AGU PUBLICATIONS

Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE

10.1002/2016MS000657

Key Points:
- The complexity and the bias introduced in small-scale phenomena due to the spectral transform method of CAM Eulerian spectral dynamical core are prominent
- The classification tree algorithm with objective thresholding is successful in detecting different types of precipitation features with high spatial complexity
- An efficient and informative study about the biases produced by GCMs should involve daily (or hourly) output (rather than monthly mean) analysis over local scales

Correspondence to:
M. S. Yorgun,
soner.yorgun@noaa.gov

Citation:
Soner Yorgun, M., and R. B. Rood (2016), A decision tree algorithm for investigation of model biases related to dynamical cores and physical parameterizations, *J. Adv. Model. Earth Syst.*, *8*, 1769–1785, doi:10.1002/2016MS000657.

Received 21 FEB 2016
Accepted 29 SEP 2016
Accepted article online 4 OCT 2016
Published online 11 NOV 2016

A decision tree algorithm for investigation of model biases related to dynamical cores and physical parameterizations

M. Soner Yorgun[1,2] and Richard B. Rood[3]

[1]Earth System Research Laboratory, Global Systems Division, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA, [2]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, USA, [3]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, Michigan, USA

Abstract An object-based evaluation method using a pattern recognition algorithm (i.e., classification trees) is applied to the simulated orographic precipitation for idealized experimental setups using the National Center of Atmospheric Research (NCAR) Community Atmosphere Model (CAM) with the finite volume (FV) and the Eulerian spectral transform dynamical cores with varying resolutions. Daily simulations were analyzed and three different types of precipitation features were identified by the classification tree algorithm. The statistical characteristics of these features (i.e., maximum value, mean value, and variance) were calculated to quantify the difference between the dynamical cores and changing resolutions. Even with the simple and smooth topography in the idealized setups, complexity in the precipitation fields simulated by the models develops quickly. The classification tree algorithm using objective thresholding successfully detected different types of precipitation features even as the complexity of the precipitation field increased. The results show that the complexity and the bias introduced in small-scale phenomena due to the spectral transform method of CAM Eulerian spectral dynamical core is prominent, and is an important reason for its dissimilarity from the FV dynamical core. The resolvable scales, both in horizontal and vertical dimensions, have significant effect on the simulation of precipitation. The results of this study also suggest that an efficient and informative study about the biases produced by GCMs should involve daily (or even hourly) output (rather than monthly mean) analysis over local scales.

1. Introduction

With the recent advances in climate modeling and computational capabilities, general circulation models (GCMs) can produce simulations of the future climate at increasing spatial resolutions. This leads to representation of weather-scale phenomena, which produce highly structured GCM output over local features (e.g., orography). The resolution of weather-scale features in climate models represents a significant advance in the scientific foundation of climate simulations and also improves the potential of models to be relevant to climate-change adaptation applications.

The resolution of weather-scale features in climate models helps to link weather science and climate science and motivates the development of new GCM evaluation techniques. The comparison of modeled fields to observations and/or reanalysis products can be regarded as the primary method to analyze the realism of models by identification of biases and variability and descriptions of the uncertainties. Widely used quantification methods for such comparison are visual analysis of select parameters of model simulations, and summary statistics (e.g., root mean squared error) calculated on a grid point-by-grid point basis [*Endris et al.*, 2013; *Haslinger et al.*, 2013; *Landgren et al.*, 2014; *Martynov et al.*, 2013; *Sillmann et al.*, 2013; *Sylla et al.*, 2013]. Such approaches are important since they are informative of the temporal and geographical distribution of model characteristics. However, they fail to address the source of a particular bias within a model, which is crucial to production of more realistic simulations of the climate system.

