UNIVERSITY OF MARYLAND

# Investigating Current and Future Precipitation Frequency Estimates for the State of Maryland: Challenges of Applying Machine Learning for Temporal Downscaling of Climate Model Projections


AZIN AL KAJBAF, MICHELLE BENSI, KAYE BRUBAKER

UNIVERSITY OF MARYLAND

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING



Ellicott City, Maryland, after flash flooding in late May 2018. Photo by Howard County government

# Introduction

- Climate change has altered the meteorological and hydrological characteristics of precipitation events in recent decades.
- Extreme rainfall events appear to be occurring more frequently.
- The contiguous United States has experienced an increase in mean average precipitation in each decade (1951-2013) [1].
- Increasing trends in extreme precipitation events are more pronounced in the Northeast of the United States [2-4].
- Associated with this trend, Maryland communities have experienced multiple flash flood events (e.g., Ellicott City flash floods in 2016 and 2018).
- These impacts are expected to worsen due to climate change.

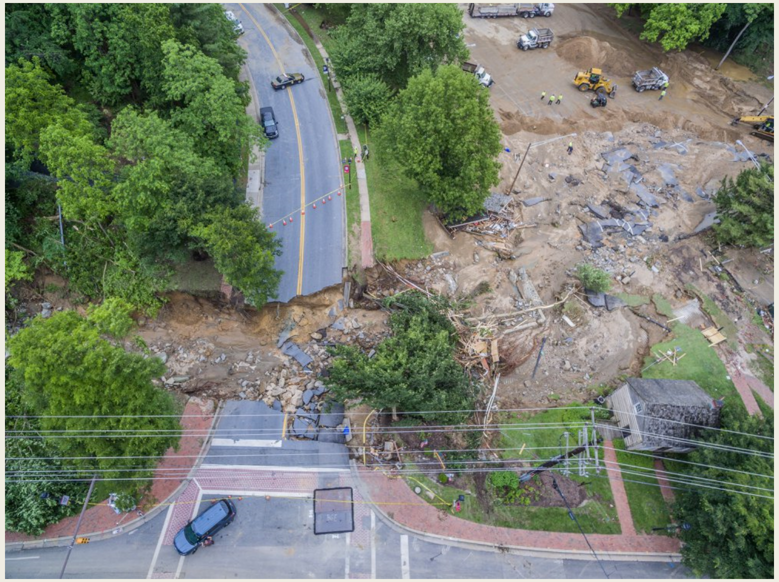

**Figure 1.** Cars overturned by flash flooding in Ellicott City, Maryland. Photo: Howard County government.

**Figure 2.** Ellicott Mills Drive at Main Street in Ellicott City on May 28. Photo: Howard County government.

# Objective

- This study evaluates multiple Machine Learning (ML) algorithms that are used for preparing precipitation data for analyzing current and future climate Intensity/Depth Duration Frequency (IDF/DDF) for the state of Maryland using the North American Regional Climate Change Assessment Program (NARCCAP) model output at 50-km spacing [5].

- The high-resolution projections of precipitation generated by NARCCAP, provided in 3-hour intervals, must be temporally disaggregated to obtain IDF/DDF curves for shorter duration rainfall events.

- This study implements multiple ML algorithms, including Artificial Neural Network (ANN), Boosted Trees (BT), and Support Vector Regression (SVR), to disaggregate 3-hour precipitation to durations of 2 hours, 1 hour, 30 minutes, and 15 minutes.

- The ML models are trained using observational data, then applied to NARCCAP output.

- Response functions are presented for further investigation of the behavior of the ML models under varying inputs.

# Literature Review

A. James Clark School of Engineering

- A brief literature review is provided on studies that have used ML methods for temporal downscaling in **Table 1.**

- Information regarding location, ML method, predictors and temporal conversion and summary of the study is presented.

