Aligning Climate Models With Stakeholder Needs: Advances in Communicating Future Rainfall Uncertainties for South Florida Decision Makers

Johnna M. Infanti, Ben P. Kirtman, Nicholas G. Aumen, John Stamm, and Colin Polsky

1University Corporation for Atmospheric Research, Boulder, CO, USA, 2Cooperative Institute for Marine and Atmospheric Studies, University of Miami Rosenstiel School of Marine and Atmospheric Sciences, Miami, FL, USA, 3Florida Center for Environmental Studies and Department of Geosciences, Florida Atlantic University, Boca Raton, FL, USA, 4Now at NOAA/Climate Prediction Center and Innovim, LLC, College Park, MD, USA, 5U.S. Geological Survey, Reston, VA, USA

Abstract Changes in future precipitation are of great importance to climate data users in South Florida. A recent U.S. Geological Survey workshop, “Increasing Confidence in Precipitation Projections for Everglades Restoration,” highlighted a gap between standard climate model outputs and the climate information needs of some key Florida natural resource managers. These natural resource managers (hereafter broadly defined as “climate data users”) need more tailored output than is commonly provided by the climate modeling community. This study responds to these user needs by outlining and testing an adaptable methodology to select output from ensemble climate-model simulations based on user-defined precipitation drivers, using statistical methods common across scientific disciplines. This methodology is developed to provide a “decision matrix” that guides climate data users to specify the subset of models most important to their work based on each user’s season (winter, summer, and annual) and the condition (dry, wet, neutral, and no threshold events) of interest. The decision matrix is intended to better communicate the subset of models best representing precipitation drivers. This information could increase users’ confidence in climate models as a resource for natural resource planning and can be used to direct future dynamical downscaling efforts. This methodology is based in dynamical processes controlling precipitation via remote and local teleconnections. We also suggest that future climate studies in South Florida include high-resolution climate model runs (i.e., ocean eddy resolving) in conjunction with dynamical downscaling to adequately capture precipitation variability.

1. Introduction

The Everglades in South Florida (Figure 1a) is an important natural resource that helps to balance the salinity of estuarine areas along the gulf, Florida Bay, and Atlantic coasts, provides an important habitat for birds, mammal, fishes, other fauna, contributes to the South Florida water supply, and potentially serves as an important regional carbon sink (Aumen et al., 2015, and references therein). Over the years, the Everglades have been compartmentalized and drained, and freshwater flows have been restricted and redirected, causing harm to ecosystems. State and Federal agencies have spent decades and millions of dollars devising plans to help restore some of the historical flows to mitigate the adverse effects of historical alterations, ultimately to regain the health of Everglades ecosystems. These plans were not designed with the consideration of future changes in climate (McLean et al., 2008), though the National Academy of Sciences recommended a midcourse assessment that includes future changes (National Academies of Sciences, 2018). There is now a recognition that planning of this scale must consider the potential for changes in spatial and temporal patterns of precipitation among other aspects of climate. Climate projections for South Florida, however, have been variable—some projecting wetter future conditions, some drier, and there is considerable uncertainty in which models are best predictors of future conditions (Infanti et al., 2019; Obeysekera et al., 2011). Climate data users desire increased confidence in climate projections as restoration plans are costly and can take years to implement. Increased confidence for future climate conditions is
critical to effective management of the resource. The purpose of this paper is to describe a decision matrix generated to identify a set of climate models most appropriate for use by Everglades restoration planners, as well as other natural resource planners in the study area.

The ecosystem of the Florida Everglades and the built environment of South Florida can be heavily impacted by changing precipitation (Aumen et al., 2015; Obeysekera et al., 2011). Though we now have better understanding of potential South Florida precipitation changes to 2,100, the associated uncertainties remain high (Infanti et al., 2019). Moreover, it can be difficult for non-climatologists to interpret and use large climate data sets. Absent a sustained engagement with climatologists to translate the model outputs, this difficulty often leads to the climate model output being ignored or underutilized. As an example of a user need, climate data users may desire a smaller suite of models to drive an ecological model due to storage and time constraints, but it is difficult to determine a singular subset of models that outperforms others across all seasons and parameters (Hagedorn et al., 2005). As a result of the 2017 U.S. Geological Survey-Florida Atlantic University (USGS-FAU) Workshop on Improving Confidence in Precipitation Projections for Everglades Restoration (http://www.ces.fau.edu/usgs/downscaling-2.0/), this manuscript seeks to develop a decision matrix for climate data users who wish to determine which model(s) most confidently represent precipitation variability for time periods or conditions most necessary for their planning. The purpose of this study is to describe how this decision matrix was constructed and show results and comparisons of precipitation projections from subsetted/nonsubsetted models.

Climate models, including those in the Coupled Model Intercomparison Project Version 5 (CMIP5, Taylor et al., 2011), are all formulated differently, with differences in spatial resolution, coupling methods, physics packages, and so on. All historical (i.e., 1850–2005) climate modeling experiments are forced by a single set of observed atmospheric composition changes that reflect composite anthropogenic and natural sources. Historical simulations thus form a range of potential climate variability, representing a variety of best effort attempts to simulate the climate system (Taylor et al., 2011). One of the major sources of uncertainty in climate models is the models’ abilities to simulate internal variability, that is, variability attributable to natural causes that are quasi-regular such as the El Niño–Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) (Deser et al., 2012; Enfield et al., 2001; Hawkins & Sutton, 2009, 2011). Realistic climate models should simulate these large-scale climatic patterns similarly to observations with respect to overall temporal and spatial variability; however, because historical simulations are initialized from an arbitrary point of a control run, modeled large-scale climatic patterns do not specifically match the year-to-year variations of observations (Taylor et al., 2011).
These large-scale climatic patterns drive regional precipitation in South Florida (teleconnections; Goly & Teegavarapu, 2014; Moses et al., 2013) and affect the predictability of regional variations. Key large-scale climatic patterns that affect Florida climate include ENSO (on seasonal timescales; Gershunov, 1998; Gershunov & Barnett, 1998; Ropelewski & Halpert, 1986), the Atlantic Multidecadal Oscillation (AMO, on multidecadal timescales; Enfield et al., 2001), and the Pacific Decadal Oscillation (PDO, on decadal timescales; Misra et al., 2011). Because these modes of natural variability are so important to Florida climate, particularly precipitation, it is also important for climate models to simulate linkages to large-scale modes of variability with some accuracy (e.g., Bellenger et al., 2014; Fuentes-Teegavarapu, 2014; Moses et al., 2013). Moreover, in a changing climate, these teleconnections may change or shift, causing changes to Florida precipitation (Oh et al., 2014). Thus, our model selection process is based on how well models simulate the relationship between regional precipitation and large-scale sea surface temperature (SST) and local 2-m temperature (temperature at 2 m above the ground surface) patterns, which are used as proxies for these large-scale climatic patterns and local precipitation drivers (i.e., those processes that impact precipitation variability), and selects models based on an assumption that precipitation is correctly represented for the right (dynamical) reasons. This is opposed to selecting models based on precipitation alone. While some models may appear to simulate precipitation well, the processes driving precipitation may be incorrect.

