

SEPTEMBER 2020

# Project Hyperion - Narrative Case Study Report: *South Florida*

Kripa Jagannathan ([kajagannathan@lbl.gov](mailto:kajagannathan@lbl.gov)) and Andrew Jones ([adjones@lbl.gov](mailto:adjones@lbl.gov))

Lawrence Berkeley Laboratory



## Contributors

- Paul Ullrich, University of California, Davis (Project Team Leader)
- Bruce Riordan, Climate Readiness Institute (Engagement Facilitator)
- Abhishekh Srivastava & Richard Grotjahn, University of California, Davis
- Carolina Maran, South Florida Water Management District
- Colin Zarzycki, Pennsylvania State University
- Dana Veron and Sara Rauscher, University of Delaware
- Hui Wang, Tampa Bay Water
- Jayantha Obeysekera, Florida International University
- Jennifer Jurado, Broward County
- Kevin Reed, Stony Brook University
- Simon Wang & Binod Pokharel, Utah State University
- Smitha Buddhavarapu, Lawrence Berkeley Laboratory

# Contents

Introduction .....	3
1. Co-production in Hyperion.....	4
2. Regional hydro-climatic context & challenges .....	5
3. Climate information needs for water management.....	6
3.1. Overview .....	6
3.2. List of decision-relevant metrics and their importance .....	6
4. Key scientific activities and results from Hyperion.....	10
4.1. Precipitation IDF curves .....	10
4.1.1. Background .....	10
4.1.2. Key Results .....	12
4.1.3. Discussion and Conclusions .....	14
4.2. Sea Breeze.....	15
4.2.1. Background and Methods .....	15
4.2.2. Key Results .....	17
4.2.3. Discussions and Conclusions .....	18
4.3. Tropical Cyclones.....	19
4.3.1. Background and Methods .....	19
4.3.2. Key Results .....	20
4.3.3. Discussions and Conclusions .....	22
Acknowledgements and way forward .....	23
References .....	33

## List of Tables

Table 1: Examples of decision-relevant metrics for each region. ....	7
Table 2: Average number of sea breezes detected per year over the 2005-2012 time period for observations (Obs) and for CAM-REF-VR28 (Model) output. ....	18
Table A1: List of metrics and summary of scientific activities pursued by Hyperion project.....	24

## List of Figures

Figure 1: Co-production process and timeline .....	4
Figure 2: NA CORDEX models' predictions precipitation intensity estimates for a 24-hour duration storm.....	12
Figure 3: 24-hr IDF estimates in the bias-corrected historical and future simulations. ....	13
Figure 4: Changes in 24-hr precipitation for 2, 5, 10, 25, 50 and 100 year return periods computed from pooled models. ....	14
Figure 5: Summertime daily probability of precipitation (PoP) for classic sea breeze occurrence .....	17
Figure 6: Percent of precipitation during summertime (JJA) days where a classic sea breeze is detected .....	18
Figure 7: Comparison of observed and CAM 5 simulated landfalling TC tracks at different resolutions. ....	21
Figure 8: Rx5day, TC-related Rx5day, and % of Rx5day events due to TCs for observations, CAM5 variable resolution and CAM5 conventional climate model resolutions.....	22
Figure A1: 24-hr IDF estimates in the bias-corrected historical and future simulations for 74 weather stations.....	28
Figure A2: Model rankings for different NA-CORDEX models and for different precipitation metrics based upon Taylor diagram. ....	29
Figure A3: Ranking of NA-CORDEX models based upon IVSS, for different precipitation metrics. ....	30
Figure A4: Overall rankings based on both Taylor diagram and IVSS for NA-CORDEX models, for various precipitation metrics .....	31
Figure A5: Overall rankings based on both Taylor diagram and IVSS for NA-CORDEX models for annual maximum precipitation .....	32

# Introduction

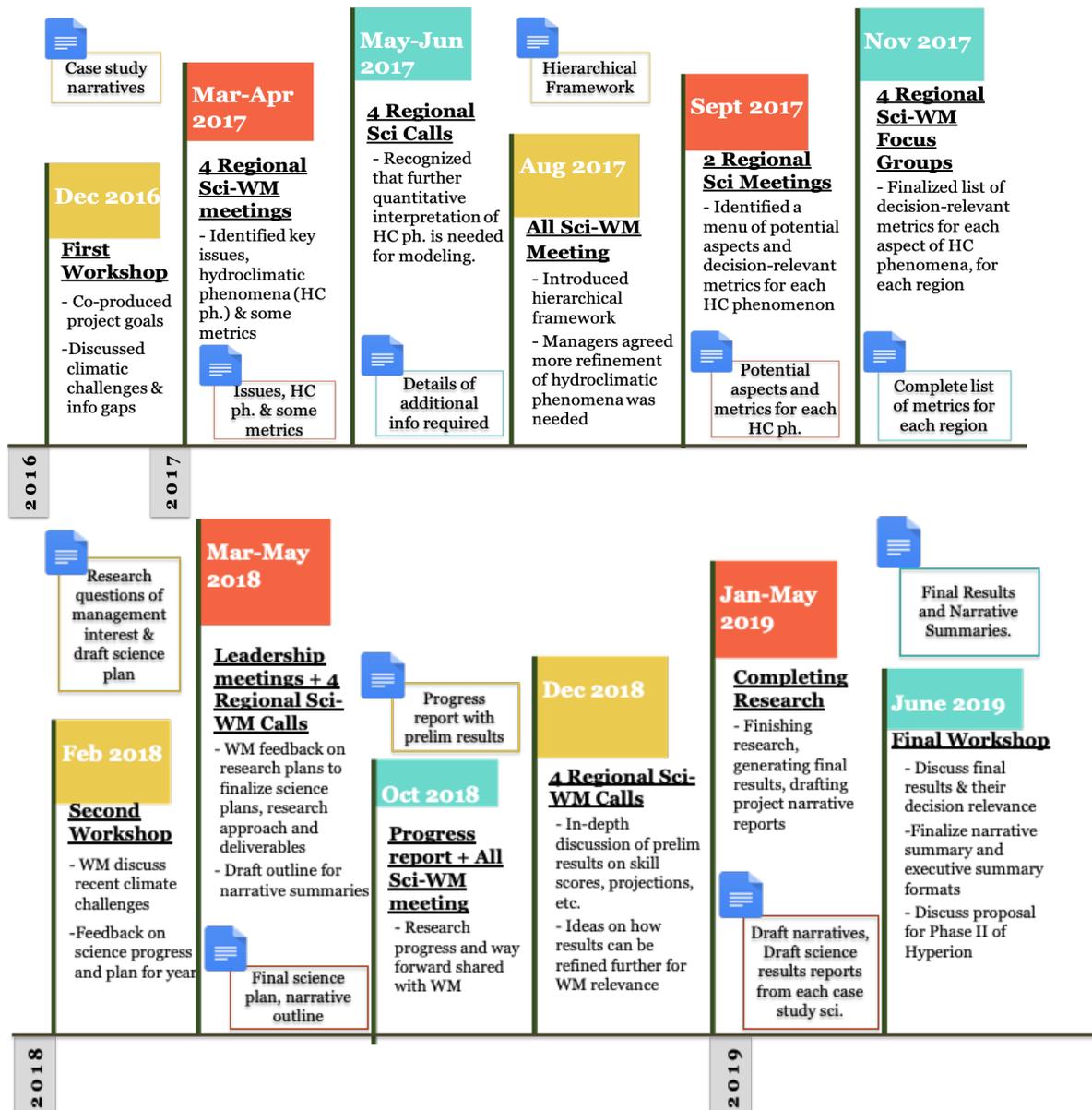
This narrative case study report is a synthesis of key discussions and preliminary scientific results for the South Florida region, undertaken as part of the Hyperion project (2016-19). Project Hyperion (now continuing as the HyperFACETS project) is a basic science project that aims to advance climate modelling by evaluating regional climate datasets for decision-relevant metrics. While there has been an explosive growth in the number of regional climate datasets available to users, there is limited understanding of the credibility and suitability of these datasets for use in different management decisions. Hyperion aims to address this need by developing comprehensive assessment capabilities to evaluate the credibility of regional climate datasets, understand the processes that contribute to model biases, and improve the ability of models to predict management relevant outcomes.

Since decision-relevance is a core motivation for the project, Hyperion is designed on the principles of co-production. The project brings together scientists from nine research institutions and managers from twelve water agencies in four watersheds: Sacramento/San Joaquin, Colorado Headwaters, South Florida, and Susquehanna. The project structure explicitly allows for both the groups to co-develop the science plan and research questions, in addition to co-producing the science itself. The scientists include atmospheric and earth system scientists as well as hydrologists. The water managers, depending on the agency, have functions including planning, operating and managing water quality, water supply, stormwater management, flood control, and water infrastructure design.

This narrative report provides an overview of the co-production process in Hyperion (Chapter 1), the regional hydro-climatic context and challenges (Chapter 2), broad climate information needs of water management agencies (Chapter 3), and short summaries of the key scientific activities undertaken for the region (Chapter 4). This information is based on the project's co-production engagements and preliminary scientific results. Some of these preliminary results may be updated or refined as they go through the peer-review process. While this report is based on the perspectives of water management agencies that were part of Hyperion, we hope that the insights and methodologies that were developed are broadly applicable to other agencies in the region as well.

# 1. Co-production in Hyperion

In Hyperion, as far as possible, the research questions, approaches and results were co-produced through regular structured and unstructured engagements between scientists and managers (Figure 1). Structured engagement methods included workshops, remote and in-person focus-group discussions, and quarterly project update calls. There were also continual less-structured, informal conversations over telephone calls and emails.



**Figure 1: Co-production process and timeline**

Summarizes key engagement activities along with important outcomes at each stage (depicted by the blue document icon). 'Sci' refers to Scientists, 'WM' refers to Water Manager and 'HC ph.' refers to Hydroclimatic Phenomena.

## 2. Regional hydro-climatic context & challenges

The South Florida region is generally considered the region south of the Orlando Florida area, and includes the heavily populated south-eastern coastal counties of Dade, Broward, and Palm Beach and their municipalities such as Miami, Fort Lauderdale, West Palm, Tampa Bay and Fort Myers metropolitan areas on the Gulf Coast. The region has a population of more than 9 million people and is rapidly growing. It is home to the Florida Everglades, one of the most diverse ecosystems in the world, and includes over 3 million acres of agriculture that has undergone centuries of transformation. The area also contains substantial high value coastal infrastructure. There are a few large restoration projects being undertaken to address growing urban and agricultural water needs that also seek environmental protection and restoration. Flood protection, water supply, sea-level rise, maintaining natural systems and water quality, are some of the key water management priorities of the region.

Overall, the key challenge for the region is balancing the quantity, quality, timing and distribution of water for an increasing urban population, restoration of natural systems especially around the Kissimmee River, and the Florida Everglades, and protecting water quality from both nonpoint and point sources. Sea level rise and flooding are critical climate related threats. High intensity rainfall events are also a related risk. With increasing temperatures, managing extreme heat events and maintaining water supply for all sectors are becoming important issues. Saltwater intrusion into surface and groundwater systems are also a matter of concern.