**Table 1.** Literature review

| Authors | Location | Data Type | ML Method | Predictors | Temporal Conversion |
|---|---|---|---|---|---|
| Burian et al. (2001) [6] | Alabama, USA | 3 Rain gauge stations | ANN | Three sequential hours of rainfall amounts in a long-term hourly rainfall record. | Downscaling hourly precipitation to 15 minutes |
| Mirhosseini et al. (2014) [7] | Alabama, USA | 1- Historical rainfall data of NOAA (34 15-mins rain gauges station)<br><br>2- Simulated historical precipitation from NARCCAP | ANN, Stochastic method | 3-hour precipitation (P3), daily precipitation, monthly precipitation, and daily maximum and minimum temperatures | Downscaling 3h precipitation to 15, 30, 45, 60, 120 minutes |
| Alam and Elshorbagy (2015) [8] | Saskatoon, Canada | 1- Daily, hourly and 5-minutes observations<br>2- Daily precipitation of climate models of CanESM2, HadGEM2-ES, IPSL, CNRM, CSIRO, BCC, MRI and MIROC | K-nearest neighbors (K-NN) | An optimal window size was chosen on both sides of a disaggregation period | Downscaling from daily to hourly and from hourly to sub-hourly |
| Sachindra et al. (2018) [9] | Victoria, Australia | 1- Monthly observations of precipitation (48 stations)<br><br>2- NCEP-NCAR Reanalysis Data | Genetic Programming, ANN, SVR, Relevance Vector Machine | Air temperature at surface and geopotential heights relative and specific humidity at surface, zonal and meridional wind speeds, sea level pressure, pressure at surface and precipitable water content | Downscaling reanalysis data to monthly precipitation |

# Study Area and Data

- The rectangular area in **Figure 3** indicates the study area.
- Hourly and 15-minute stations in the states of MD, VA, WV, PA, DE and NJ are used as observations for training the ML models and are illustrated in **Figure 3**.
- The observations are collected from Climate Data Online database of the NOAA National Center for Environmental Information (NCEI).
- Hourly and 15-minute stations are used when daily summaries information is available at the same station.

**Table 2.** Length of record in each type of station

| Type of station | Duration of data |
|---|---|
| Hourly | Hourly: 1950-2014<br>Daily summaries: 1950-2014 |
| 15-Minute | Hourly: 1970-2014<br>Daily summaries: 1970-2014 |