GCMs are composed of multiple components that are connected together yielding a complex system. The main components of a GCM are the dynamical core [*Williamson*, 2007], which solves the governing fluid and thermodynamic equations on resolved scales, and the physical parameterizations, which represent subgrid-scale processes and other processes not included in the dynamical core such as radiative transfer [*Rood*, 2011]. These components can be composed of several subcomponents (e.g., cloud macro and

microphysics, tracer transport, etc.) communicating with each other through couplers or mediators. The complexity of GCM structure contributes to the difficulty of determining the causality of model errors. Model errors can be local, produced by a particular component or by the combination of multiple components. Analysis methods using summary statistics likely average out such errors, losing knowledge on causality. This study aims to address this problem by adopting an object-based approach to quantify the bias, describe the uncertainty, and identify their sources within a GCM (explained in the next section). By the application of a classification tree algorithm, we aim to partition the precipitation field to isolate the different types of precipitation features (as identified by *Yorgun and Rood* [2014]) and analyze them separately with the expectation that different features (originating from different underlying processes) will exhibit different statistical characteristics. Additionally, with the partitioning of the precipitation field, we are also motivated by differences between the directly forced precipitation and the precipitation that is produced internally, their characteristics, and organization at intermediate and small scales (i.e., crossing the scales from grid to fully resolved). Capturing and analyzing these differences is valuable in terms of pinpointing the shortcomings of different models and their causes. We focus on a comparative study of the Eulerian spectral and the finite volume (FV) [*Lin and Rood*, 1996, 1997] dynamical core components of the National Center for Atmospheric Research (NCAR) Community Atmosphere Model (CAM) version 5.0 [*Neale et al*., 2010] and the simulation of orographic precipitation. In our previous studies [*Yorgun and Rood*, 2014, 2015], we focused on the monthly mean simulations and quantified the differences between the two dynamical cores. In this study, we add the time dimension to the analysis by analyzing daily simulations. It should be noted that this is not a model verification study, but rather a model-to-model comparison in order to quantify the differences between CAM Eulerian and CAM FV dynamical cores in their simulation of the precipitation over Sierra Nevada [*Yorgun and Rood*, 2014].

1.1. Object-Based Approach

An object is defined as a coherent system with an associated set of measurable parameters [*Douglass*, 2000]. In an object-based model evaluation approach, features such as fronts, rain bands, clouds in deep convection, etc. are defined and isolated as objects [*Posselt et al*., 2012; *Xu et al*., 2005] in modeled and/or observed fields. Process and object-based evaluation preserves information in quantitative analyses by avoiding the need for extensive spatial and temporal averaging. Object-based evaluation has been used in evaluation of weather forecasts and climate simulations [*Byna et al*., 2011; *Davis et al*., 2006; *Ebert and McBride*, 2000; *Micheas et al*., 2007; *Skok et al*., 2013; *Wernli et al*., 2008; *Wick et al*., 2013]. In object-based methods, the meteorological features are identified/detected in the domain of interest, and then isolated and compared with their analogs between modeled and/or observed fields. This makes it possible to evaluate processes in models without needing to reproduce the time and location of, for example, a particular observed cloud system or a cyclone.

The detection of an object is traditionally done by setting a threshold for a variable and evaluating the grid points in the field relative to that threshold (i.e., is the grid value above or below the threshold?). This evaluation is done to identify if the grid point (or observation) belongs to the class of the feature to be detected. The main disadvantages of these methods are the subjective decisions in thresholding (user defined fixed-values), which leads to misses in detection. The results of the current detection methods are highly dependent on these fixed threshold values and even the use of multiple thresholds [*Xu et al*., 2005; *Xu*, 2009; *Posselt et al*., 2012; *Skok et al*., 2013] does not solve the problem of misdetection. This study involves development of a more efficient detection strategy that will alleviate the fixed user-defined threshold problem using a pattern recognition algorithm (classification trees). The algorithm has the flexibility to use information from other fields. In our problem, orographic precipitation features are detected and isolated by using information both from precipitation rates and the topography.