The model selection methodology presented here seeks to reduce the uncertainty surrounding natural or internal variability in climate models. This is achieved by selecting the models we are most confident in based on their representation of the natural remote and local drivers of Florida precipitation. Specific to Florida, precipitation is highly seasonal (e.g., Obeysekera et al., 2011), and remote teleconnections can also vary seasonally. For example, the influence of El Niño (La Niña) on wintertime precipitation is that the southeast is wetter (drier), but in the summer months the response is much weaker (this relationship has been studied by many authors, including Konrad, 1997; Mo & Schemm, 2008; Mo, 2010; Ropelewski & Halpert, 1986). While this manuscript does not offer a detailed review of remote drivers, the interested reader is encouraged to examine Kirtman et al. (2017) for more information. The local driver used here is 2-m temperature, which follows the typical assumption that rainfall increases soil moisture, which increases local evaporation (latent heat flux), and thus decreases air temperature (e.g., Cong & Brady, 2012). Both of these relationships are assumed to be contemporaneous.

It should be emphasized that this model subsetting approach assumes that mechanisms associated with these background states lead to greater probability of wet, dry, or neutral events. We thus select models on the basis of their representation of the understood/assumed primary physical drivers of Florida rainfall in recognition of the specific use need, and in recognition that using the full suite of model realizations is impractical for most uses. Our intention is for this decision matrix to be used such that whether an end user is interested in evaluating conditions of increased or decreased rainfall, they can identify which model output best reflects their needs. In addition, we assert that this decision matrix can be used to drive future dynamical downscaling efforts in South Florida by targeting the models that best represent the key drivers of Florida rainfall.

This manuscript details a simple methodology for model subsetting that is based on model representation of physical drivers of Florida rainfall focusing on contributions from SST and 2-m temperature. We use monthly and daily climate-model data from raw and downscaled sources. The purpose of this paper is to describe the construction of a decision matrix generated to identify a selected set of climate models most appropriate for use by Everglades restoration planners, as well as other natural resource planners in the study area. We also discuss limitations to this approach such as the large noise component in precipitation leading to uncertainty, model representation of convective precipitation, and so on. Finally, we provide guidance for future climate modeling work in South Florida.

2. Data

Climate model data used in this study are CMIP5 raw precipitation, SST, and 2-m temperature. We use the first realization of each model for this assessment (r1i1p1). These data are regridded to 1° × 1° for consistency and are primarily used for comparison between historical observations and models. Because these data are too low resolution for many users in South Florida, we also use statistically downscaled data for presentation of future precipitation changes. These data were obtained from the Bureau of Reclamation public archive of
Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections (Reclamation, 2013). A discussion of statistical downscaling is beyond the scope of this manuscript, but we encourage the interested reader to consider Wood et al. (2004) and Pierce et al. (2014). We used monthly Bias Corrected Spatially Disaggregated (BCSD) and daily Localized Constructed Analogue (LOCA) precipitation from Representative Concentration Pathway (RCP) 4.5 (moderate emissions) and RCP8.5 (high emissions) in this study. Remaining RCPs (RCP2.6 and 6.0) have higher uncertainty or unrealistic future information (Infanti et al., 2019), and we have elected not to include these here. For more information on RCPs, see van Vuuren et al. (2011). BCSD and LOCA data are available monthly and daily, respectively, which informed our choice of downscaled data for this study. Future changes are presented for the near-, middle-, and long-term periods of 2019–2045, 2045–2073, and 2074–2099, respectively. Thirty-two downscaled models are available for monthly analysis, and 15 for daily. A list of these models is included in the supporting information. Examples are shown with respect to monthly data.

Historical climate model data are compared to historical observations for 1950–2005 (monthly), and 1986–2005 (daily). Daily data have a shorter period of record due to data availability. Observation data sets considered in this study are Global Precipitation Climatology Centre Version 7 (GPCC, Schneider et al., 2016), Ziese et al., 2011, National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed Sea Surface Temperature Version 5 (ERSSTv5; Huang et al., 2017), and Global Historical Climatology Network Version 2 and the Climate Anomaly Monitoring System (GHCN-CAMS) 2-m temperature (Fan & Van den Dool, 2008). We also utilize the North American Regional Reanalysis for convective and large-scale precipitation as these data are not readily available in observations (Mesinger et al., 2006). Convective precipitation is precipitation produced by convection, or the upward movement of a local air mass that is warmer than its surroundings, which then cools and condenses as it rises. Convective precipitation is typically short lived, localized, and intense. Large-scale precipitation is related to large-scale atmospheric circulation/frontal passages and is comparatively more widespread and longer lived. Total precipitation (noted later in the manuscript) is the sum of convective and large-scale precipitation.

3. Decision Matrix Creation

To create the decision matrix, we use the following briefly outlined steps (explained in detail throughout the section with examples for monthly data). For a visual representation of Steps 1–5, please see the flow chart in the supporting information (Figure S1).

1. Linearly regress global SST anomaly (SSTA) and Southeast United States 2-m temperature onto observed filtered South Florida (region in Figure 1a) October–March (ONDJFM), April–September (AMJJAS), or annual mean wet, dry, and neutral precipitation events (defined below) (example regression equation for SSTA and precipitation: \( y_{\text{precip}} = mx_{\text{SSTA}} + b \)). We base the analysis on ONDJFM and AMJJAS given typical agricultural seasons (e.g., Keellings, 2016).