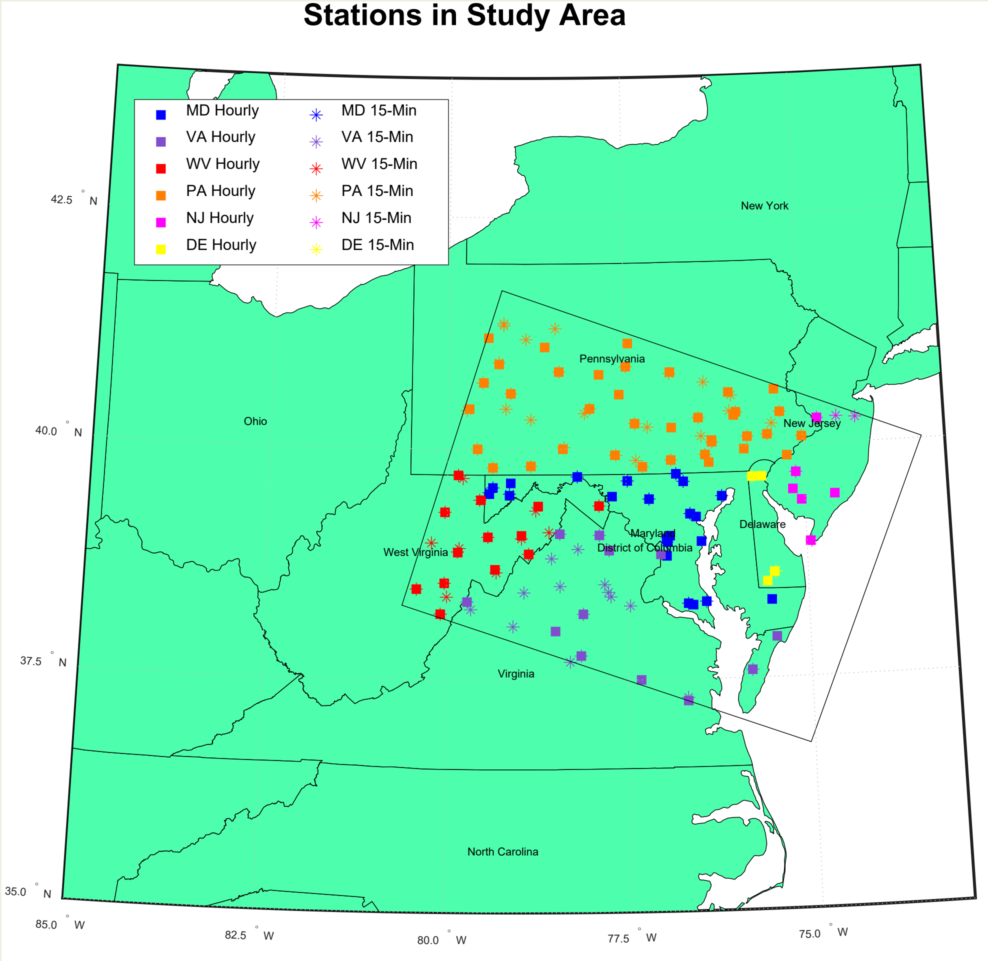


**Figure 3.** Location of hourly and 15-minute stations in the study area.

# Challenges Associated with Preparing the Datasets

A. JAMES CLARK SCHOOL OF ENGINEERING

- Since daily summaries information (e.g., temperature, daily precipitation) are used in conjunction with hourly and 15-minute precipitation data for training the ML models, stations where daily summaries information are not available, cannot be used in training.

- 3-hour precipitation is mapped to durations of 2 hours, 1 hour, 30 minute, and 15 minute. Since hourly durations of 2 hours and 30 minutes are not available, they are generated from aggregation of 1 hour and 15-minute duration datasets for consecutive events.

- In hourly and 15-minute stations, no observation is recorded when precipitation is zero. To efficiently use the available data for generating longer durations, 45-minute and hourly stations are zero padded for an extra 45 minute and 2 hours, respectively.

- At some locations, reported hourly or 15-minute precipitation exceeds associated reported daily precipitation. Since the exact time of recording for daily precipitation is not clear, observations where daily precipitation is smaller than hourly or n-minute precipitation, are eliminated from the datasets for training.

# Input Parameters for Each Duration

- The higher duration precipitation that is used as predictor has the highest importance and a significant difference with other parameters in predicting target durations and hence is not shown in the plots.

- Other parameters including latitude, longitude and elevation of stations were used for prediction. Using geographic information improved the performance of the ML models, however, when applied to NARCCAP dataset, the results were inconsistent. Therefore, they are eliminated from the final set of input parameters.

**Table 3.** Input parameters for training ML models in each duration

| Input parameter | Target precipitation |
|---|---|
| 3h pr, Daily pr, Max daily temp, Min daily temp, Max monthly pr, Max yearly pr | 2 hour |
| 2h pr, Daily pr, Max daily temp, Min daily temp, Max monthly pr, Max yearly pr | 1 hour |
| 1h pr, Daily pr, Max daily temp, Min daily temp, Max monthly pr, Max yearly pr | 30 minutes |
| 30 min pr, Daily pr, Max daily temp, Min daily temp, Max monthly pr, Max yearly pr | 15 minutes |

