogs, was generated from computed values of heat index across 23 American weather stations. Its results were compared to the NAM using four scalar accuracy measures, and was found to surpass the NAM by a significant margin when processed with predictor weighting and quantiling.

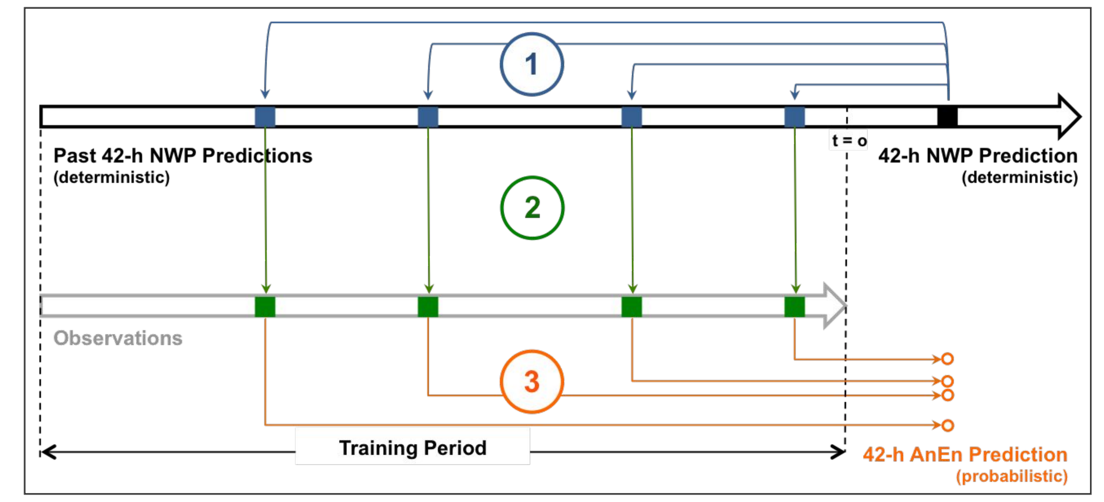


Figure 4: A schematic representation of the process for finding four members of AnEn at one forecast lead time[7]. This poses benefits and challenges to note.

Currently, the AnEn technique has only been applied as a deterministic forecast, where each forecast lead time is represented by a single value. Further analysis must be undertaken to assess the effectiveness of a probabilistic ensemble, where each of the FLTs represent multiple forecasts with an individually-assigned likelihood, in the prediction of heat waves. In addition to quantiling, a bias correction could potentially be applied to AnEn with the effect of decreasing forecast bias.

References

1. Wilhelmi, O. V., & Hayden, M. H. (2010). Connecting People and Place: A New Framework for Reducing Urban Vulnerability to Extreme Heat. *Environmental Research Letters, 5*(1), 014021. doi: 10.1088/1748-9326/5/1/014021
2. Guo, Y., *et al.* (2017). Heat Wave and Mortality: A Multicountry, Multicommunity Study. *Environmental Health Perspectives, 125*(8), 087006. doi: 10.1289/ehp1026
3. Teng, H., *et al.* (2013). Probability of US Heat Waves Affected by a Subseasonal Planetary Wave Pattern. *Nature Geoscience, 6*(12), 1056–1061. doi: 10.1038/ngeo1988
4. Steadman, R. G. (1979). The Assessment of Sultriness. Part I: A Temperature-Humidity Index Based on Human Physiology and Clothing Science. *Journal of Applied Meteorology, 18*(7), 861–873. doi: 10.1175/1520-0450(1979)018<0861:taospi>2.0.co;2
5. The National Weather Service. (2019, December 15). What Is the Heat Index? Retrieved from https://www.weather.gov/ama/heatindex.
6. Argüeso, D., Evans, J. P., Pitman, A. J., & Luca, A. D. (2015). Effects of City Expansion on Heat Stress under Climate Change Conditions. *Plos One, 10*(2). doi: 10.1371/journal.pone.0117066
7. Monache, L. D., Eckel, F. A., Rife, D. L., Nagarajan, B., & Searight, K. (2013). Probabilistic Weather Prediction with an Analog Ensemble. *Monthly Weather Review, 141*(10), 3498–3516. doi: 10.1175/mwr-d-12-00281.1

I would like to thank Dr. Guido Cervone for providing me with the opportunity to conduct research in his lab.