2. Follow the same approach for raw (regridded to 1° × 1°) historical climate model data, resulting in 32 patterns of modeled regressions of SSTA and 32 patterns of modeled regressions of 2-m temperature (32 due to the number of available models) for each season (ONDJFM, AMJJAS, and annual) and event (wet, dry, and neutral).

3. Spatially correlate each of the 32 SSTA regressions and thirty-two 2-m temperature regression patterns with the observed SSTA regression and 2-m temperature regression (i.e., calculate the correlation coefficient over the given latitude and longitude domain using Pearson’s \( r \), resulting in 1 correlation coefficient for each model/observation regression pattern pair), thus 32 SSTA correlations and thirty-two 2-m temperature correlation coefficients for each season (ONDJFM, AMJJAS, and annual) and event (wet, dry, and neutral).

4. Of these 32 SSTA correlation coefficients and thirty-two 2-m temperature correlation coefficients, choose the top 10. This provides the 10 models that best estimate the regression pattern of SSTA and the regression pattern of 2-m temperature.

5. Because both SSTA and 2-m temperature are key drivers of Florida precipitation, determine the models that appear on both top 10 lists. These models represent those that best represent the observed regression pattern of both SSTA and 2-m temperature. These models are those that fill out the decision matrix (Table 1).
Table 1

| Decision Matrix Based on Regression of Sea Surface Temperature Anomalies and 2-m Temperature Regressed Onto South Florida Precipitation |
|-------------------------------------------------|
| ONDJFM       | AMIJJAS       | Annual      |
| Dry events (SPI less than −0.5)                  | BNU-ESM       | FGOALS-g2   |
|             | CCSM4         | GFDL-ESM 2g |
|             | GFDL-CM3      | HadGEM2-ES  |
|             | GISS-e2-h-cc  | MPI-ESM-LR  |
|             | HadGEM2-es    | inmcm4      |
|             | BCC-CSM1-1    |             |
| Wet Events (SPI >0.5)                            | ACCESS1-3     | GFDL-CM3    |
|             | CCSM4         | HadGEM2-CC  |
|             | CESM1-BGC     | BCC-CSM1-1-m|
|             | CanESM2       | BCC-CSM1-1  |
|             | GFDL-CM3      | inmcm4      |
|             | GFDL-ESM 2m   |             |
| Neutral Events (SPI greater than −0.5, SPI <0.5) | CMCC-CM       | CSM4        |
|             | MPI-ESM-LR    | CMCC-CM     |
|             | GISS-e2-r     | GISS-e2-r   |
|             | BCC-CSM1-1-m  | BCC-CSM1-1-m|
| Overall (no threshold)                           | ACCESS1-3     | CSM4        |
|             | BNU-ESM       | CESM1-CAM5  |
|             | CCSM4         | GISS-e2-r   |
|             | CNRM-CM5      | INMCM4      |
|             | GFDL-CM3      |             |

Note. Decision matrix is based on monthly 1° × 1° raw Coupled Model Intercomparison Project Version 5 (CMIP5) data and applied to Bias Corrected Spatially Disaggregated (BCSD) downscaled data. Model selection is outlined in section 3, Steps 1 through 6. The matrix shows winter (ONDJFM), summer (AMIJJAS), and annual mean model selections for dry, wet, neutral, and no threshold events. An example of usage is a user most interested in what models best represent winter-time wet events would use the six models noted in Column 2, Row 3. More information on models can be found in the supporting information and BCSD data can be accessed online (ftp://gdo-dcp.ucllnl.org/pub/dcp/archive/cmip5/bc). Model names in matrix directly correspond to available data downloads. SPI = Standardized Precipitation Index. ONDJFM = October–March. AMIJJAS = April–September.

6. Once these sets of models are selected from historical model data, use the smaller subsets to calculate precipitation changes in downscaled RCP4.5 and RCP8.5. For the purposes of demonstrating the subsetting methodology, show how these compare to the full set of models.

As an example of our subsetting methodology, we show results for ONDJFM wet events. For the purpose of this paper, “wet events” are defined as a value of greater than 0.5 on the 6-month Standardized Precipitation Index (SPI6) (Figure 1b, with precipitation composite shown in Figure 1c). SPI is a measure of precipitation based on a gamma distribution and can be used to determine wetter or drier than normal (e.g., Hayes et al., 1999). “Dry events” are defined as a value of less than 0.5 on the SPI6. Our threshold for wet and dry events is conceptually similar to precipitation ±0.5 standard deviations from the mean. In order to calculate SPI6, precipitation data are spatially aggregated in the South Florida region (land points in Figure 1a) using latitude weighting.

Figures 1d and 1e show results from Step 1 above, regressions of observed wet ONDJFM South Florida precipitation onto SSTA and 2-m temperature, respectively. Figure 1d shows that for ONDJFM wet events there is a strong signal in the tropical Pacific that can be attributed to ENSO, where increasing precipitation is associated with increasing SSTA (red shading). There is also a strong negative relationship with extratropical Pacific SSTA that can be associated with the PDO (blue shading, indicating that as precipitation increases, SSTA decreases, and vice versa; e.g., Kurtzman & Scanlon, 2007). Finally, there is also a negative relationship in the Gulf Stream region, which is a potential player in Florida rainfall (i.e., Infanti & Kirtman, 2018; Siqueira & Kirtman, 2016). The 2-m temperature anomalies are overall negative on the order of the extratropical SSTA regression amplitude (Figure 1e). During El Niño, we expect the southeast to be cool and wet, so these patterns are expected.