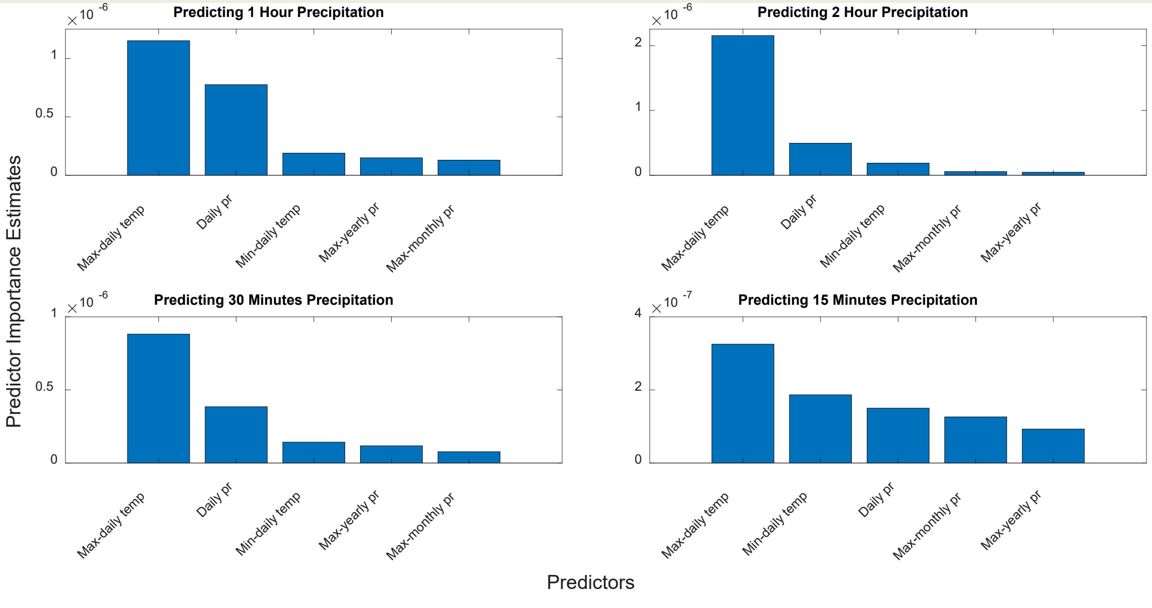


**Figure 4.** Estimates of predictor importance in predicting different durations.

# ML Models’ Performances

A. JAMES CLARK
SCHOOL OF ENGINEERING

- ML models of ANN, BT and SVR are developed, and their performance is evaluated before applying to NARCCAP points.

- Models are trained using 70% of randomly selected data (k-fold=10) and tested on remaining 30% percent. The results are shown as vertical bars where top and bottom bars show maximum and minimum, and the center symbol shows the average value of the performance measures in 10 folds.

- ANN and BT constantly outperform SVR. The variation in results of BT is lower than that in ANN and SVR.