Climate and hydrological data are used in some infrastructure plans and operations management such as in asset management and 50-year capital improvement plans. Regional and integrated hydrologic and hydraulic modelling activities are being undertaken, that include detailed descriptions of surface water, groundwater, and the intricacies of the infrastructure used to control their interaction. As most of the southeast region is only slightly elevated above sea-level, water is tightly controlled through the canals, gates, pumps, and locks to control the flow and movement of water, particularly storm-flows, as it interacts with the natural, built, and coastal environments. Land use assessments and plans also consider climate, hydrological data, geographic data, urban footprints, and infrastructure location, and in several places such as the greater Miami region, sea-level rise and saltwater intrusion, are also considered.

On the west coast that includes the greater Tampa Bay region, the area has historically relied on groundwater pumping for the overwhelming majority of its water supply needs. Environmental issues like saltwater intrusion and over-pumping triggered a long series of changes in the region, and now Tampa Bay Water, formed in 1998 as a regional supply agency, has built a mixed portfolio of water supply to meet regional demand. Over the past 20 years, the agency has built an integrated water supply system which includes a surface water system, groundwater wells, and a seawater desalination plant. This has enabled the agency to shift from being 100 percent reliant on groundwater to a mix of sources with an increasing reliance on surface waters. Ongoing efforts at Tampa Bay Water aim to address potential impact of changing climate on its water supply system.

High-resolution rainfall and ET information (such as gridded, daily data) are currently not available. The level of uncertainty and lack of skill in future projections are also important information gaps. For instance, information on the combined impacts of floods, storm surges and sea level rise need to be developed, e.g. IDF curves (intensity, duration and frequency) for future scenarios. Climate change information needs to be more effectively incorporated in future infrastructure planning and design. Further, understanding of integrated risk management, including insurance and reinsurance needs, is also required. Overall, a vision for effective & dynamic adaptation would need to be developed.

## 3. Climate information needs for water management

### 3.1. Overview

Flooding and water supply are two key issues for South Florida. A lot of planning relies on extreme precipitation metrics, and runoff/flow related metrics. In addition, certain drought or dry spell metrics are also of interest for water supply planning. Several of the metrics suggested by the stakeholders were on sub-seasonal (daily/weekly) time scales and skill of the models in capturing such short time periods can be an issue. In terms of water supply, Tampa Bay was particularly interested in summer rainfall which provides 60-70 percent of the region's water. They were particularly interested in changes in the timing or duration or start date of the summer rainfall, say 20 or 30 years from now.

In addition to the decision-relevant metrics, the managers also showed interest in certain 'upstream metrics' and questions about drivers of events, particularly as they relate to hurricanes/storms and their characteristics going into the future. There was also interest in better understanding the modes of multi-decadal variability in the region, and how models are able to represent these modes. Further, the managers also wanted to understand whether there is a change to be expected in these multi-decadal modes, in the future and how that would impact local scale meteorological phenomena.

Climate change projections for 2030 or 2060, up to the end of the century were stated as useful time scales. In terms of spatial scale, currently there is a focus on point location-based estimates (different rain gauges or weather stations placed across different basins) and using an area extrapolation from those estimates. But, stakeholders showed interest in basin scale average values as well.

### 3.2. List of decision-relevant metrics and their importance

In order for science to be actionable, resource managers need information on decision-relevant climatic metrics. Therefore, one of the first goals of Hyperion was to co-produce the decision-relevant metrics for different management decisions in each of the case study regions. From the water managers' perspective, such metrics quantitatively describe climatic phenomena that are directly related to practical management problems; changes in these quantities would

necessitate shifts in water infrastructure planning and operations. From the scientists' perspective, these metrics can be used to test model fidelity for decision-relevant phenomena and hence push model development and scientific inquiry in more use-inspired directions. Table 1 represents the decision relevant metrics, along with their potential importance, that were developed through iterative engagements between Dec 2016 to Nov 2017. This table is referred from the [published journal article](#) titled "The making of a metric: Co-producing decision-relevant climate science" by Jagannathan, Jones and Ray.<sup>1</sup>

**Table 1: Examples of decision-relevant metrics for each region.**

The table highlights management issues, hydroclimatic phenomena, aspect of phenomena and then each decision-relevant metric. The last column also describes some of the potential decisions or uses for these metrics that were identified by the case study water managers.

Issue	Hydroclimatic Phenomenon	Aspect of Phenomenon	Decision-relevant Metric	Decision/Use
Floods	Streamflow	Peakflow	Multi-year probability distribution of instantaneous maximum flow on an annual scale. i.e. Annual Maxima	Flood management and planning.
Floods	Streamflow	Peakflow	Timing of annual maxima	Flood management and planning.
Floods	Streamflow	Peakflow	5-year 1-hour volume, 5-year 1-day volume, 10-year 3-day volume, 25-year 3-day volume, 100-year 3-day volume	Stormwater infrastructure design criteria (for different infrastructure), and flood management, especially planning water storage facilities.
Water Supply	Streamflow	Low-end Streamflow	7-day 10 year low flow (7Q10)	Water supply planning, especially water quality for dry years, and drought planning.
Water Supply	Streamflow	Low-end Streamflow	Flow anomalies at monthly timescales	Water supply planning to calculate potential water shortages, and drought monitoring.
Water Supply	Streamflow	Annual/Monthly flow	Average Monthly or Annual Flow volumes	Water supply planning for restoration activities.

<sup>1</sup> Jagannathan, K., A. D. Jones, and I. Ray, The making of a metric: Co-producing decision-relevant climate science. *Bull. Amer. Meteor. Soc.*, doi: <https://doi.org/10.1175/BAMS-D-19-0296.1>.

<b>Issue</b>	<b>Hydroclimatic Phenomenon</b>	<b>Aspect of Phenomenon</b>	<b>Decision-relevant Metric</b>	<b>Decision/Use</b>
Water Supply	Streamflow	Monthly flow	Percentage of annual flow occurring in each month	Water supply planning for restoration activities, and reservoir operations management such as managing reservoir use-re-fill cycle.
Water Supply	Streamflow	Variability of Streamflow	Standard deviation of monthly or annual flow volumes	Planning for variability in water supply and for restoration activities.
Water Supply	Streamflow	Peakflow	Daily max flow	Water quality management, e.g. planning water source rotations.
Floods	Rainfall	Extreme Rainfall	Intensity Duration Frequency or IDF curves, specifically, 1-day, 3-day and up to 7-day rainfall events, for 10, 25, 50 and 100 year frequency intervals.	To calculate applicable discharge rates for different stormwater management infrastructure. Design criteria used for drainage and flood protection are in terms of IDFs. In other words, designing standard engineering practices for infrastructure.
Floods	Rainfall	Extreme Rainfall	Change in temporal frequency of historic storms of particular return frequencies (10-100 yrs)	To understand how design storm criteria for different infrastructure may change in the future (roads 10 yrs, houses 100 yrs.), and for recurrent/nuisance flooding monitoring.
Floods	Rainfall	Extreme Rainfall	Frequency and duration of rainfall events greater than 3 or 5 inches in summer season (July, August, September)	For water supply preparedness, to understand cumulative water availability in key supply months.
Floods	Rainfall	Extreme Rainfall	Probable maximum precipitation. For 1-day, 3-day and maybe up to 7-day events	Large storage infrastructure design (like high dams).

<b>Issue</b>	<b>Hydroclimatic Phenomenon</b>	<b>Aspect of Phenomenon</b>	<b>Decision-relevant Metric</b>	<b>Decision/Use</b>
Water Supply	Rainfall	Annual Rainfall	Total annual rainfall volumes	Water supply planning, and drought monitoring.
Water Supply	Rainfall	Monthly Rainfall	Monthly rainfall distributions	Water supply planning, and drought monitoring.
Water Supply	Rainfall	Variability in Rainfall	Rainfall anomalies at Monthly time scales	Water supply planning, and drought monitoring.
Water Supply	Rainfall	Extreme Rainfall	Rainfall Geometric curve (analogous to the SWE triangle). Including start of wet season, duration of wet season, peak volume and other related parameters	Understanding how the annual water season would change in the future.
Water Supply	Droughts/Dry Spells	Drought metrics	Frequency of 1-, 2-, 3- year duration droughts.	Drought prediction and management, and planning future water infrastructure investments.
Water Supply	Droughts/Dry Spells	Drought metrics	Change in temporal frequency of historic multi-year droughts of a particular return period (e.g. 1 in 10 year droughts)	Drought prediction and management, and planning future water infrastructure investments.
Water Supply	Droughts/Dry Spells	Low Rainfall	Probability distribution of annual rainfall totals: particularly focusing on low end of that distribution and how often that occurs	Water supply planning, drought monitoring, and planning future water infrastructure investments.
Water Supply	Droughts/Dry Spells	Demand-Supply gap in Streamflow	Streamflow curves showing the general shape and timing of runoff supply and water demand, and	Water supply planning to calculate potential water shortages, and water conservation measures (reduction targets), drought

Issue	Hydroclimatic Phenomenon	Aspect of Phenomenon	Decision-relevant Metric	Decision/Use
			the differential (or gap) between the two	monitoring, and planning future water infrastructure investments.

## 4. Key scientific activities and results from Hyperion

From the above long list of decision-relevant metrics, project Hyperion’s managers and scientists collectively developed case study science plans that identified a shorter list of scientific activities and metrics that will be a focus of the project (Table A1 in the Appendix).

Out of this long list of decision-relevant metrics, project Hyperion’s stakeholders and scientists collectively decided to focus on the following metrics and scientific activities, as outlined in the case study science plans. These key scientific activities are as follows: Precipitation IDFs, Sea Breezes, and Tropical Cyclones. The rest of this section presents a narrative description of these three short-listed scientific activities. The key motivation, methods, results and limitations from each of the three scientific activities, are summarised below.

### 4.1. Precipitation IDF curves

#### Summary

- This work analyzes how different climate models vary in their IDF estimates for the past and the future. It also proposes a framework that allows for examining IDF estimates for longer return periods, where the data sample size can be a limitation.
- The study finds that there is considerable variability within and across models in both predicting historical IDFs as well as in IDF projections of the future.
- A method is proposed that employs pooling of model data based upon historical performance of models. The models selected for pooling are bias corrected and then used for estimation of non-stationary IDF curves. The proposed method reduces estimation uncertainty due to enhanced sample size.