The regression between raw CMIP5 historical precipitation and SSTA/2-m temperature can differ from the observed relationships. While we do not depict each model regression of SSTA and 2-m temperature

INFANTI ET AL.
patterns, we note that there is a large spread in the patterns across models. Figures 2a and 2c show the average regression pattern resulting from regressing ONDJFM SSTA and 2-m temperature onto wet precipitation events for all models (remaining seasons and events are depicted in the supporting information for nonsubsetted and subsetted). Figures 2b and 2d show the average regression pattern resulting from regressing ONDJFM SSTA and 2-m temperature onto wet precipitation events for

**Figure 2.** October–March (ONDJFM) full and subsetted multimodel mean regression of sea surface temperature anomalies (SSTA) and 2-m temperature onto South Florida wet precipitation events. Units in °C/standard deviation. (a) SSTA regression for full suite of models. (b) SSTA regression for subsetted models. (c and d) As in (a) and (b) but for 2-m temperature regression.

| Table 2 | As in Table 1, but for Daily LOCA Data |
|---------|----------------------------------------|
|         | ONDJFM                                | AMJAS        | Annual          |
| Dry Events (SPI less than −0.5) | BNU-ESM | CESM1-BGC | ACCESS1-0 |
|         | CNRM-CM5 | CNRM-CM5 | CCSM4 | |
|         | MPI-ESM-MR | CanESM2 | CNRM-CM5 | |
|         | MRI-CGCM3 | MIROC-ESM-CHEM | MRI-CGCM3 | |
| Wet Events (SPI >0.5) | CSSM4 | BNU-ESM | IPSL-CM5A-LR |
|         | CESM1-BGC | CSSM4 | IPSL-CM5A-MR | |
|         | IPSL-CM5A-LR | CanESM2 | IPSL-CM5A-MR | |
|         | IPSL-CM5A-MR | MIROC-ESM-CHEM | MRI-CGCM3 | |
| Neutral Events (SPI greater than −0.5, SPI <0.5) | CanESM2 | ACCESS1-0 | ACCESS1-0 |
|         | GFDL-CM3 | CESM1-BGC | CanESM2 | |
|         | GFDL-CM3 | GFDL-CM3 | MIROC-ESM | |
|         | MPI-ESM-LR | MIROC-ESM | MRI-CGCM3 | |
| Overall (no threshold) | ACCESS1-0 | CNRM-CM5 | ACCESS1-0 |
|         | BNU-ESM | MIROC-ESM-CHEM | BNU-ESM | |
|         | CCSM4 | CNRM-CM5 | CCSM4 | |
|         | CNRM-CM5 | IPSL-CM5A-LR | CanESM2 | |
|         | IPSL-CM5A-LR | MIROC5 | IPSL-CM5A-LR | |
|         | IPSL-CM5A-MR | MIROC5 | IPSL-CM5A-MR | |
|         | MIROC5 | MPI-ESM-LR | MIROC5 | |
subsetted models (figure corresponds to Step 2, above). We do not expect the model regression patterns to perfectly match observations as models still have uncertainties related to representation of natural variability; however, we note that subsetted models show a better match with the observed patterns versus considering the full suite of models. For example, the Gulf Stream region in Figure 2a (nonsubsetted models) shows a weakly positive regression, while observations show a strong negative regression (Figure 1c). The negative Gulf Stream regression pattern is seen in the subsetted models (Figure 2b). For ONDJFM, 2-m temperature did not show much impact in overall structure from subsetting, but the changes are much stronger (Figures 2c and 2d). The SSTA regression pattern is stronger in the winter months, while the 2-m temperature regression pattern is stronger in the summer months (not shown, see supporting information).

Steps 3 and 4 are not explicitly shown as they result in numerical lists of correlations. This approach is then applied to AMJJAS, the annual mean, dry events (SPI less than –0.5), neutral events (SPI greater than –0.5 and less than 0.5), and overall (i.e., no threshold) precipitation. The full decision matrix for monthly data with results from each season and parameter is shown in Table 1 (corresponding to Step 5).

We follow a similar methodology for daily precipitation events, using 1986–2005 as a reference period due to data size. LOCA data have fewer models available (15; list of models in the supporting information).
to compute the regressions, we aggregate daily data to 6-month means and perform a similar analysis to the above. Overall, regression results are similar to monthly mean data, and the subsetted daily model list is shown in Table 2. We include both daily and monthly data to incorporate all user needs.

Simulated spatial distribution of the change in precipitation (in percent and mm/day) in near-, middle-, and long-term periods for nonsubsetted (Figures 3 and 5) and subsetted (Figures 4 and 6) simulations based on the above approach are calculated. This corresponds to Step 6. We include changes in RCP4.5 (Figures 3 and 5) and 8.5 (Figures 4 and 6). Though we do not specifically discuss each figure, we have included both RCPs and full/subsetted suites of models for completeness. This matrix should be read as, for example, ONDJFM dry events are projected to get wetter (green shading/increased precipitation), as are ONDJFM wet events (green shading/increased precipitation). Similar results from the full and subsetted suite of models point to general agreement across models with respect to physical drivers of precipitation changes. Projected changes that are largely dissimilar between these sets indicate a less robust signal across the models. In practice, this may reduce user confidence in the models. Showing results by category, for example, the wet versus dry events, enables users to identify cases when models tend to show relatively high levels of agreement. While overall we find similar patterns of change in ONDJFM wet, dry, and no threshold between nonsubsetted and subsetted (Figure 3 compared to Figure 4, ONDJFM column), there are large differences between results for neutral events in all seasons (Figure 3 compared to Figure 4, neutral row). The neutral nonsubsetted results mainly show increased precipitation and the subsetted suite shows a decrease. This difference is likely because neutral events are difficult to predict/project (Van Den Dool & Toth, 1991), as they are not as associated with large-scale drivers and are overall noisier than the extremes.

