Severe thunderstorm outbreaks are not uncommon in Upstate New York, and are usually predicted with reasonable accuracy. However, within these outbreaks, determining the type of severe threat posed by an individual thunderstorm is more difficult. A series of multiple regression models were developed that determine the probability of severe wind and severe hail from an individual thunderstorm based on the values of 48 parameters at the time and location of the thunderstorms' initiation. The models can be used in real-time operational forecasting, and an analysis of the predictive power of the individual parameters can provide insight into how these parameters explain thunderstorm development

11 different multiple regression models were built to develop the probability determination. Eight were built to predict the severe threat; four for severe wind and four for severe hail. The datasets of these parallel sets of models differ on the basis of storm type; one model was built with all severe wind reports, one was built with severe wind reports from supercells, and likewise for clusters and squall lines, and then four models for severe hail were built in similar fashion. The remaining three models were built to determine storm type (supercell, cluster, or squall line). The raw outputs of these models were used to construct a probability distribution which was then used to determine a final output probability.

By using input data from the time of thunderstorm formation, the models aim at being both useful operationally and in thunderstorm development research. While the models that produce the probability of type of severe threat take as input all 48 parameters that were collected, the 11 models were also re-created using the 24 parameters that demonstrate the strongest correlation with the severe threat. Comparing the similarities and differences in these strongly correlated parameters across the models can lead to avenues of inquiry into the development of severe thunderstorms.

Data to create the models was taken from severe thunderstorm outbreaks that occurred across Upstate New York from 2013 and 2014. Severe storm reports and archived radar data were used to identify severe storms. In total, 513 storm reports were used, and 70% of these (359) were randomly selected to build the model while the remaining 30% (154) were used as data to test the models against. An addition 1000-1500 storm reports from severe thunderstorm outbreaks in 2013-2017 plan to be added to the overall dataset and the process of building and testing the models will be repeated to increase accuracy and confidence. These additional data points will enable us to split the dataset geographically. Instead of considering all severe thunderstorms across all of Upstate New York, we will be able to build and analyze models that include only thunderstorms that occurred only in the Adirondack Mountains and only in the Hudson River Valley, for example. This geographic split, and the differences and similarities among the parameters with the strongest correlations between regions, will allow us to identify which factors are most important in the local terrain. In addition, future null data points (thunderstorms that did not produce severe wind or hail) are planned to be added to the models in order to not be able to predict how a thunderstorm might be severe, but the likelihood that it actually does end up producing severe threats.

The models, once fully built and tested, can have significant operational usefulness in their ability to determine the type of severe threat a thunderstorm is most likely to pose and the probability of the severe threat occurring. The details of the models themselves can also open avenues of future research into the thunderstorm formation process, and can be compared to similar research done for thunderstorms in other regions of the US. As it currently stands, a process has been developed to use the values of 48 environmental parameters at the time of thunderstorm formation to predict the type of severe threat posed by that storm. The predictive power of this process will be determined, and topics of potential future study within thunderstorm formation will be analyzed. In the spring of 2019, 1000-1500 additional data points will be added to the models, and the process will be re-assessed.