#### 4.1.1. Background

IDF or Intensity Duration Frequency estimates are used for planning and management of extreme precipitation events. The curves specify the magnitude (i.e., intensity) of precipitation events across a range of durations and return periods (i.e., frequencies). These estimates provide information to support a wide variety of civil activities such as designing flood protection

structures and urban drainage systems. However, there are significant uncertainties and variability in climate models' predictions of extreme precipitation. In the case of IDFs, the estimation uncertainty increases as one considers longer return periods since larger sample sizes are needed to estimate rarer events (e.g., assessing IDFs for 100-yr return period requires at least 100 years of data). Not many studies have critically examined the variability among different models in predictions of IDFs. In addition, the few studies that provide projections of IDFs for the future, either take a mean or median of IDF estimates from different models which may not address the issue of uncertainty due to small sample sizes and variability across models. Therefore, this research also proposes a new methodology that can help to reduce some of the issues associated with limited sample size for IDF estimations. The underlying hypothesis for this work is that, due to data (sample size) limitations for IDF estimations, and uncertainties, new and novel methods of combining model data of IDFs may be needed to better evaluate this metric. This research focuses on the following key research questions:

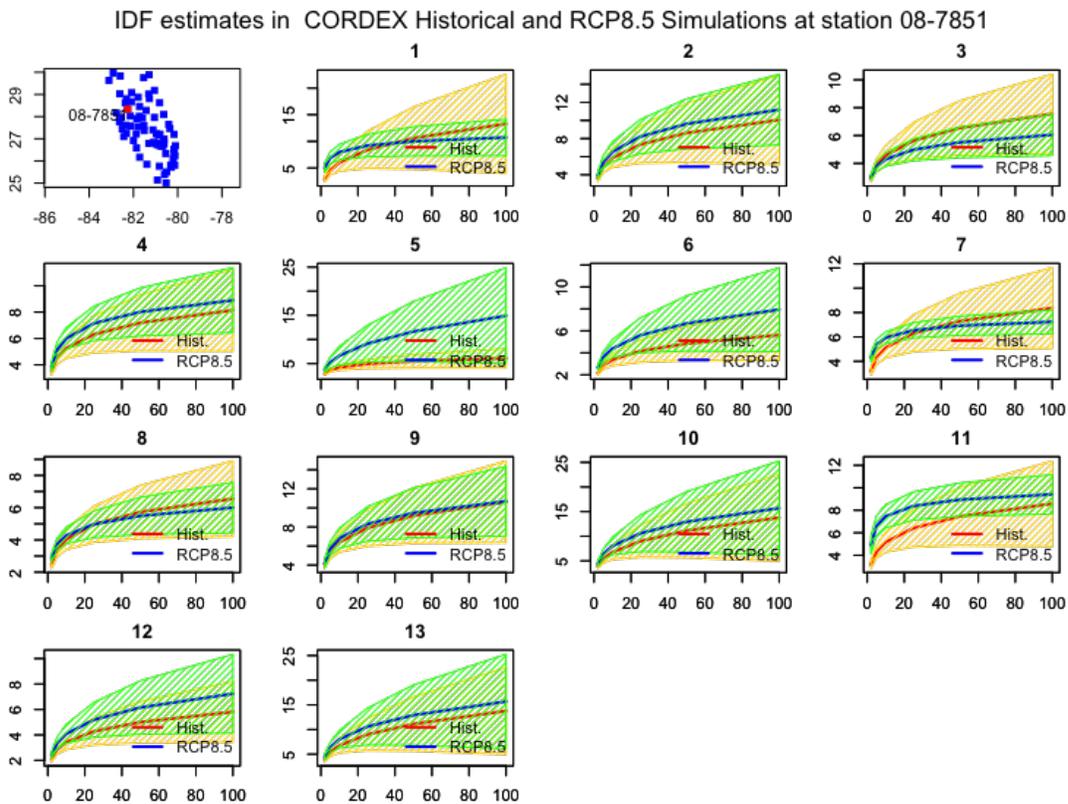
1. ***How do climate models vary in their IDF estimates of the past and future? What are the differential capabilities of climate models in predicting historical IDFs?***
2. ***Do models show a statistically significant change in IDF estimates in future time periods as compared to historical?***
3. ***What framework allows for analyzing changes in IDF estimates despite sample size limitation, natural variability across space, and variability across models?***

IDF estimates were computed for historical (1956-2005) and RCP8.5 simulations (2049-2098) of the NA-CORDEX models. To provide station-wise results, model data was interpolated to station locations using nearest neighbor interpolation. The reference weather station data was obtained from NOAA Atlas 14, 24-hr precipitation data from GHCN archive. Every station that had 50 years data between 1950-2005 was included. The 12 NA-CORDEX models (with 0.25x0.25deg resolution) that were evaluated are: CanESM2.CanRCM4 (A or 1), CanESM2.CRCM5-OUR (B or 2), CanESM2.CRCM5-UQAM (C or 3), GFDL-ESM2M.CRCM5-OUR (D or 4), GFDL-ESM2M.RegCM4 (E or 5), GFDL-ESM2M.WRF (F or 6), HadGEM2-ES.RegCM4 (G or 7), HadGEM2-ES.WRF (H or 8), MPI-ESM-LR.CRCM5-OUR (I or 9), MPI-ESM-LR.CRCM5-UQAM (J or 10), MPI-ESM-LR.RegCM4 (K or 11), MPI-ESM-LR.WRF (L or 12), and MPI-ESM-MR.CRCM5-UQAM (M or 13). Other datasets such as Variable Resolution CESM and LOCA downscaled data were also analyzed but are not presented here for brevity.

The IDF estimations are based on univariate extreme value analysis that uses the method of maximized likelihood estimation. The generalized extreme value (GEV) distribution was then fitted to the data in a non-stationary framework. IDF estimates were computed from a sample of 50 years, for 24-hour duration events at; 2, 5, 10, 25, 50, and 100-year return periods. Historical and future IDF estimates from different models were computed for each weather station, and the resultant intra and inter-model variability in IDF results was examined. Since the 50-year sample size was limiting (especially for assessing longer return periods), models were then bias-corrected using quantile matching so that all models have the same historical distribution as the observations. The models (ones that accurately capture space and time variability of select precipitation metrics) were pooled together to develop a long time-series of data (i.e. if 5 models with 20 years of data can be pooled, it can yield 100 years of data).

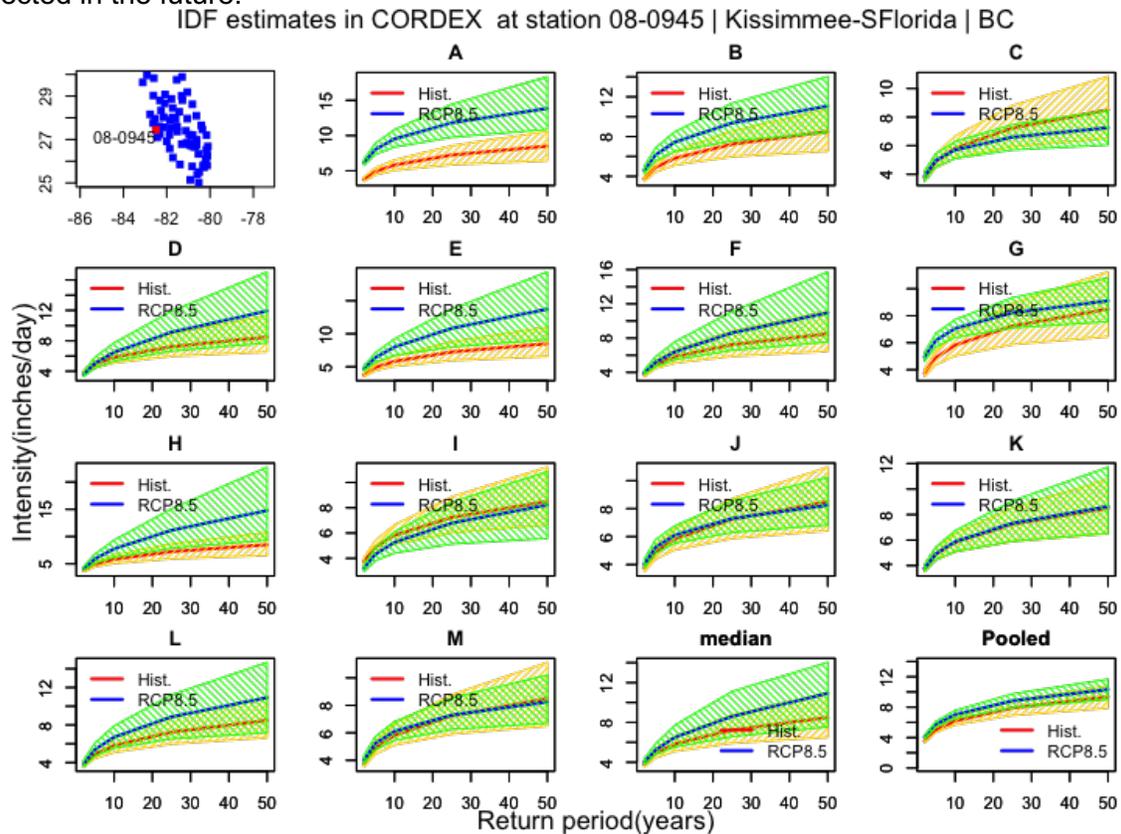
## 4.1.2.Key Results

Figure 2 shows the results of a comparative skill evaluation of models in predicting historical IDF for a 24 hour duration storm at one weather station. This analysis was done for over 50 weather stations across the region. The results show that there is considerable variability across models in predicting both the historical and future IDF. Also, the estimation uncertainty is large in models. For example, model 1 estimates historical rainfall intensity for 24 hour duration storms of different return periods between 2 and 22 inches whereas model 2 predicts the same at between 2 and 14 inches. The estimation uncertainty particularly increases with higher return periods. Apart from other factors this could be due to small sample size. Figure 2 also gives IDF projections for the weather station from the different models. The figure suggests there may be an increase in IDF estimates in the future, but this change (change between historical and projected IDF for each model) may not be statistically significant. Further, some models (e.g. 1,3 and 7) do not show a clear increase in precipitation for all return periods, there is also a large variability between the projections in different models i.e. both intra and inter model variability in projected IDF is high, and hence a statistically significant change signal is not seen.



**Figure 2: NA CORDEX models' predictions precipitation intensity estimates for a 24-hour duration storm**  
The first panel shows weather stations and their associated latitude longitude. Panels labelled 1-13 represent the different NA CORDEX models' predictions of historical and future precipitation intensity estimates for a 24-hour duration storm. Red is the historical IDF estimate, and the yellow shaded area is the 95% confidence interval (CI) around it. Similarly, blue is the future IDF estimate and the green area is the 95% CI around it.

In order to overcome the limitation of a small sample size, a methodology for pooling different models' data to create a large sample size was developed. This methodology required bias correction of data. Bias corrected and pooled results are presented in Figure 3. From the figure we can gather the pooling model data reduces some of the estimation uncertainty by increasing the precision of the results due to larger sample size (as compared to using individual models or taking the median results). This pooling enables the detection of a statistically significant change or IDF estimates is seen between historical and RCP 8.5, showing that when the models are combined together the changed signal is clearer. This approach overcomes the small sample size limitations, thereby providing a clearer picture of the change in IDF estimates that may be expected in the future.