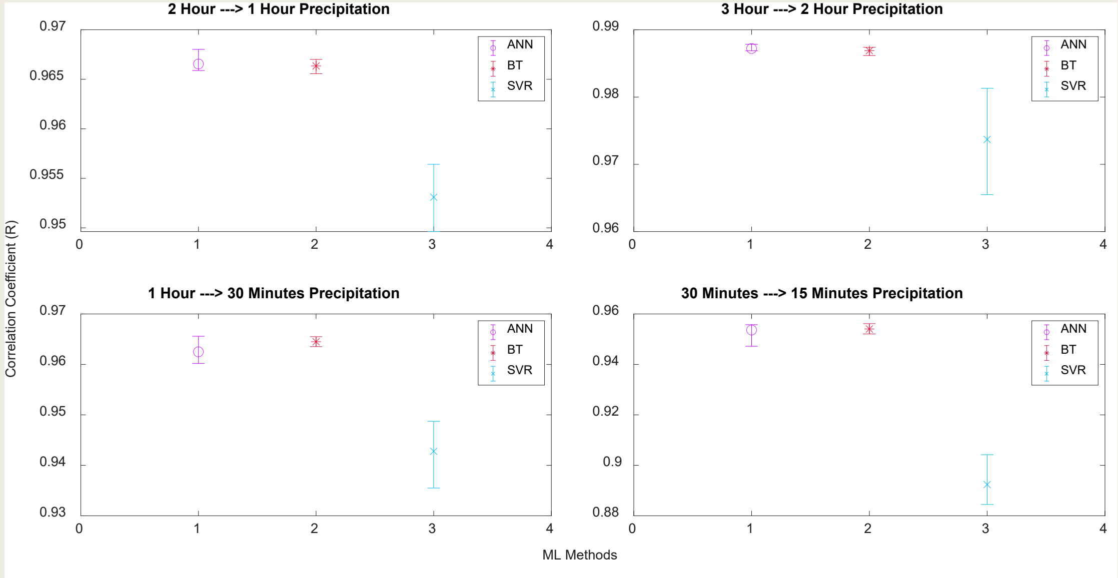


**Figure 5.** Correlation coefficient of ML models in predicting different durations.

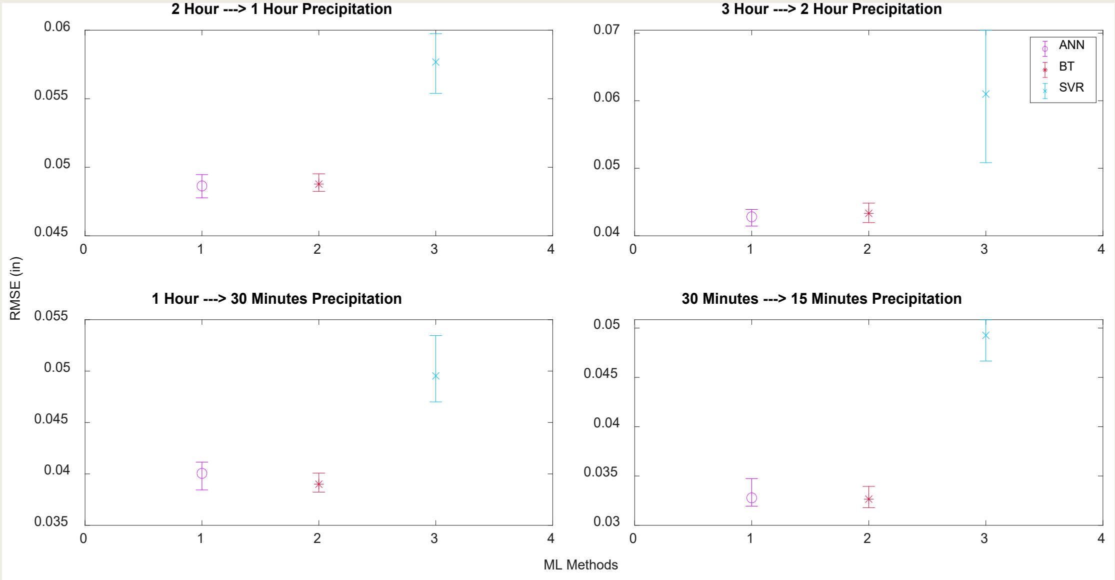


**Figure 6.** RMSE of ML models in predicting different durations.

# Response Functions



- For developing response functions, all variables are locked at a specific value (average of the range considered for each variable) except for two variables which are shown on x and y axes. The colors on the plot show the response of the models with respect to each set of variables.

- For both models the first three figures from the left shows the response function when one of the parameters is a higher duration precipitation which is the leading contributor to the precipitation value.

- Response functions also depict the different mechanisms of the ML models in predicting target precipitation.