### Table 2: Regression Results

| Region    | ONDJFM | AMJJAS | Annual |
|-----------|--------|--------|--------|
|           | NT     | MT     | LT     | NT     | MT     | LT     |
| Dry Events (SPI < -0.5) | ![Graph](image1.png) | ![Graph](image2.png) | ![Graph](image3.png) | ![Graph](image4.png) | ![Graph](image5.png) |
| Wet Events (SPI > 0.5)  | ![Graph](image6.png) | ![Graph](image7.png) | ![Graph](image8.png) | ![Graph](image9.png) | ![Graph](image10.png) |
| Neutral Events (SPI < -0.5, SPI > 0.5) | ![Graph](image11.png) | ![Graph](image12.png) | ![Graph](image13.png) | ![Graph](image14.png) | ![Graph](image15.png) |
| No Threshold           | ![Graph](image16.png) | ![Graph](image17.png) | ![Graph](image18.png) | ![Graph](image19.png) | ![Graph](image20.png) |

![Figure 4](image.png)

As in Figure 3, but for subsetted bias-corrected spatially disaggregated (BCSD) data.
In contrast, AMJJAS shows an overall dry-to-wet pattern of spatial change (dry in the southern part of the domain and wet in the north) in both non-subsetted and subsetted (Figure 3 compared to Figure 4, AMJJAS column). The AMJJAS change is weak (compared to annual mean and NDJFM changes) and displays large variations across models, particularly in the location of the shift from dry to wet. This change is discussed in more detail in Infanti et al. (2019). The leading hypothesis for this pattern of change is that inter-ocean temperature differences cause Caribbean drying (Lee et al., 2010; Rauscher et al., 2010). This pattern of change also aligns with the 2-m temperature pattern shown in the second empirical orthogonal function (EOF2) of 2-m temperature regressed onto South Florida precipitation (discussed further below). We also note that the no-threshold changes in AMJJAS are overshadowed by ONDJFM; thus, the annual mean projected change is wet. These projected changes in the non-subsetted and in the subsetted suite agree with those in the Intergovernmental Panel on Climate Change (IPCC) report and Infanti et al. (2019).

Though this manuscript mainly focuses on RCP4.5 results, we include RCP8.5 in Figures 5 and 6 for completeness. The key difference between RCP4.5 and RCP8.5 is stronger emissions in RCP8.5. We expect that overall results will be similar until the long-term period, as RCP differences more strongly impact changes toward the end of the period, though short- and middle-term results can differ due to differences in other trace gases (see Infanti et al., 2019, for results over Florida and references therein for more information). We encourage the user to determine which emissions scenario best meets their needs and use the subsetted results of this manuscript accordingly.

### Table 1

|                | ONDJFM | AMJJAS | Annual |
|----------------|--------|--------|--------|
|                | NT | MT | LT | NT | MT | LT | NT | MT | LT |
| **Dry Events** |     |     |     |     |     |     |     |     |     |
| (SPI < 0.5)    | ![Map](image1) | ![Map](image2) | ![Map](image3) | ![Map](image4) | ![Map](image5) | ![Map](image6) | ![Map](image7) | ![Map](image8) | ![Map](image9) |
| **Wet Events** |     |     |     |     |     |     |     |     |     |
| (SPI > 0.5)    | ![Map](image1) | ![Map](image2) | ![Map](image3) | ![Map](image4) | ![Map](image5) | ![Map](image6) | ![Map](image7) | ![Map](image8) | ![Map](image9) |
| **Neutral Events** |     |     |     |     |     |     |     |     |     |
| (SPI < -0.5, SPI > 0.5) | ![Map](image1) | ![Map](image2) | ![Map](image3) | ![Map](image4) | ![Map](image5) | ![Map](image6) | ![Map](image7) | ![Map](image8) | ![Map](image9) |
| **No Threshold** |     |     |     |     |     |     |     |     |     |

**Figure 5.** As in Figure 3 but for Representative Concentration Pathway 8.5 (RCP8.5).
In addition to suggesting use of a tailored subset of all existing models, we further suggest using more than a single relevant/chosen model from the subset to capture uncertainty due to model formulation. There may be concern when considering a smaller subset of models that the spread of the change is not fully captured. We demonstrate that the nonsubsetted spread is similar to the subsetted spread in Figure 7. This figure shows the multimodel mean (dark blue line) and the minimum and maximum over the models (blue shading) from 2019–2045 (the near-term period), as well as the spread (calculated and noted on the figure) for the near-, middle-, and long-term periods for ONDJFM. The spread (calculated) is similar in all periods but is most similar in the near-term when uncertainty is mainly stemming from natural variability (result is similar for RCP4.5 and 8.5). Thus, despite using a smaller set of models, we still capture the spread and we are not overconfident in the subsetted results.

The above demonstrates the use of large-scale and regional drivers of precipitation (SSTA and 2-m temperature) to determine a subset of models that more confidently represent the given precipitation parameters by also representing these important drivers of precipitation. However, we again note that this study only explores the variability associated with SSTA and 2-m temperature and many other factors may substantially affect precipitation, which can lead to uncertainty.

4. Understanding the Uncertainty

While SSTA and 2-m temperature are important drivers of Florida precipitation, there are facets of precipitation that make the variable difficult to represent in climate models, such as convective-scale precipitation...
Figure 7. (a) Representative Concentration Pathway 4.5 (RCP4.5) change in precipitation calculated with the 6 month Standardized Precipitation Index (SPI6, unitless) for all models. Blue line indicates multimodel mean change; blue shading is the minimum and maximum of all realizations (LT). (b) As in (a) but for RCP8.5. (c) As in (a) but for subsetted models. (d) As in (b) but for subsetted models. Spread for the NT, MT, and LT is noted on the figure panels.

Figure 8. Signal to total ratio of precipitation in bias corrected spatially disaggregated (BCSD) downscaled models. Calculation is based on standard deviations (unitless) (a and b) October–March (ONDJFM), where (b) shows specifically South Florida; (c and d) as in (a) and (b) but for April–September (AMJJAS).
(due to limited climate model resolution), sea breezes leading to enhanced precipitation, and land surface interaction (Obeysekera et al., 2011). The above subsetting approach is quantitative, grounded in observational comparisons, and transferrable to other models or regions. However, once implemented, robust estimates and understanding of uncertainty are needed. Here we provide some assessment of the uncertainty.

Figure 8 shows the signal-to-total standard deviation ratio of BCSD precipitation for ONDJFM and AMJJAS (see Infanti & Kirtman, 2016; Rowell, 1998; Schubert et al., 2002). Higher values (over 0.5) indicate dominance of signal variance, and smaller values (under 0.5) indicate signal variance is a small contributor to total variance—that is, that noise is dominant. Signal variance is generally associated with large-scale teleconnections, where noise is the unpredictable part of precipitation. For the sake of completeness and comparison, we include the Southeast United States and South Florida. In a relative sense, ONDJFM has higher signal variance dominance than AMJJAS (compare Figures 8b and 8d), particularly over South Florida, though there is a region of high signal variance east of Lake Okeechobee in both seasons. However, as the values never reach over 0.5, it is clear that the signal variance is a small contributor to total variance overall. This association shows that while some of the precipitation is explained by the large-scale drivers, there is still a large portion that is unpredictable.