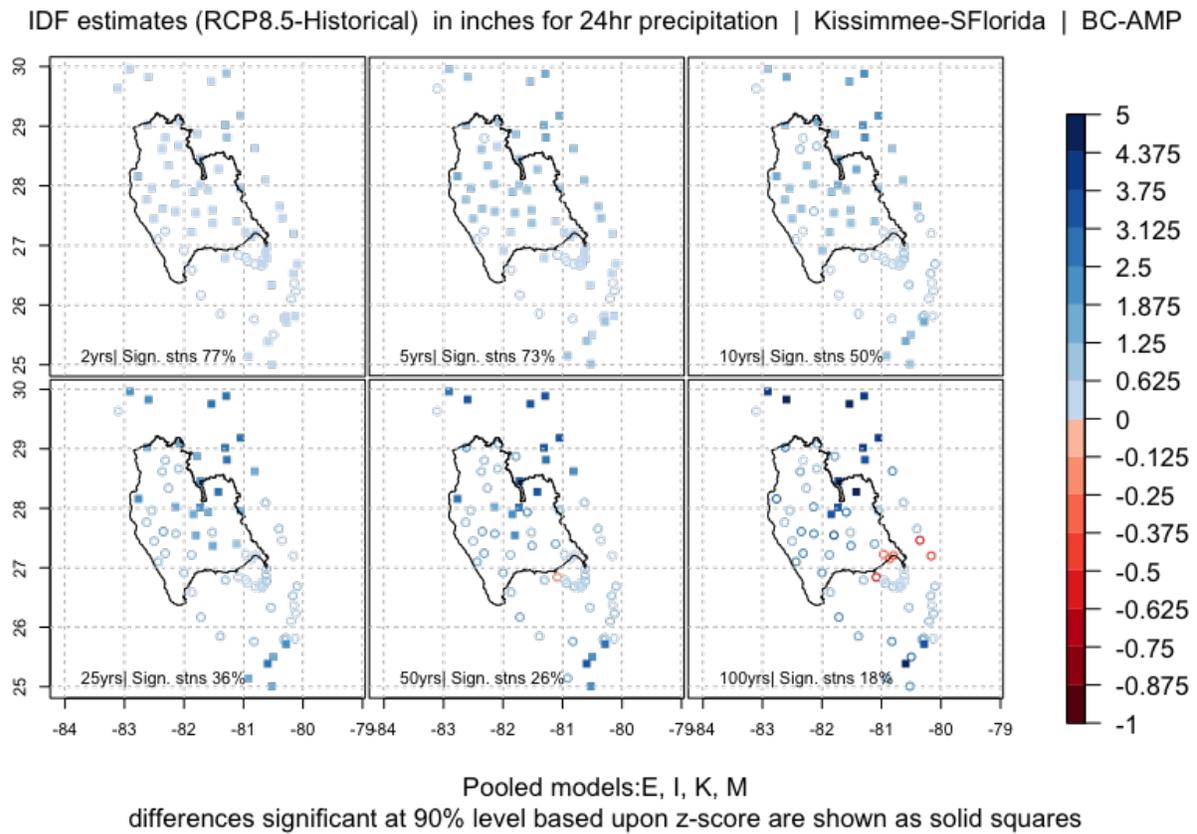


**Figure 3: 24-hr IDF estimates in the bias-corrected historical and future simulations.**

The red curve and yellow shading indicate IDF estimates and corresponding 90% confidence interval in the bias-corrected historical simulation. The blue curve and green shading indicate IDF estimates and corresponding 90% confidence interval in the bias-corrected RCP8.5 simulation. The reader should note that historical IDF estimates and their 90% confidence intervals are the same in all models. The red curve with orange shading varies between panels solely due to changes in the vertical scale. The "median-all" panel shows the median of IDF estimates from all models, while the "median-pooled" shows the median of the IDF estimates from models that are used for pooling. "Pooled" shows the IDF estimates computed from pooling of better performing models. Models that are used for pooling are shown in red letters in the top left corner of the figures. The X-axis indicates return periods in years and Y-axis indicates intensity in inches/day. This figure shows that the pooling method enables the detection of a significant change between IDF estimates of the past and the future for return periods less than 10 years.

Figure 4 presents results for all weather stations where only models with good skill score for the annual maximum precipitation (AMP) metric were pooled (a station-wise version of this figure is

provided in Appendix 1- Figure A1). The figure again showcases that a statistically significant change between historical and RCP 8.5 is seen when the models are pooled together to create a large sample of data. With this pooling method, most weather stations in the region show an increase in intensity of precipitation for the 24 hour duration storm for recurrence intervals less than 10 years. For longer return periods, the models still indicate an increase in precipitation, however this change is not significant in all the stations that were examined.



**Figure 4: Changes in 24-hr precipitation for 2, 5, 10, 25, 50 and 100 year return periods computed from pooled models.**

The differences that are significant at the 90% significance level are shown as solid squares and those not significant at 90% are shown as blank circles. The significance is computed using the z-statistic as defined in section 3.2.3 of Srivastava et al. 2019. Units are in inches/day. Significant stations shows the percentage of stations at which the differences are significant. Pooled models: (E), (I), (K) and (M).

### 4.1.3. Discussion and Conclusions

Considering widespread variability across models, using a multi-model estimate of IDF seems a better option than relying on any single model. The IDF estimates based upon bias-corrected and pooled model data offers a larger sample size that enables the detection of significant increases in future precipitation estimates at more stations than any other method. This method can be applicable to any region or spatial scale, even where models do not agree well with each other and are data limited.

Although this work improves on current capabilities for assessing IDF estimates across different models, it is to be noted that the pooling method increases precision but not accuracy of results. We note that even the “best” performing models differ quite substantially in their estimates. Hence, alternate methods for evaluating accuracy of model predictions may be needed. One such method that has been pursued by this research team, is examining model skill for other extreme precipitation indices. This work has assessed the skill of simulations of observed precipitation indices (P-indices) in the historical runs of regional climate models in the NA-CORDEX program. Some of the results from this work are presented in the Appendix (Figures A2-A5). Since the models that perform well in Susquehanna and Florida differ, more research is needed to better understand the precipitation related processes that are more dominant in different regions. Further research on the topic must also focus on understanding the causes for models performing good/ bad (resolution, dynamics, convection, parameterization). Continual refinement of the multi-model approach to estimate future precipitation changes/ IDF estimates is also needed.

## 4.2. Sea Breeze

### Summary

- This work examines the contribution of the Florida sea breeze to summertime rainfall variability over land, and analyzes how the CAM model represents sea breezes.
- A sea breeze detection algorithm was developed, which is able to detect sea breeze reasonably well (the observed summertime sea breeze detection rate between 2009-2018 was 84%).
- Results show that on days with classic sea breezes, the probability of precipitation is higher than 75% in some regions, especially when sea breeze is detected along both coasts. In terms of model skill, the CAM-REF-VR28 model shows 5 times less frequent and weaker sea breeze occurrences as compared to the observed sea breezes.

### 4.2.1. Background and Methods

This research was inspired by the fact that most studies on large scale precipitation in Florida, show that 50% or more of the precipitation is coming from unorganized convection and not associated with mesoscale or larger systems. However, the relationship between sea breeze and summertime rainfall is not very well understood because tracking sea breezes can be difficult. Further, the skill of climate models (which typically have a coarser resolution of around 25 km as compared to phenomena like sea breeze which can be at smaller scales), for predicting such local scale circulations is also not well studied. Water managers in the region are interested in better understanding the influence of such diurnal forcings on precipitation patterns during the rainy season, and how it might be changing.

This study hence aims to provide a better understanding of the relationship between sea breeze and summertime rainfall through developing a sea breeze tracking algorithm and assessing the

probability of precipitation during sea breeze days. Further, it also examines the skill of a variable resolution climate model in detecting sea breezes. Understanding how well models are able to capture this critical local phenomenon can help point to areas for model improvement, so that they are able to better predict the complex precipitation related processes in the region. This research asks two questions:

- 1. What is the contribution of the Florida sea breeze to summertime rainfall variability over land?**
  - a. Hypothesis: the presence of the Florida sea breeze circulation enhances convergence and convection and leads to an increase in daily precipitation compared to days without sea breeze.*
- 2. How well can models like the Community Atmosphere Model (CAM) represent local circulation such as the sea breeze?**
  - a. Hypothesis: the model's horizontal grid spacing may be too coarse to represent the occurrence of FL sea breeze in the variables typically used to detect and characterize sea breeze observationally.*

As a first step, a sea breeze detection algorithm was developed and tested using observed data to see how well the algorithm is able to detect observed sea breezes. To do this observational analysis, weather data from 13 stations in the Florida Automated Weather Network (FAWN) from 2009-2018 (June July August - JJA) were analyzed, to detect the presence of a sea breeze based on objective criteria involving wind and temperature. FAWN data are station data with 15-min resolution, while the detection algorithm is based on hourly averages. Sea breeze occurrence was detected based on comparing coastal weather stations with inland stations for key variables such as differences in temperature, dew point temperature, and wind speed and wind direction ( $T$ ,  $T_d$ ,  $WS$ ,  $WD$ ). Based on this, the number of days of sea breeze occurrence was recorded, and the sea breezes were characterized by coastline (East, West, Both), time of day, and date of occurrence. Synoptic types were calculated based on average low-level atmospheric flow from the daily 12UTC Tampa radiosonde.

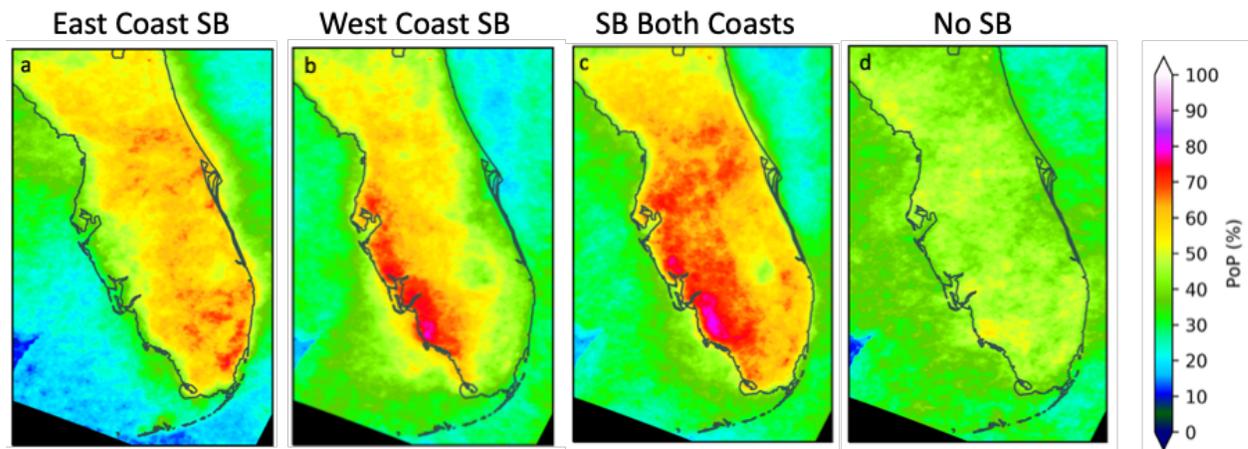
Next, the relationship between sea breeze occurrence and precipitation was analyzed using daily precipitation on days with and without sea breezes, which were compared generally and also by synoptic type. The precipitation data are NCEP 4-km gauge-corrected from 2009-2018 (hourly – accumulated to daily 12pm-3am). Probability of precipitation and percent of total precipitation was calculated.

For the model skill evaluation, a variable resolution model - CAM-REF-VR28's output from 2005-14 was analyzed. Similar model variables ( $T$ , wind speed and direction at 3-hr, 28-km resolution) were examined to compare the sea breeze occurrence statistics between observations and model output. Grid cells were selected near observational stations. Synoptic types were calculated using 6 grid-cell average of 850-hPa wind vector.

## 4.2.2. Key Results

The sea breeze algorithm that was developed is able to detect sea breeze reasonably well, the observed summertime (June, July, August- JJA) sea breeze detection rate 2009-2018 was 84%. In other words, the algorithm was able to detect a sea breeze in either the west or the east coast 84% of the time. This also showed that these observed sea breezes are most likely to occur under low-level wind speeds (i.e. weak synoptic influence with wind speeds <5 m/s) with westerly or southerly winds.