Our analysis starts with the selection of orographic precipitation features (study features) that have consistent differences in controlled GCM simulations with the CAM Eulerian spectral core and the CAM FV core. The precipitation features are mechanistically represented in simplified experimental setups using idealized test cases. A brief summary of the significant results of our previous time-averaged analysis is given in section 2 of this paper; however, the detailed results can be found in *Yorgun and Rood* [2014] and *Yorgun and Rood* [2015]. Once the study features are analyzed qualitatively, the identification and detection algorithm with classification trees (explained in section 3) are applied. The features are isolated from the whole precipitation field as objects and matched with their analogs between CAM Eulerian spectral and CAM FV

Table 1. The Mountain Specifications for Three Experimental Setups

	Center Point in Longitude (λ_c)		Center Point in Latitude (ϕ_c)		Peak Height (h_0) (m)		Half Width (d) (km)	
Case	First Peak	Second Peak	First Peak	Second Peak	First Peak	Second Peak	First Peak	Second Peak
1 Single Mountain	97°E		30°N		1500		1500	
2 Double Mountain	90°E	97°E	30°N	30°N	1500	1500	1500	1500
3 Realistic	90°E	93°E	30°N	30°N	500	3000	500	500

experimental model runs. The comparisons of the statistical characteristics (i.e., peak value, mean value, and variance) of these objects are presented in section 4. Conclusions of the study are given in section 5.

2. Selection of the Study Features and the Experimental Model Runs for Orographic Precipitation

The comparison of the 21 year January mean precipitation on the west coast of United States between Global Precipitation Climatology Project (GPCP) observations [*Beck et al*., 2005] and the Atmospheric Model Intercomparison Project (AMIP) runs with the Eulerian spectral and the FV dynamical cores reveal significant differences [*Yorgun and Rood*, 2014]. The FV model resembles the GPCP observations better in comparison to the Eulerian spectral. For example, the FV model with 0.5° resolution successfully simulates the dry region between the Coast Ranges and the Sierra Nevada, whereas the Eulerian spectral model with T170 triangular truncation merges the precipitation features related to the two mountain ranges to create a single precipitation feature [*Yorgun and Rood*, 2014]. We selected these features as our study features and conducted experimental model runs with NCAR CAM 5.0 with Eulerian spectral T85 (∼1.4°) and T170 (∼0.7°) triangular truncations, and FV 1° and 0.5° resolutions in order to reproduce and further evaluate these features.

Three topography configurations were created using Gaussian bell-shaped mountains ranging from a single mountain peak toward a more structured setup resembling the Coast Ranges and the Sierra Nevada (Table 1).

The treatment of the topography is different between the CAM Eulerian and CAM FV dynamical cores as discussed in *Yorgun and Rood* [2014], and this difference is kept as is in the experimental model runs in this study in order to analyze the standard model configurations. The physics/dynamics time-steps are also different, i.e., the physics and dynamics time steps for the CAM Eulerian Spectral model are identical, and are 1800 and 300 s for T85 and T170, respectively. For the CAM FV model, the physics time steps are 1800 and 600, and the dynamics time steps are 180 and 60 s for 1° and 0.5°, respectively. The physical parameterization is set to a simplified moist parameterization suite called Simple-Physics suite [*Reed and Jablonowski*, 2012] for both CAM FV and the CAM Eulerian Spectral dynamical cores. The suite allows physical processes that are important for orographic precipitation, which are large-scale condensation, boundary layer turbulence of horizontal momentum, temperature and specific humidity, and surface fluxes of horizontal momentum, evaporation (specific humidity), and sensible heat (temperature) from the surface to the lower atmosphere. Moisture (specific humidity) is transported to the mountains via zonal and meridional winds, followed by orographic lift and subsequent large-scale condensation and precipitation. Although the CAM 5.0 includes the subgrid-scale gravity wave drag parameterization especially to account for the deceleration of the zonal wind above the surface [*Neale et al*., 2010], the simple physics package does not have such parameterization thus the solutions presented in this study are not affected by it. The simulations start with a purely westerly wind and they evolve to have northerly, southerly, and vertical components as the simulations continue. The details of the initial conditions can be found in *Yorgun and Rood* [2014]. Three types of precipitation features are simulated in all three topographical setups, which are shown in the monthly mean plot of the double mountain setup (Figure 1).