We examine the observed relationship between ONDJFM and AMJJAS SSTA and 2-m temperature and South Florida precipitation more closely in Figure 9. The goal of this figure and discussion is to provide context on the chosen drivers of SSTA and 2-m temperature, and how much these drivers relate to South Florida precipitation. Though this information may not be directly used by a practitioner, we aim to provide the scientific basis for why the subsetting can be based on these two chosen drivers.
We compute the first, second, and third EOF of South Florida precipitation and regress SSTA and 2-m temperature onto the resulting principal component (PC) time series in Figure 9. For the regression pattern associated with the first EOF, about 75% of precipitation variance is associated with this large-scale SSTA pattern and 2-m temperature pattern in ONDJFM (Figure 9a). This result is in contrast to AMJJAS where a smaller amount of precipitation explained variance is associated with a weak SSTA pattern (Figure 9d), and comparatively, the 2-m temperature regression is stronger (Figure 9b). This pattern supports the assumption that SSTA is a larger driver of precipitation in winter months due to ENSO. The 2-m temperature is still a strong driver of ONDJFM precipitation in EOFs 2 and 3 (comparatively stronger than SSTA), which supports using both variables in our above assessment.

Also of note is the 2-m temperature pattern associated with AMJJAS EOF2 precipitation (Figure 9e). In general, in climate models, temperature increases equate to precipitation increases and vice versa, as 2-m temperature reflects cloudiness. This pattern is similar to the spatial pattern of simulated projected precipitation change seen in AMJJAS (e.g., Figure 3, AMJJAS column). In addition, the EOF2 precipitation pattern is similar to the pattern of change seen in AMJJAS (precipitation EOF2 not shown, changes shown in Figures 3–6). Though we can clearly see these ties to large-scale drivers, as noted in Figure 8, there is still significant precipitation variability that is not related to these large-scale drivers and may ultimately be unpredictable. This pattern of change is associated with temperature differences and projected results indicate that this spatial pattern will become more common in the future.

A final facet of precipitation that we examine is the existence of convective-scale precipitation and its importance in South Florida. Convective-scale precipitation is a limitation in climate models and in this study. Most climate models do not explicitly resolve convective-scale precipitation—its effect on the large-scale circulation is estimated or parameterized based on empirical constraints and large-scale variables. Climate models with higher horizontal resolution simulate total precipitation more accurately than convective precipitation (O’Brien et al., 2016; Shields et al., 2016; Wehner et al., 2014), and dynamical downscaling is generally needed to resolve convection more accurately (Wood et al., 2004). Statistical downscaling, while a pragmatic choice for downscaling due to the lack of computing power needed to achieve it, does not specifically target convective-scale precipitation variability (Wood et al., 2004). It is important to note the amount of convective-scale precipitation that exists in South Florida, as it is unlikely most climate models will capture this source of precipitation variability.

Figure 10. Observed (from reanalysis) variance ratios of convective precipitation to total precipitation, and large-scale precipitation to total precipitation for the Southeast United States (a, b, e, and f); and South Florida (c, d, g, and h) (unitless). (a) October–March (ONDJFM) ratio of convective/total precipitation. (b) ONDJFM ratio of large-scale/total precipitation. (c and d) As in (a) and (b) but for South Florida specifically. (e–h) As in (a)–(d) but for Apr–Sep (AMJJAS).
In order to demonstrate this variability, we calculate the variance ratio of convective to total precipitation, and of large-scale to total precipitation for ONDJFM and AMJJAS (estimated using reanalysis data) in Figure 10. Values below 0.5 indicate that less than half of total precipitation is controlled by convective/large-scale (depending on panel), and vice versa. For example, much of the interior southeast precipitation variance is controlled by large-scale precipitation rather than convective-scale in ONDJFM, as is expected due to ENSO teleconnections, whereas South Florida shows larger influence of convective-scale precipitation that is mainly tied to sea breezes (Obeysekera et al., 2011, 2015) (Figures 10a–10d). In AMJJAS, most of the Southeast United States coastal regions are dominated by convective-scale precipitation, and large-scale precipitation is only weakly contributing (Figures 10e–10g). This figure demonstrates that convective-scale precipitation is an important contributor to precipitation variance, though we assert that given results from the above discussion, large-scale and local drivers are still important enough contributors to use as variables in subsetting. We also note that improvements in representing horizontal ocean and atmosphere convection can lead to significant improvements in large-scale total precipitation prediction skill, depending on region considered, but this improvement would be mainly due to increasing ocean resolution to resolve Gulf stream eddy variability (Infanti & Kirtman, 2018). It is hypothesized that a combination of high-resolution (ocean eddy resolving, or 0.1° ocean, 0.5° atmosphere; see, e.g., Kirtman et al., 2012; Siqueira & Kirtman, 2016) climate models and dynamical downscaling will serve to better capture the total variability of precipitation.

5. Concluding Remarks

This effort aims to provide a simple-to-use and understandable methodology for subsetting CMIP5 climate models for use by climate data users in South Florida who desire a tailored and smaller set of model output than is typically provided by climatologists. Our approach emphasizes tailoring based on model confidence. This approach seeks to select models that simulate observed Florida precipitation comparatively well, in addition to representing two sources of precipitation variability—global SST and regional 2-m temperature patterns—allowing for greater confidence that this subset is capturing the precipitation for the right reasons. Because this subsetting approach is based on widely available data (SST, 2-m temperature, and precipitation) and methods that are easily translatable across fields (regressions, correlations, etc.), it is both easily performed and understood by a variety of practitioners and users, as well as updated given additional or updated models. Advances such as the subsetting approach shown here should lead to improvements in how the climate modeling community can serve natural resource managers who use climate data as one among many inputs in their models. Finally, this study provides a matrix of changes in future South Florida precipitation for summer and winter seasons and the annual mean, dry events, wet events, neutral events, and overall changes (no threshold) for the full set of models and RCPs 4.5 and 8.5. We envision that these matrices can be used as guidance for decision makers and users desiring more information on future changes in the parameters of precipitation. We do not provide a suggestion on using RCP4.5 or 8.5 as we believe this decision should be based on user needs.