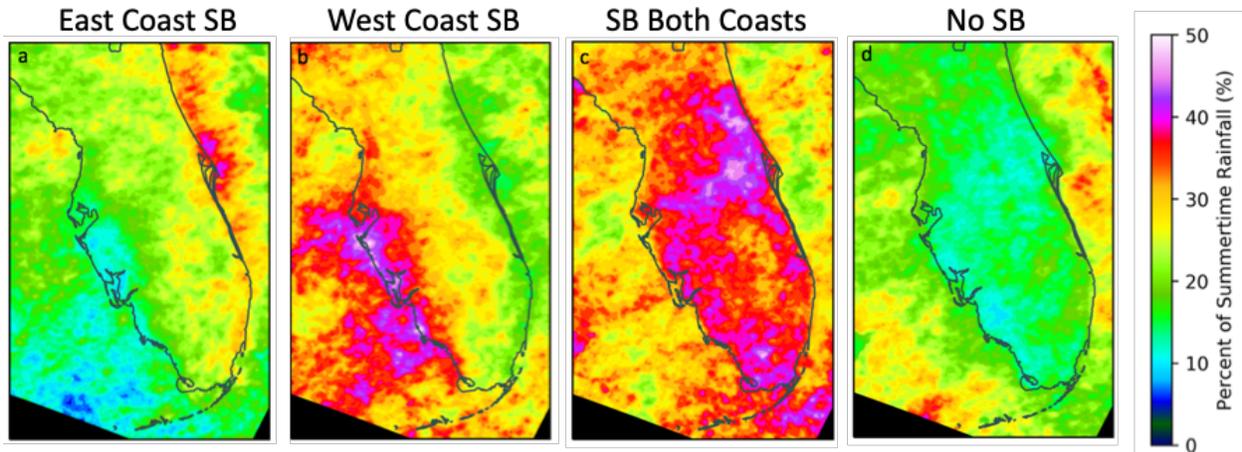
On days with classic sea breezes, the probability of precipitation is higher than 75% in some regions especially when sea breeze is detected along both coasts (Figure 5).



**Figure 5: Summertime daily probability of precipitation (PoP) for classic sea breeze occurrence**

Summertime (June-July-August) daily probability of precipitation (PoP) for classic sea breeze occurrence on the east coast (a; 22% of days), west coast (b; 31% of days), both coasts (c; 31% of days) and when no classic sea breeze was detected on either coast (d; 16% of days).

Some regions experience >50% of daily summertime accumulated rainfall on sea breeze days, even though these days only represent ~31% of summertime days (Figure 6). This suggests that there is a significant amount of precipitation associated with sea breeze days in the summer, although these results cannot identify the additional precipitation falling due to sea breeze (as other local pop-up thunderstorms may also be contributing to this precipitation).



**Figure 6: Percent of precipitation during summertime (JJA) days where a classic sea breeze is detected**  
 Percent of precipitation during summertime (JJA) days where a classic sea breeze is detected by at least one station on the east coast (a; 22% of days), west coast (b; 31% of days), both coasts simultaneously (c; 31% of days), and days where a classic sea breeze is not detected by any test station (d; 16% of days).

In terms of model skill, the current analysis of model data shows model sea breezes for 8 years (2005-2012). Sea breeze occurs about 5 times less frequently in the model when the same detection criteria are used for observations (refer Table 2). When the detection thresholds were lowered, particularly for the difference in temperature between test and reference station, then more sea breezes were detected (refer row for Model-adj, in Table 2). This and other sensitivity analyses indicate that these results are somewhat sensitive to the wind speed criteria, and more sensitive to the temperature gradient criteria. Further, the model shows more frequent occurrences of easterly winds when compared to observations when applying the criteria for the 6-pixel average. This model bias can be an important finding for improving model performance for this local phenomenon.

**Table 2: Average number of sea breezes detected per year over the 2005-2012 time period for observations (Obs) and for CAM-REF-VR28 (Model) output.**

Average number of sea breezes detected per year over the 2005-2012 time period for the FAWN observations (Obs) and for CAM-REF-VR28 (Model) output. The analysis was repeated with a lower temperature threshold for the model results (Model-adj), which increased the average number of detected days with sea breezes.

Average Number of SB Detected per year

	East Coast	West Coast
Obs	22	30.3
Model	5.7	4.5
Model - adj	7.8	9.3

### 4.2.3. Discussions and Conclusions

Sea breeze related precipitation is an important contributor to the overall summertime precipitation, especially when there are west coast sea breezes or sea breezes on both coasts. More than 35% of rainfall across agricultural regions in Florida occurs on days where a sea breeze was detected along both coastlines (31% of days). However, no significant difference in

rainfall intensity was found when sea breezes occur. In terms of model skill, the CAM-REF-VR28 model shows 5x less frequent and weaker sea breeze occurrences as compared to the observed sea breezes. The model also shows more equal partitioning of sea breezes between the east and west coasts, and appears to have larger frequency of easterly winds than seen in observations. The broader takeaway from this research is that, peninsular regions prone to sea and/or lake breezes can see an enhancement of convergence and uplift leading to greater probability of rainfall on days when sea breeze is occurring, although some climate models may not detect the occurrences of these sea breezes accurately.

Inland stations that are used as “control stations” in this analysis, are near Lake Okeechobee so may have some lake breeze biases, although these have been accounted for to a certain extent, by eliminating the ones that are very close to the lake. The current work cannot say how much additional precipitation falls due to sea breeze, although it can say that probability of precipitation is significantly higher on days with sea breeze. Separating out sea breeze effects from other local thunderstorm outflows is difficult from an image processing standpoint. For the model outputs only two types of sea breezes with wind (wind speed, wind direction) and temperature shifts are detected. More work is needed to detect the sea breezes due to difference in humidity, although it is unclear whether and how this will change results. It has been suggested in WRF studies that sea breeze is expressed more strongly in model humidity than temperature or wind vector. Further analysis is also needed to determine the sensitivity of these results. Additional sensitivities relating to the land masking criteria in the model will also be explored.

## 4.3. Tropical Cyclones

### Summary

- This study investigates the skill of variable resolution models in tracking the frequency and spatial distribution of landfalling Tropical Cyclones (TCs) along the Eastern US, as well as the mean and extreme precipitation associated with these storms.
- The variable resolution CAM5 configuration was able to produce reasonable frequency and spatial distribution of landfalling TCs along Eastern U.S, representing a significant improvement from the conventional climate model.
- Compared to observations, the variable resolution simulations underestimate the percent of extreme precipitation from tropical cyclones. Despite this low bias, it is a significant improvement over the conventional climate simulations (100 km spatial resolution).

### 4.3.1. Background and Methods

Coastal storms, such as tropical and extratropical cyclones (TCs and ETCs), are responsible for a substantial amount of disaster related losses in the U.S. every year due to a combination of rainfall, high-winds and storm surge. About 39% of the U.S. population live in counties directly on the coastline, where a significant amount of the nation’s critical energy infrastructure is also

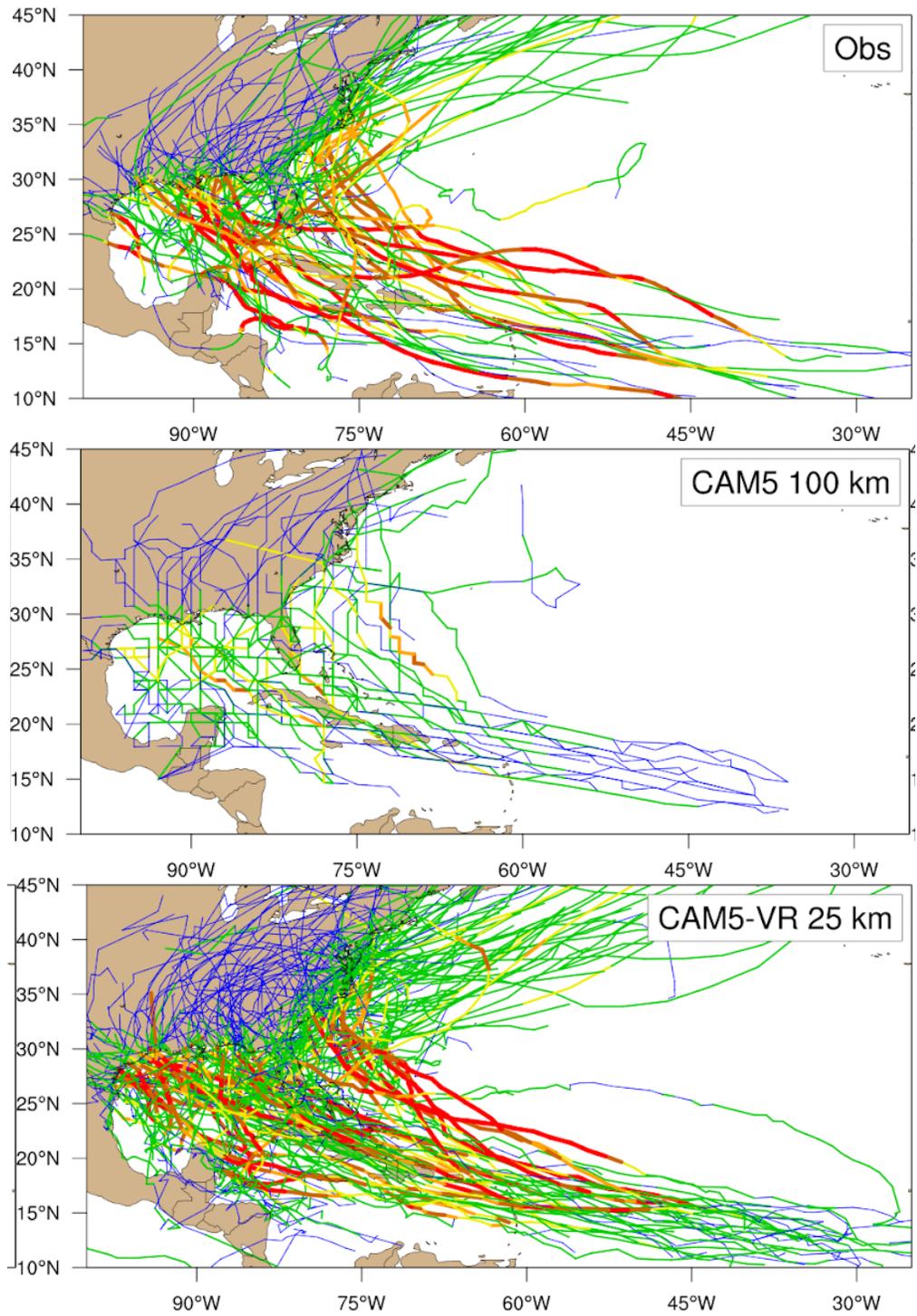
located. A prime example of this is Florida, where tropical cyclones can lead to heavy amounts of precipitation over much of the state's area. Credible simulation of these events can help better understand and prepare for future changes in precipitation associated with these events. However, conventional climate models often struggle to resolve tropical cyclones and therefore the rainfall associated with these storms. Evidence suggests that high-resolution climate models at grid spacings of approximately 25 km may be needed for appropriately simulating aspects of regional scale extreme precipitation events. However, there is no clarity on the exact skill of these models for specific phenomena. Since increasing model resolution can be costly in terms of resources and time, better understanding of the overall benefits of increased resolution can be useful for effective use of resources. This research asks the following research questions:

- 1. How well do variable resolution models track the frequency and spatial distribution of the landfalling Tropical Cyclones (TCs) along the Eastern US?**
- 2. How do TCs relate to mean & extreme precipitation, and how well does a variable resolution model simulate mean and extreme precipitation associated with TCs?**

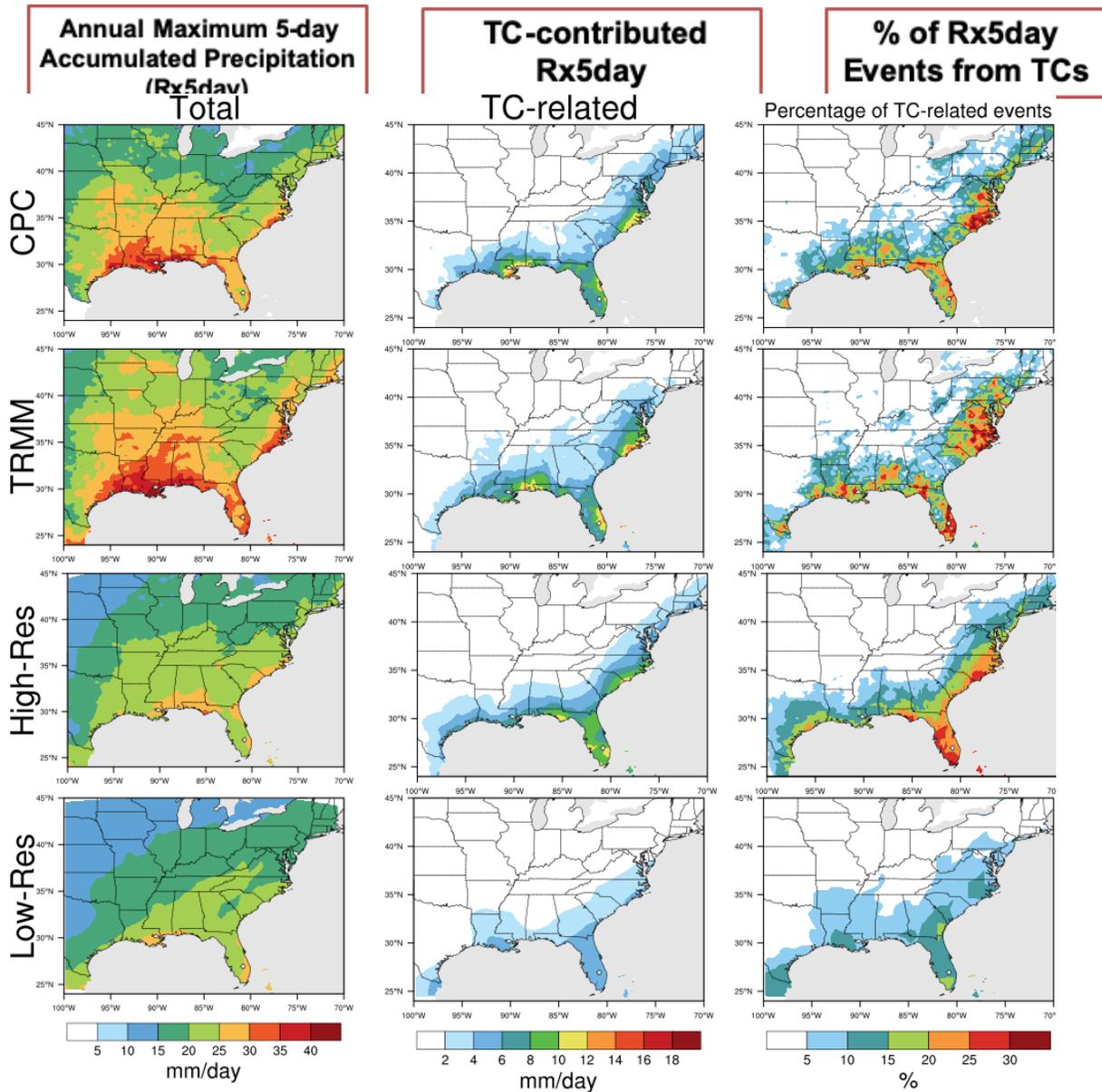
A tropical cyclone rainfall analysis tool is developed with TempestExtremes to extract precipitation associated with individual TCs, and quantify their contribution to average and extreme (Rx5day) precipitation in the Eastern U.S. Occurrence of storms and the precipitation associated with them is analyzed from a suite of CESM-CAM5 variable resolution simulations with high-resolution domains (25 km) over the North Atlantic and Eastern. These high-resolution simulations are compared to a conventional climate simulation (100 km), and to observations.

#### 4.3.2. Key Results

The variable resolution CAM5 configuration was able to produce reasonable frequency and spatial distribution of landfalling TCs along Eastern U.S. This model simulated 12.5 TC/yr as compared to the observed 13 TC/yr in the North Atlantic. This represents a significant improvement from the conventional climate model which was able to track only 2.6 TC/yr (Fig 7). The variable resolution simulations suggest that although these storms only contribute 1-2% to the average annual precipitation, their contribution to extreme precipitation is much larger, approaching 30% in some regions along the coast. When compared to observations, the variable resolution simulations slightly underestimate the percentage of extreme precipitation from TCs. For example, observations suggest that Florida receives about 17-18% of its extreme precipitation from TCs, but CAM5 with variable resolution simulates 15%, which is an improvement over the standard resolution CAM5 (13%) (Fig 8). Despite this slight low bias it is a significant improvement over the conventional climate simulations (100 km spatial resolution).



**Figure 7: Comparison of observed and CAM 5 simulated landfalling TC tracks at different resolutions.** Comparison of observed Eastern United States landfalling TC tracks to CAM5 simulated landfalling TC tracks at a conventional climate model resolution of 100 km and CAM5-VR with 25 km resolution. The colors of the lines represent the intensity of the storm as measured by the Saffir-Simpson scale. The CAM5-VR configuration performs well at simulated TC landfalls compared to the conventional climate model approach.



**Figure 8: Rx5day, TC-related Rx5day, and % of Rx5day events due to TCs for observations, CAM5 variable resolution and CAM5 conventional climate model resolutions.** Rx5day (annual maximum 5-day accumulated precipitation) [mm/yr] (left column), TC-related Rx5day [mm/yr] (middle column), and percentage of Rx5day events due to TCs (right column) for observations (top), CAM5 variable resolution (middle) and CAM5 conventional climate model resolutions. The CAM5 variable resolution configurations shows an improvement in capturing the TC extreme precipitation magnitude and percentage of events over Florida. Top two rows are observed precipitation from Climate Prediction Center (CPC) and from Tropical Rainfall Measuring Mission (TRMM) data.

#### 4.3.3. Discussions and Conclusions

This research has developed a novel analysis tool for tracking tropical cyclones and quantifying the precipitation associated with them. In general, the high-resolution CESM model along with the domains used for Hyperion, offer a significant improvement in simulating tropical cyclone

frequency in the North Atlantic. This improvement also leads to a reasonable simulation of extreme precipitation due to tropical cyclones (albeit slightly reduced), particularly for the Florida region. Work in the near-future will focus on analyzing future climate simulations for these storm related precipitation metrics. Additional future work will focus on understanding changes in storm structures and how this impacts precipitation associated with the storms.

## Acknowledgements and way forward

We are deeply grateful to all of Project Hyperion's water managers and scientists who patiently participated in the many back-and-forth engagements that form the basis of this report. We are also thankful to Bruce Riordan who co-led the engagements, Paul Ullrich for his agile leadership of the project, and Smitha Buddhavarapu for her careful review and edits of this report. Hyperion's successor project "HyperFACETS" is currently underway (2019-present) and will expand on the project's research activities and further work on creating broadly applicable tools for co-producing actionable climate science.

This work was supported by the Office of Science, Office of Biological and Environmental Research, Climate and Environmental Science Division, of the U.S. Department of Energy under contract DE-AC02-05CH11231 as part of the Hyperion Project, An Integrated Evaluation of the Simulated Hydroclimate System of the Continental US (award DE-SC0016605).

## Appendix 1

**Table A1: List of metrics and summary of scientific activities pursued by Hyperion project**

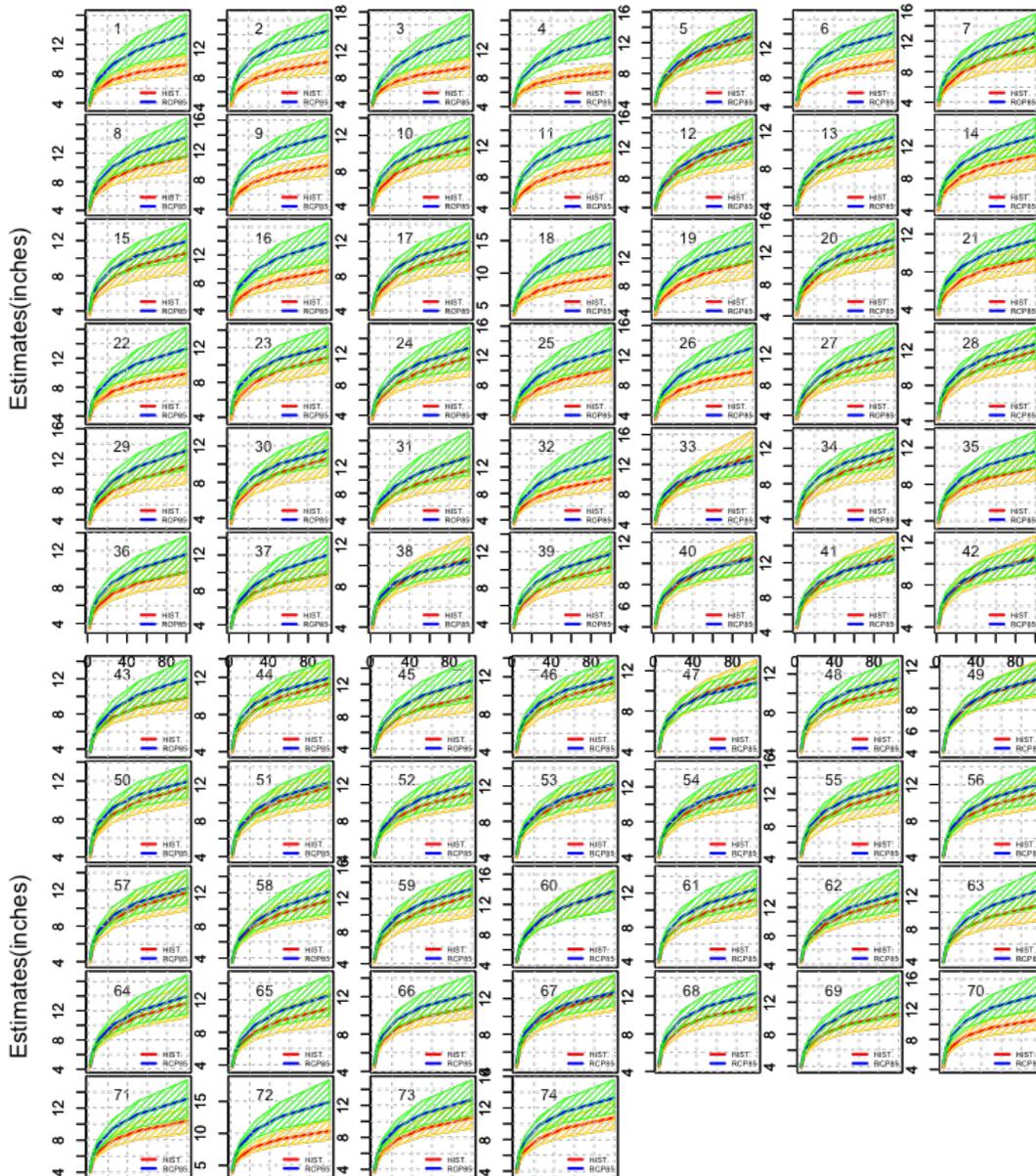
S No.	Science Activity	Lead Scientists	Description
1.	Precipitation IDF curves: Model skill and future projections	Abhishekh Srivastava and Richard Grotjahn	<p>This work analyzed annual maximum precipitation (AMP) in historical simulations of VR-CESM and NA-CORDEX models. Models have considerable biases with respect to the station based AMP. A new approach is adopted wherein models are selected based upon their historical performances. The historical and future (RCP8.5) simulation data of selected models are then bias-corrected and pooled to estimate non-stationary changes in the IDF estimates of 24-hr precipitation.</p> <p><b>Related papers:</b> Srivastava et al., 2019 “<b>Quantifying changes in 24-hr precipitation extremes by pooling NA-CORDEX models: Susquehanna watershed and Florida peninsula</b>” (under review in Water Resources Research).</p>
2.	Precipitation metrics (other than IDF): Skill evaluation	Abhishekh Srivastava and Richard Grotjahn	<p>Precipitation indices (ETCCDI) namely annual mean P, SDII, CDD, CWD, Rx1day, Rx5day, R10mm, R20mm and Fr95T have been analyzed in the historical NA-CORDEX models. The two performance criteria (1) Taylor diagram and (2) interannual variability skill score (IVSS0 are used to estimate model performance against station-based data. The results are summarized in the form of “heat map” (also called stop-light diagram or portrait diagram). The analysis indicates that models have moderate skills in simulating both the observed spatial and temporal patterns of ETCCDI indices.</p>
3.	Sea Breeze: Model skill evaluation	Dana Veron and Sara Rauscher	<p>This work examines the contribution of the Florida sea breeze to summertime rainfall variability over land, and analyzes how the CAM model represents sea breezes. A sea breeze detection algorithm was developed, which is able to detect sea breeze reasonably well (the observed summertime sea breeze detection rate between 2009-2018 was 84%). Results show that, on days with classic sea breezes, the probability of precipitation is higher than 75% in some regions especially when sea breeze is detected along both coasts. In terms of model skill, the CAM-REF-VR28 model shows 5x less frequent and weaker sea breeze</p>