The features as indicated by their corresponding numbers in the CAM Eulerian spectral T170 simulation (Figure 1b) are:

1. Large-scale features due to stable upslope ascent

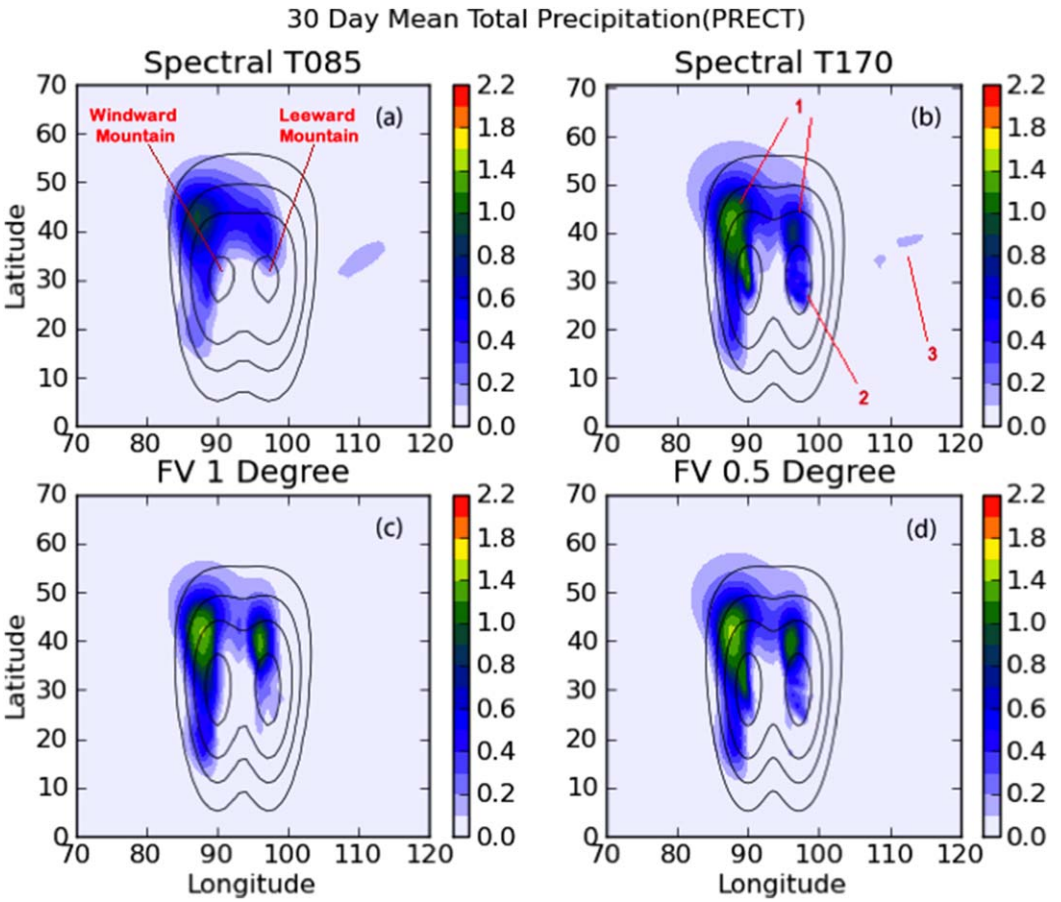

Figure 1. 30 day mean total precipitation (mm/d), surface geopotential (topography) for double mountain setup of experimental test cases, simulated by (a) CAM Eulerian spectral T85, (b) CAM Eulerian spectral T170, (c) CAM FV 1°, and (d) CAM FV 0.5° dynamical cores. Surface geopotential contours are 500, 2000, 5000, and 11000 m^2s^{-2}, with the peak contour at the interior of the mountain. The numbers 1, 2, and 3 on Figure 1b indicate the large-scale feature due to stable upslope ascent, small-scale feature due to local evaporation and lee-side convergence, and the feature due to leeward baroclinic waves, respectively.