Our methodology for subsetting involves using the large-scale (SSTA) and regional (2-m temperature) drivers of precipitation as tools for subsetting. Only the models that best represent the observed regression patterns of SSTA and 2-m temperature onto South Florida precipitation are included in the subset. The list of subsetted models for BCSD and LOCA data are included in Tables 1 and 2. Though we provided two to six models in each cell of the matrix, this does not mean that we lack confidence in the other models, rather if we had to choose only a subset, these are our suggested models. Given that SSTA, 2-m temperature, and precipitation are widely available variables in climate models and are generally included in climate predictions, projections, and downscaled data sets, this methodology can be easily adapted to include additional models or, for example, CMIP6 data when released. We note that while the methodology can be easily adapted, improvements to climate models, shifting the seasons considered, and so on, will change the list of subsetted models.

The main conclusion of this manuscript is that this methodology involving the large-scale and regional drivers of precipitation is effective for South Florida, but there are a number of results surrounding changing precipitation in South Florida that are also discussed. The pattern of projected change in the subsetted models agrees broadly with results from other studies, including Infanti et al. (2019), indicating that subsetting
does not significantly alter the expected results. For example, ONDJFM changes are overall wet, though ONDJFM neutral events show differences between subsetted and nonsubsetting models owing to the difficulty in climate model representation of neutral events. We also find a dry-to-wet pattern of spatial change in AMJJAS agreeing with prior results, though annual mean change is overall wet due to the larger winter change. The AMJJAS precipitation change is associated with a temperature pattern that is projected to become more common in the future, while the ONDJFM change is associated with the expected dynamical response to a warming atmosphere, where a warmer atmosphere holds more water, and also indicates that a more El Niño-like SST pattern may become more common in a warming climate. For more detailed information on projected change in South Florida, the reader is encouraged to see Infanti et al. (2019). The manuscript also details the scientific background of Florida precipitation drivers and impacts to the decision matrix creation. EOF analysis shows the percentage of precipitation variance that is explained by either SSTA or 2 m temperature, and we find that SSTA and 2 m temperature explain about 75% of precipitation variance in the winter, whereas the 2 m temperature relationship is stronger in summer.

Though large-scale drivers can explain some precipitation variance, there is still a noise variance component thus a portion of variance that is uncertain. We note the limitations to predictability such as the large amount of noise involved in precipitation and the dominance of convective-scale precipitation in South Florida. There is a definite contribution to precipitation variance due to large-scale and regional drivers; however, limitations will always exist in low-resolution model runs and statistically downscaled data particularly related to convective-scale precipitation. In combination with results from Kirtman et al. (2012), Siqueira and Kirtman (2016), Laurindo et al. (2019), and Infanti and Kirtman (2018) suggesting that increasing horizontal ocean and atmosphere resolution significantly impacts the predictability of precipitation, dynamical downscaling must be used in addition to increased horizontal resolution to adequately capture all facets of precipitation. An added use for the subsetted models is that they can be used as guidance for future modeling studies and for dynamical downscaling as one must choose the combination of driving and regional models used. As an example, we suggest increasing ocean resolution to 0.1° latitude × 0.1° longitude (eddy resolving) and increasing the atmosphere resolution to at least 0.25° latitude by 0.25° longitude in order to aid with representation of large-scale teleconnections. This increase in resolution is within bounds of current computational capacity. We also suggest dynamical downscaling of this high-resolution data for areas with large dominance of convective-scale precipitation, such as South Florida.

Our suggestion for those interested in utilizing the subset of models is as follows. We show that the spread of models is similar when considering a subset of models versus the full set. We strongly suggest that users consider all models in the subset in order to sample uncertainty. Use of one realization alone will not capture the uncertainty of future climate. These subsetted models are provided to be used as guidance but selecting which set of models may be best for individual user needs (e.g., a set of models best suited for winter season wet events) is up to the discretion of the user.

We acknowledge that model culling or subsetting is controversial and challenging. In initiating this work, it was our hypothesis that the subsetting would be straightforward and less controversial because of the singular focus—South Florida rainfall. Moreover, the approach taken here is based on known large-scale climate drivers that are well established. Nevertheless, substantial challenges remain. More objective criteria such as soil moisture, AMO, and PDO should be considered in the future. Though the methodology presented has some limitations, we aim to provide an easy to follow and adaptable methodology for subsetting. This study provides methodology, analysis, data, and suggestions for future climate modeling work in South Florida. We envision that the information provided in this manuscript can be of use by climate data users in South Florida, particularly the Everglades restoration effort, and in addition by the climate science community for future work.