S No.	Science Activity	Lead Scientists	Description
			<p>occurrences as compared to the observed sea breezes.</p> <p><b>Related papers:</b>  A Characterization of Sea Breeze Enhanced Rainfall in South Florida, Moore et al. 2020, under review in JAMC.</p> <p>Moore, D. P., and <b>D. E. Veron</b>, 2020: Automated detection algorithm for sea breeze fronts using ground-based radar, <i>Journal of Atmospheric and Oceanic Technology</i>, to be submitted Jan 2020.</p> <p>Moore, D., and <b>D. E. Veron</b>, 2019: Automated Detection of Sea Breeze Circulations: A Climatology of Sea Breeze-Initiated Precipitation, American Meteorological Society's 23<sup>rd</sup> Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS), 99<sup>th</sup> Annual Meeting, Phoenix, AZ, USA.</p> <p>Moore, D.P., 2019. Detection and Analysis of Sea Breeze and Sea Breeze Enhanced Rainfall: A Study of the Florida and Delmarva Peninsulas (Master's thesis, University of Delaware).</p>
4.	Extra Tropical Cyclones: Model skill and sensitivity to variable resolution domain	Colin Zarzycki	<p>This project developed metrics and software packages in order to evaluate coastal storms in gridded climate data, and evaluated the performance of these metrics in reanalyses, which allows a direct comparison to observations. Metrics were iteratively improved and showed capability to automatically extract key hydrologic events. The study also evaluated the performance of a current class of Earth System Models (ESMs) at ~1deg resolution to reproduce ETCs over historical period. Through use of variable-resolution ESMs, the project evaluated the resolution sensitivity and value add associated with finer grid spacings and ETCs.</p> <p><b>Related papers:</b>  P. A. Ullrich and <b>C. M. Zarzycki</b> (2017), TempestExtremes: A framework for scale-insensitive pointwise feature tracking on unstructured grids, <i>Geosci. Model Dev.</i>, <b>10</b>, 1069-1090, doi:10.5194/gmd-10-1069-2017. (<a href="#">Github repository</a>)</p> <p><b>C. M. Zarzycki</b> (2018), Projecting changes in</p>

S No.	Science Activity	Lead Scientists	Description
			societally-impactful northeastern U.S. snowstorms. <i>Geophys. Res. Lett.</i> , <b>45</b> , 12067–12075, doi:10.1029/2018GL079820.
5.	Tropical Cyclones: Model skill and sensitivity to variable resolution domain	Kevin Reed	<p>These studies investigate the skill of climate models and variable resolution models in tracking the frequency and spatial distribution of landfalling Tropical Cyclones (TCs) along the Eastern US, as well as the mean and extreme precipitation associated with these storms. The work also explores the large-scale climate controls on TCs in the North Atlantic, as well as model skill in storm intensity.</p> <p>The variable resolution CAM5 configuration was able to produce reasonable frequency and spatial distribution of landfalling TCs along Eastern U.S, representing a significant improvement from the conventional climate model.</p> <p>When compared to observations the variable resolution simulations underestimate the percentage of extreme precipitation from tropical cyclones. Despite this low bias it is a significant improvement over the conventional climate simulations (100 km spatial resolution).</p> <p><b>Related Papers:</b>  Chavas, D. R., <b>K. A. Reed</b> and J. A. Knaff (2017), Physical understanding of the tropical cyclone wind-pressure relationship, <i>Nature Communications</i>, 8, 1360, doi: 10.1038/s41467-017-01546-9.</p> <p><b>Reed, K. A.</b>, J. T. Bacmeister, <b>J. J. A. Huff</b>, X. Wu, S. C. Bates and N. A. Rosenbloom (2019), Exploring the impact of dust on North Atlantic hurricanes in a high-resolution climate model, <i>Geophys. Res. Lett.</i>, 46, 1105–1112, doi: 10.1029/2018gl080642.</p> <p><b>Stansfield, A. M.</b>, <b>K. A. Reed</b>, <b>C. M. Zarzycki</b> and <b>P. A. Ullrich</b> (2019), Assessing Tropical Cyclones' Contribution to Precipitation over the Eastern United States and Sensitivity to the Variable-Resolution Domain Extent, <i>J. Hydrometeorol.</i>, submitted.</p>
6.	Future projections from VR-CESM	Colin Zarzycki and Kevin Reed	This study applied CESM Large Ensemble model runs to evaluate the projected changes in snowfall and total precipitation associated with ETCs over the northeastern United States during a mid-century and end-of-century period.

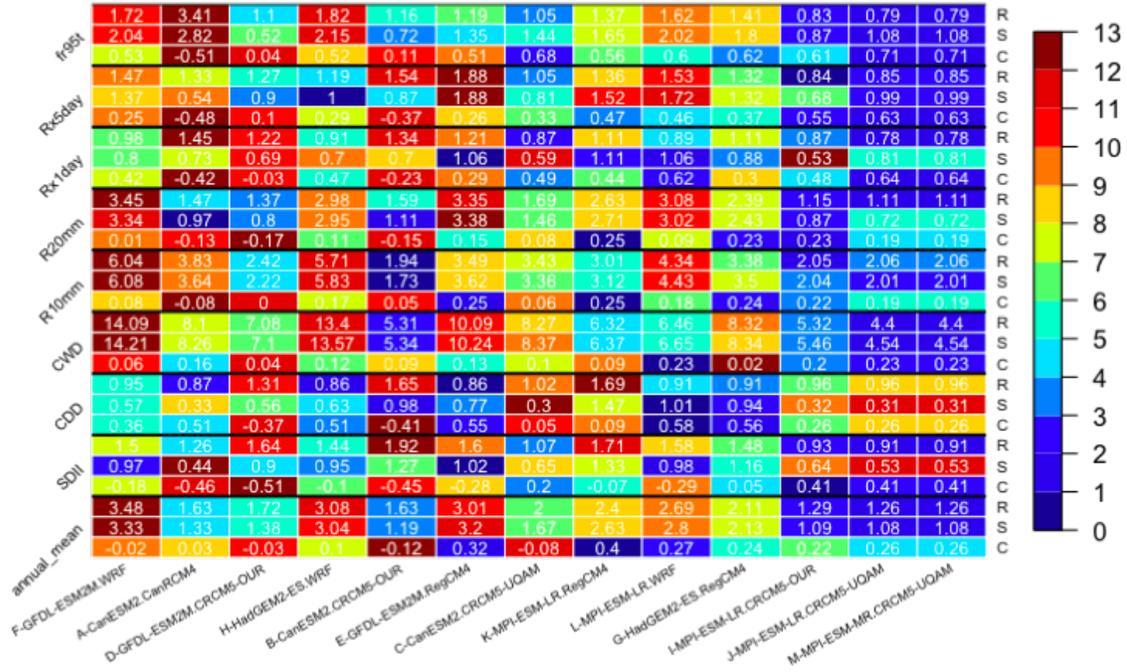
S No.	Science Activity	Lead Scientists	Description
			<p><b>Related papers:</b>  <b>C. M. Zarzycki</b> (2018), Projecting changes in societally-impactful northeastern U.S. snowstorms. <i>Geophys. Res. Lett.</i>, <b>45</b>, 12067–12075, doi:10.1029/2018GL079820.</p> <p>This study applied the TC-rainfall methodology to explore changes in TC extreme precipitation, intensity and size and their impacts for the Eastern United States under future climate warming scenarios.</p> <p><b>Stansfield, A. M., K. A. Reed and C. M. Zarzycki</b> (2019), Changes in Precipitation from North Atlantic Tropical Cyclones under RCP Scenarios in the Variable-Resolution Community Atmosphere Model, <i>Geophys. Res. Letts</i>, to be submitted.</p>
7.	Mesoscale Convective Systems (MCS) metrics modeling	Simon Wang and Binod Pokharel	Metrics were developed to track the MCS over the northeast US that covers the Susquehanna. The NARCCAP Data were also utilized.