2. Small-scale features due to local evaporation and lee-side convergence
3. Feature due to leeward baroclinic waves

The large-scale features (1) appear in the windward side of the mountains early in the simulation (Day 1) covering a relatively large area whereas the small-scale features (2) manifest later in the simulation (around Day 15) closer to the peaks of the mountains. The peak precipitation values within the large-scale features are less than that of the small-scale features. The features due to the baroclinic waves also appear early in the simulation and they consistently producing light precipitation with their shape resembling the wave structure. These features are different in their origins and evolutions, thus it is important to identify and evaluate them as separate objects. In *Yorgun and Rood* [2015], we focused on the large-scale stable upslope ascent features (#1 in Figure 1b). In that study, we made an analysis of the monthly mean simulations of these features and quantified the differences between the two dynamical cores using *k*-means cluster analysis for identification, and variography for comparison of the features. In this study, we add the time dimension to the analysis by looking at daily simulations of both the large-scale stable upslope ascent and the small-scale features due to local evaporation and lee-side convergence (#2 in Figure 1b). This way we aim to quantify the spatiotemporal differences in the representation of these features between the CAM Eulerian spectral and the CAM FV dynamical cores. The features due to leeward baroclinic waves (#3 in Figure 1b) are away from the topography, therefore they are excluded from current analysis and the focus is kept on the features that are over the topography in this study. We focus on the double mountain, and the realistic topographical setups (Table 1) using classification trees for the detection and isolation of the precipitation features.

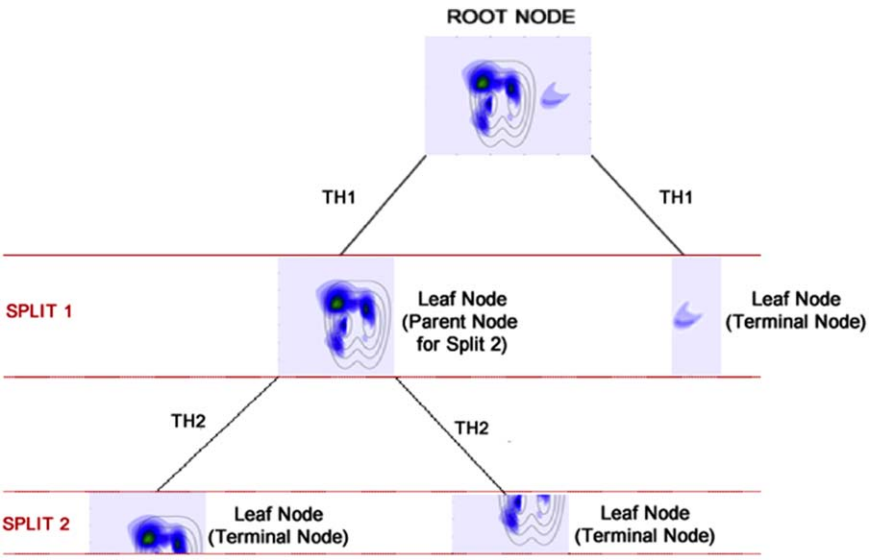

Figure 2. An example classification tree with two rectangular splits at day 15 total precipitation rate for the double mountain case, simulated by CAM FV 0.5°. TH is threshold.

3. Method

A classification tree [*Breiman et al*., 1984] algorithm was implemented for detection and isolation of precipitation features simulated in idealized test cases. Classification trees is a machine-learning method for constructing prediction models from data by recursively partitioning the data space and fitting a simple model into each partition. Given a class variable (Y) with values 1, 2, . . ., k, and p predictor variables (X_1, X_2, . . ., X_p), a classification strategy aims to construct a model to predict the class value of Y from values of X. The solution is a partition of the X space into k disjoint sets S_1, S_2, . . ., S_k, such that the predicted value of Y is j if X belongs to S_j for $j = 1, 2, \ldots, k$. Classification trees yield rectangular sets S_j by recursively partitioning the data field on X [*Loh*, 2011].

The data field (in our case the precipitation rate) is partitioned into rectangular boxes that are called nodes. The initial field (the root node) is split into a series of parent and leaf nodes, where each parent node is split into two leaf nodes. The splitting is done until a stopping criterion is met for a leaf node in which case the leaf node is called a terminal node. Both the split and the stopping decisions are made via an impurity measure. The measure of node impurity (explained below) is based on the distribution of the observed Y values in the node. Our algorithm finds a subset over all X values for the split that minimizes the difference between the impurity of the node and the impurity of its two leaf nodes. An example classification tree of a precipitation field simulated by CAM FV 0.5° is given in Figure 2.