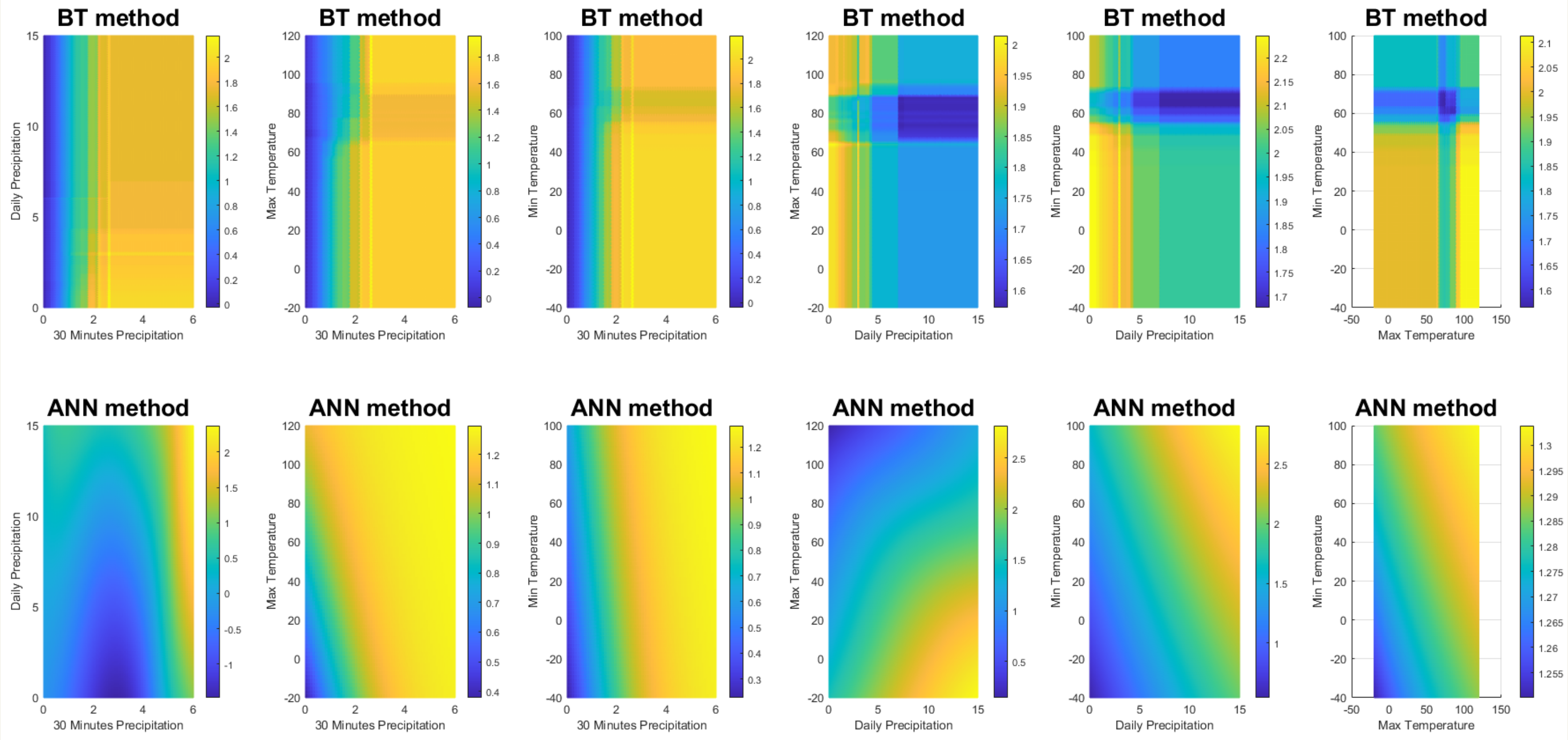


**Figure 7.** Estimates of predictor importance in predicting 15-minute precipitation.

# NARCCAP Points in EPC2-gfdl Climate Scenario

A. JAMES CLARK SCHOOL OF ENGINEERING

- The NARCCAP regional model points (centers of the grid cells) that are located in Maryland are extracted. The unit of precipitation in the regional models is $kg\ m^{-2}s^{-1}$ and is converted to inches.

- The values of precipitation in regional models are provided over grid cells at 50-km spacing, or 2500 $km^2$ in area. They are converted to point values using point/area adjustment following LeClerk and Schaake [10].