P-F estimates in inches for 24hr precipitation | Kissimmee-SFlorida | BC-AMP



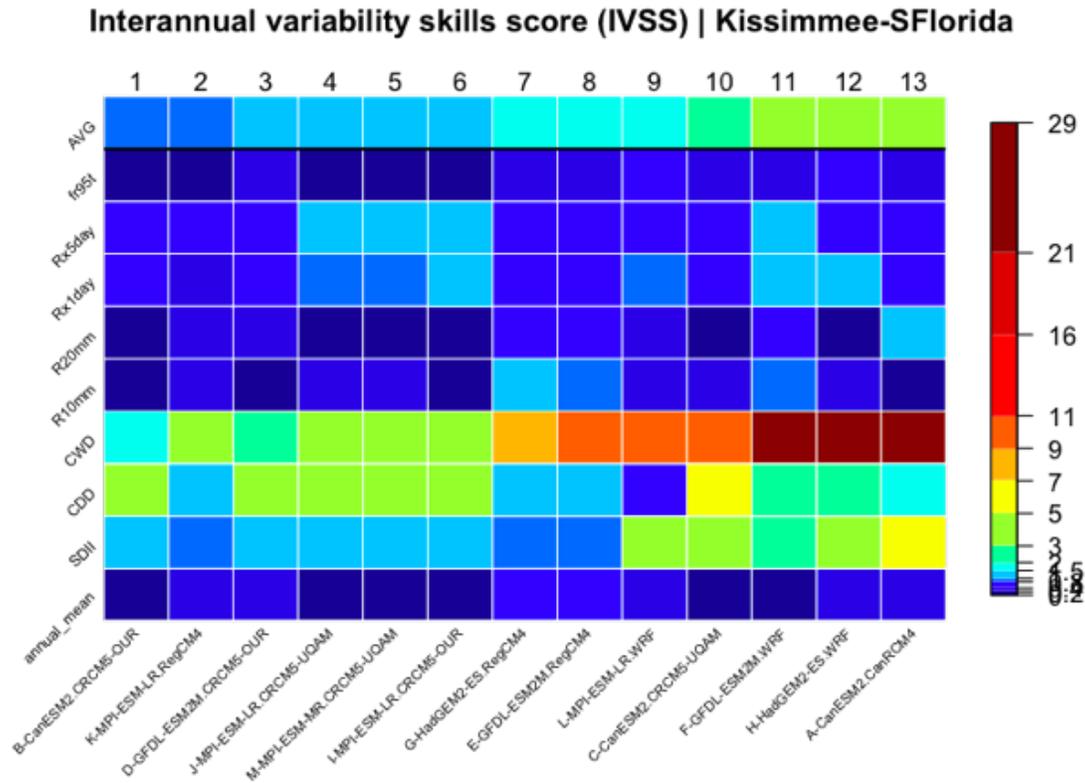
**Figure A1: 24-hr IDF estimates in the bias-corrected historical and future simulations for 74 weather stations.** IDF estimates above are from pooling better performing models (models E,I,K,M in Section 3.1). The red curve and yellow shading indicate IDF estimates and corresponding 90% confidence interval in the bias-corrected historical simulation. The blue curve and green shading indicate IDF estimates and corresponding 90% confidence interval in the bias-corrected RCP8.5 simulation. The X-axis is return periods in years and the Y-axis is intensity in inches/day.

## Ranking based upon Taylor Diagram | Kissimmee-SFlorida



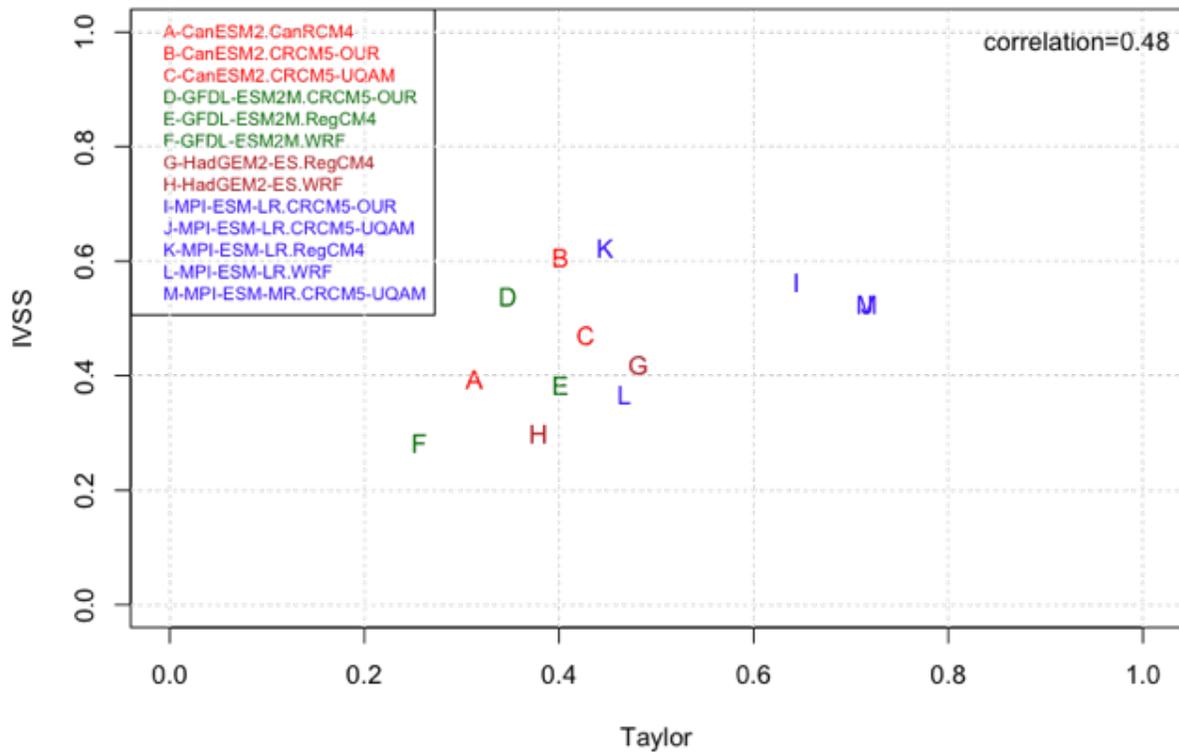
**Figure A2: Model rankings for different NA-CORDEX models and for different precipitation metrics based upon Taylor diagram.**

The diagram compares the spatial pattern of the long term means of P-indices in terms of their standard deviation, root mean square error (RMSE) and spatial correlation wrt. observation. Color bar indicates ranking of models between 1 and 13. The letters on the right vertical axis have the following meanings. C: spatial correlation between model and the observation, S: Standard deviation in model divided by the standard deviation in the observation, R: RMSE between model and the observation divided by the standard deviation in the observation. The numbers inside the boxes indicate their actual values. The figure shows that MPI set of models are among the best performers based upon Taylor diagram.



**Figure A3: Ranking of NA-CORDEX models based upon IVSS, for different precipitation metrics.** IVSS is a temporal variability skill score based upon the ratio of temporal standard deviation of the indices in model to the standard deviation of the indices in the observation at each station. Color bar indicates IVSS (score); the smaller the score the better is the model performance. The numbers on top indicate model ranking between 1 and 13 based upon IVSS. The figure shows that nearly half of the models have reasonably good (and similar) skills in simulating the temporal variability in the observation. Nearly all models have trouble in simulating CDD and CWD.

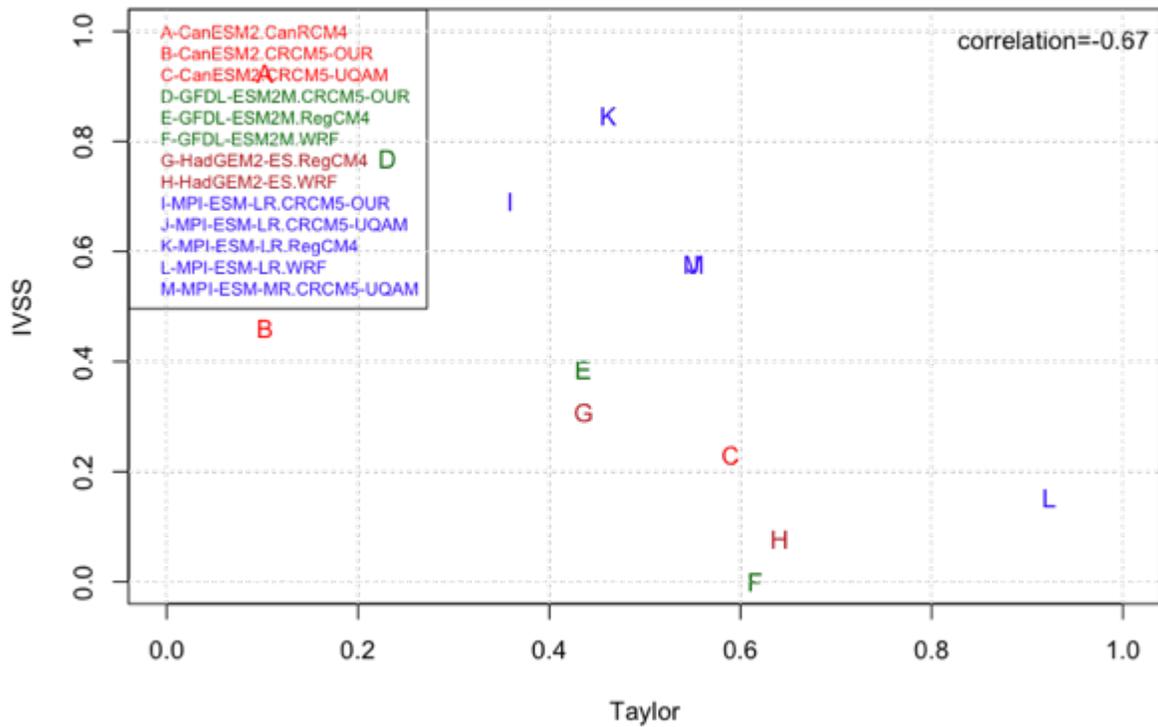
### Ranking based upon Taylor diagram and IVSS score | Kissimmee-SFlorida



**Figure A4: Overall rankings based on both Taylor diagram and IVSS for NA-CORDEX models, for various precipitation metrics**

The correlation between model ranking scores from Taylor diagram and IVSS is shown in the upper left corner. The correlation coefficient indicates the degree to which NACORDEX models performing well in spatial simulation of indices also perform well on temporal scale. The figure shows that most of the models have moderate skills in simulating the spatial and temporal patterns of the observed indices- models lie in the center of the scatter diagram.

## Ranking based upon Taylor diagram and IVSS score | AMP | Kissimmee-SFlorida



**Figure A5: Overall rankings based on both Taylor diagram and IVSS for NA-CORDEX models for annual maximum precipitation**

The correlation between model ranking scores from Taylor diagram and IVSS is shown in the upper left corner. The correlation coefficient indicates the degree to which models performing well in spatial simulation of indices also perform well on temporal scale. This figure shows that for the metric of Annual Maximum Precipitation, models that perform well in simulating the spatial pattern perform badly in capturing the temporal standard deviation of the annual maximum precipitation. This result was in contrast with results of the same analysis for Susquehanna and Sacramento/San Joaquin, where the models ranked in a more orderly fashion, and there were models that performed well on both temporal and spatial rankings. In this case, the good-performing models were ones that had moderate skill in both temporal and spatial rankings ( K, I, M, and E).

## References

<sup>1</sup>Jagannathan, K., A. D. Jones, and I. Ray, The making of a metric: Co-producing decision-relevant climate science. *Bull. Amer. Meteor. Soc.*, doi: <https://doi.org/10.1175/BAMS-D-19-0296.1>.