Only two splits are shown in Figure 2 for presentation purposes. The root node is the initial precipitation field for the double mountain case simulated by CAM FV 0.5°. The first split is done over a longitude line, which partitions the initial field such that the large-scale upslope ascent and the small-scale local evaporation and lee-side convergence features remain in one node (westward of the split boundary), and the feature due to the leeward waves remain in the other (eastward of the split boundary). The resulting two nodes of split 1 are both leaf nodes of the root node (i.e., the combination of the two leaf nodes gives the root node). In this example, it is assumed that the impurity measure of the node containing the feature due to the leeward baroclinic wave meets the stopping criterion; therefore, it can also be labeled as a terminal node. A further split is done over the node containing the large and small-scale features; therefore, it can also be labeled as a parent node. The second split is done over a latitude line on the parent node resulting from split 1, which separates the small-scale features in one node, and the large-scale features on the other. Both nodes are the leaf nodes of their parent node; however, there are no more splits in this example so they can also be labeled as terminal nodes. As can be seen in Figure 2, there are two threshold values employed for each split (TH1 and TH2). Although these threshold values are fixed values, they are determined objectively after each split by the characteristics of the field thus they are not user-defined fixed

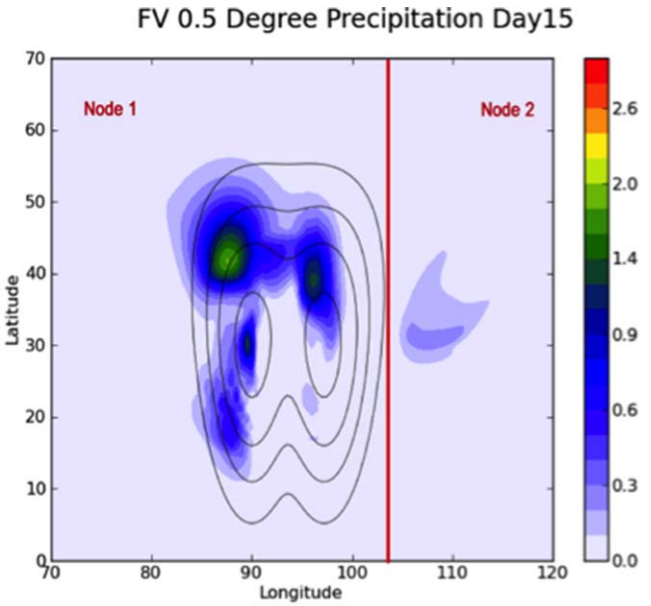

Figure 3. Day 15 total precipitation rate, surface geopotential (topography) for double mountain case, simulated by FV 0.5°. The longitudinal split boundary is shown with the red line.

values for the whole application. The decision of these threshold values, the decision of the longitude (or latitude) line as the split boundary due to an impurity measure, and the decision of stopping criteria after each split are the key factors in the formation of such a classification tree. These factors are explained in detail by looking at the first split given in Figure 3.

In Figure 3, the whole precipitation field is the root node and is split into two leaf nodes at longitude 104° W. For each partition, the algorithm calculates the impurity of each node via a measure called Gini index:

$$Gini(node) = \sum_{y=1}^{Y} p_y(1-p_y) \quad (1)$$

where p_y is the frequency of occurrence of grid-points belonging to class y. We defined two classes (Y = 1, 2) in our calculations, one being the rain (i.e., the grid-point that belongs to the object of interest) and the other being the nonrain (i.e., any grid-point outside objects). The classes and the corresponding node impurities are defined via thresholds. The decision of the split is done via maximizing the Qsplit parameter:

$$Qsplit = Impurity(parent) - \left[\frac{N_{node1}}{N_{parent}} Impurity(node1) + \frac{N_{node2}}{N_{parent}} Impurity(node2)\right] \quad (2)$$