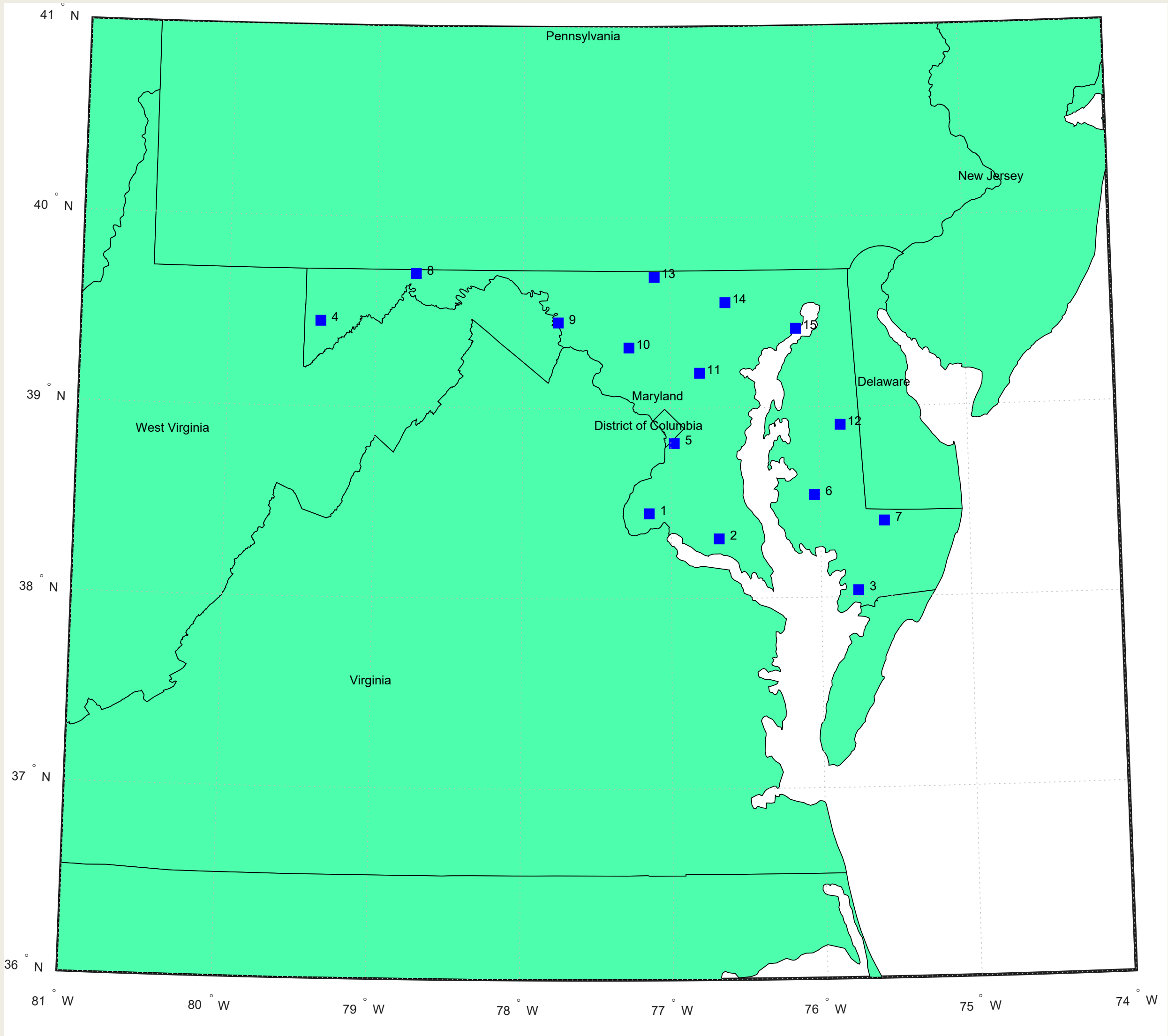


**Figure 8.** The location of NARCCAP points in state of MD in EPC2-gfdl scenario

# Applying the ML models to NARCCAP Points

- The trained ML models of ANN and BT are applied to 3-hour precipitation information of NARCCAP points in Maryland.

- Although the results of both models seems to be close, the BT model shows a better performance in predicting smaller values of precipitation.