Data Availability Statement
Climate model data are provided by the Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections archive (http://gdo-dcp.ucarlrl.org/downscaled_cmip_projections/). The free software NCAR Command Language was used to create the plots and analyze data.
Acknowledgments
The authors acknowledge funding from the Postdocs Applying Climate Expertise (PACE) Fellowship, administered by the University Corporation for Atmospheric Research (UCAR), and from the U. S. Geological Survey’s (USGS) Greater Everglades Priority Ecosystem Sciences (GEPES) program. We thank two anonymous reviewers and USGS internal reviewers; their comments greatly improved the manuscript. We acknowledge the support of University of Miami (UM), Florida Atlantic University (FAU), USGS, and the National Oceanic and Atmospheric Administration (NOAA). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References
Aumen, N. G., Havens, K. E., Best, G. R., & Berry, L. (2015). Predicting ecological responses of the Florida Everglades to possible future climate scenarios: Introduction. *Environmental Management, 55*, 741–748. https://doi.org/10.1007/s00267-014-0439-z
Bellenger, H., Guilyardi, E., Leloup, J., Lenaigiane, M., & Viallard, J. (2014). ENSO representation in climate models: From CMIP3 to CMIP5. *Climate Dynamics, 42*(7–8), 1999–2018. https://doi.org/10.1007/s00382-013-1785-3
Cong, R. G., & Brady, M. (2012). The interdecadal dependence between rainfall and temperature: Copula analyses. *The Scientific World Journal, 2012*, 405675. https://doi.org/10.1100/2012/405675
Deser, C., Phillips, A., Bourdette, V., & Tong, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics, 38*, 527–546. https://doi.org/10.1007/s00382-010-0977-x
Enfield, D. B., Mestas-Nuñez, A. M., & Trimble, P. J. (2001). The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental US. *Geophysical Research Letters, 28*(10), 2077–2080. https://doi.org/10.1029/2000gl012745
Fan, Y., & Van den Dool, H. (2008). A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research, 113*, D01103. https://doi.org/10.1029/2007JD008870
Fuentes-Franco, R., Giorgi, F., Coppola, E., & Kucharski, F. (2016). The role of ENSO and PDO in variability of winter precipitation over North America from twenty first century CMIP5 projections. *Climate Dynamics, 46*, 3259–3277. https://doi.org/10.1007/s00382-015-2767-y
Gershunov, A. (1998). ENSO influence on intraseasonal extreme rainfall and temperature fin the contiguous United States: Implications for long-range predictability. *Journal of Climate, 11*, 3192–3203. https://doi.org/10.1175/1520-0442(1998)011<3192:EIIOTER>2.0.CO;2
Gershunov, A., & Barnett, T. P. (1998). Interdecadal Modulation of ENSO Teleconnections. *Bulletin of the American Meteorological Society, 79*, 2715–2725. https://doi.org/10.1175/1520-0479(1998)7927<2715:IMOTSE>2.0.CO;2
Goly, A., & Teegavarapu, R. S. V. (2014). Individual and coupled influences of AMO and ENSO on regional precipitation characteristics and extremes. *Water Resources Research, 50*, 4686–4709. https://doi.org/10.1002/2013wr014540
Hagedorn, R., Doblas-Reyes, F. J., & Palmer, T. N. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus A, 57*, 219–233. https://doi.org/10.1111/j.1600-0870.2005.00103.x
Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society, 90*, 1095–1107. https://doi.org/10.1175/2009BAMS2607.1
Hawkins, E., & Sutton, R. (2011). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics, 37*, 407–418. https://doi.org/10.1007/s00382-010-0810-6
Hayes, M. J., Svoboda, M. D., Willhite, D. A., & Vanyarzko, O. V. (1999). Monitoring the 1996 drought using the Standardized Precipitation Index. *Bulletin of the American Meteorological Society, 80*, 429–438. https://doi.org/10.1175/1520-0479(1999)080<0429:MSDUSD>2.0.CO;2
Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., et al. (2017). Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *Journal of Climate, 30*(20), 8179–8205. https://doi.org/10.1175/JCLI-D-16-0836.1
Infanti, J. M., & Kirtman, B. P. (2016). Prediction and predictability of land and atmosphere initialized CCSM4 climate forecasts over North America. *Journal of Geophysical Research: Atmospheres, 121*, 12690–12701. https://doi.org/10.1002/2016JD024932
Infanti, J. M., & Kirtman, B. P. (2018). A comparison of CCSM4 high-resolution and low-resolution predictions for south Florida and southeast United States drought. *Climate Dynamics*. https://doi.org/10.1007/s00382-018-4553-0
Infanti, J. M., Kirtman, B. P., Aumen, N. G., Stamm, J., & Polsky, C. (2019). Assessment of uncertainty in multi-model means of downscaled South Florida precipitation for projected (2019–2099) climate. *International Journal of Climatology, 40*(5), 2764–2777. https://doi.org/10.1002/joc.6365
Keehlings, D. (2016). Evaluation of downscaled CMIP5 model skill in simulating daily maximum temperature over the southeastern United States. *International Journal of Climatology, 36*(12), 4172–4180. https://doi.org/10.1002/joc.4612
Kirtman, B., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., et al. (2012). Impact of ocean model resolution on CCSM climate simulations. *Climate Dynamics, 39*(6), 1303–1328. https://doi.org/10.1007/s00382-012-1500-3
Kirtman, B., Misra, V., Burgman, R., Infanti, J., & Obeysekera, J. (2017). Florida Climate Variability and Prediction. In E. P. Chassignet, J. Fuentes, M., 2015). *The Fourth Decadal Oscillation impacts on precipitation in the southern and central United States: Evaluation of spatial distribution and predictions. Water Resources Research, 43*, W10427. https://doi.org/10.1029/2007WR005863
Laurindo, L. C., Siqueira, L., Mariano, A. J., & Kirtman, B. P. (2019). Cross-spectral analysis of the SST/10-m wind speed coupling resolved by satellite products and climate model simulations. *Climate Dynamics, 52*, 5073–5098. https://doi.org/10.1007/s00382-018-4434-6
Lee, L. K., Enfield, D. B., & Wang, C. (2010). Future impact of differential interbasin ocean warming on Atlantic Hurricanes. *Journal of Climate, 24*, 1264–1275. https://doi.org/10.1175/2010JCLI3884.1
McLean, A. R., Ogden, J. C., & Williams, E. E. (2008). Comprehensive Everglades restoration plan
Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafston, P. C., Ebisuzaki, W., et al. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society, 87*(3), 343–360. https://doi.org/10.1175/BAMS-87-3-343
Misra, V., Carlson, R., Craig, K., Enfield, D., Kirtman, B., Landling, W., et al. (2011). *Climate Scenarios: A Florida-Centric View, Florida Climate Change Task Force. http://floridaclimate.org/whitepapers/*
Mo, K. C., & Schemm J. E. (2008). Relationships between ENSO and drought over the southeastern United States. *Geophysical Research Letters, 35*, L15701. https://doi.org/10.1029/2008gl034656
Mo, K. C. (2010). Interdecadal Modulation of the Impact of ENSO on Precipitation and Temperature over the United States. *Journal of Climate, 23*(13), 3639–3656. https://doi.org/10.1175/2010jcli3553.1
Moses, C. S., Anderson, W. T., Saunders, C., & Sklar, F. (2013). Regional climate gradients in precipitation and temperature in response to climate teleconnections in the Greater Everglades ecosystem of South Florida. *Journal of Paleolimnology, 49*, 5–14. https://doi.org/10.1007/s10933-012-9635-0