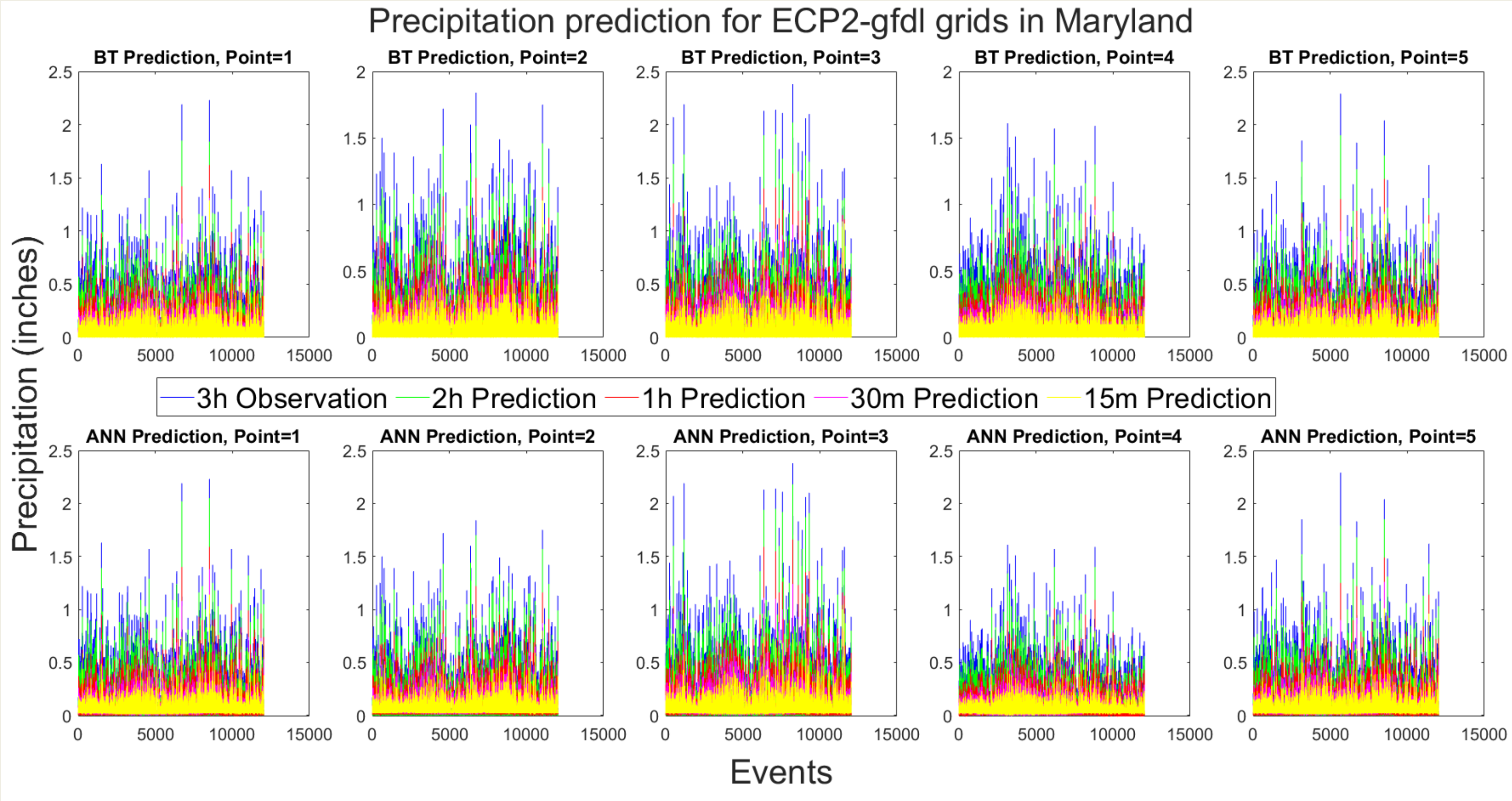


**Figure 9.** Results of applying BT and ANN models to NARCCAP points for different durations.

# Conclusions

A. JAMES CLARK SCHOOL OF ENGINEERING

- Multiple ML methods (e.g., ANN, BT, SVR) are implemented for temporal disaggregation of high-resolution projections of precipitation generated by NARCCAP, provided in 3-hour intervals to durations of 2 hours, 1 hour, 30 minutes, and 15 minutes.

- Challenges that are associated with preparing the datasets for training the ML methods include:
  - unavailability of daily summaries information at some hourly and 15-minute stations (these stations are not considered in the analysis).
  - Inconsistency of precipitation data recorded by daily summaries sets and hourly and 15-minute stations (the observations with this inconsistency are eliminated from dataset).
  - Hourly and 15-minute stations does not record data when precipitation is zero which can be mistaken with periods of missing data (zero padding is applied).

- In selection of set of appropriate input parameters for training ML models, aspects beyond performance metrics (e.g., R, RMSE) that are used for evaluation of such models must be considered.

- Even though comparison of R and RMSE shows that ANN has similar performance (sometimes ANN is slightly better) to BT, when applied to NARCCAP points, BT shows to be better in predicting precipitation values close to zero.

- Response function provide valuable insight regarding behavior of ML models in predicting target response in wide range of input variables.

A. JAMES CLARK SCHOOL OF ENGINEERING

# References:


[1] L. Mearns et al., “North American Regional Climate Change Assessment Program dataset.” UCAR/NCAR, p. approximately 40 TB, 2007, doi: 10.5065/D6RN35ST.

[2] S.-Y. Wu, “Changing characteristics of precipitation for the contiguous United States,” Climatic Change, vol. 132, no. 4, pp. 677–692, 2015.

[3] D. J. Leathers, S. E. Brasher, K. R. Brinson, C. Hughes, and S. Weiskopf, “A comparison of extreme precipitation event frequency and magnitude using a high-resolution rain gage network and NOAA Atlas 14 across Delaware,” International Journal of Climatology, 2019.

[4] L. Agel, M. Barlow, J.-H. Qian, F. Colby, E. Douglas, and T. Eichler, “Climatology of daily precipitation and extreme precipitation events in the northeast United States,” Journal of Hydrometeorology, vol. 16, no. 6, pp. 2537–2557, 2015.

[5] K. E. Kunkel et al., “Monitoring and understanding trends in extreme storms: State of knowledge,” Bulletin of the American Meteorological Society, vol. 94, no. 4, pp. 499–514, 2013.

[6] Burian, S.J., Durrans, S.R., Nix, S.J. and Pitt, R.E., 2001. Training artificial neural networks to perform rainfall disaggregation. Journal of Hydrologic Engineering, 6(1), pp.43-51.

[7] Mirhosseini, G., Srivastava, P. and Fang, X., 2014. Developing rainfall intensity-duration-frequency curves for Alabama under future climate scenarios using artificial neural networks. *Journal of Hydrologic Engineering*, *19*(11), p.04014022.

[8] Alam, M.S. and Elshorbagy, A., 2015. Quantification of the climate change-induced variations in Intensity–Duration–Frequency curves in the Canadian Prairies. *Journal of Hydrology*, *527*, pp.990-1005.

[9] Sachindra, D.A., Ahmed, K., Rashid, M.M., Shahid, S. and Perera, B.J.C., 2018. Statistical downscaling of precipitation using machine learning techniques. *Atmospheric research*, *212*, pp.240-258.

[10] Leclerc, G. and Schaake, J.C., 1972. Derivation of hydrologic frequency curves. Report 142. Mass. Inst. of Technol., Cambridge, MA, USA.