where N_{parent}, N_{node1}, and N_{node2} are the total number of grid-points included in the parent node and subsets of leaf nodes (nodes 1 and 2), respectively. Each impurity value (i.e., the impurity values for parent, node1, and node2) in equation (2) is calculated by the Gini index given in equation (1) (i.e., *Impurity(node) = Gini(node)*). The split in Figure 3 is achieved via maximizing the Qsplit with the threshold value of 0.5 mm/d (i.e., if a grid-point value is above 0.5 mm/d, it is considered rain class, if it is below this threshold, it is considered nonrain class). After this first split, the algorithm continues to split each leaf node until all the resulting nodes become terminal, which is decided when the impurity of nodes drops below a certain value (i.e., until they become pure). This value is taken as 0.05 in this study, which ensures very low impurity in the resulting nodes. The splitting is done sequentially in latitudes and longitudes (if the first split is in longitudes, then the second is in latitudes, third is in longitudes etc.); however, this sequence can be changed depending on the user's choice.

This nature of classification trees that allow step-by-step splitting of the root node into terminal leaf nodes allows the resetting of the threshold value according to the characteristics of each parent node. In Figure 3, node 1 includes two types of features discussed in section 2, namely the large-scale upslope ascent feature and the small-scale feature due to local evaporation and lee-side convergence, whereas node 2 includes the feature due to the leeward waves. These features are different in their manifestations, characteristics, and evolutions in time; therefore, they should not be classified with the same threshold. The algorithm objectively calculates the threshold before each split by looking at the statistics of the parent node to be split. In this study, the threshold is defined by adding half of the standard deviation of the node to the mean value of the node. This measure is selected empirically based on the qualitative analysis on the precipitation features we have previously conducted [*Yorgun and Rood*, 2014] and can be changed according to the nature of the problem. The types, shapes, and structures of the precipitation features were analyzed and visually differentiated in that analysis, and that knowledge is used as validation of the resulting isolated objects by the algorithm in this study. We used several different statistical combinations (e.g., mean, 2 x

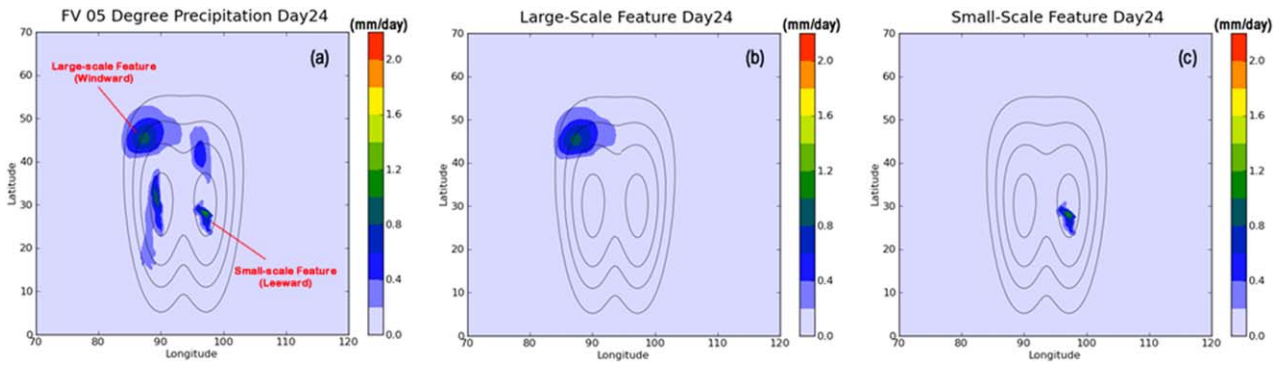

Figure 4. Day 24 total precipitation rate, surface geopotential (topography) for the double mountain case, simulated by (a) CAM FV 0.5°. The isolated large-scale stable upslope ascent feature on the windward mountain and the small-scale feature on the leeward mountain are shown in b and c, respectively.

mean, mean + standard deviation, etc.) and observed that the mean + 0.5 x standard deviation results matched the expected shape/structure of the different precipitation features better.

As indicated previously, the decision tree algorithm partitions the initial field into rectangular boxes. This introduces a limitation to our problem since the precipitation features are not rectangular. Depending on the threshold value, the algorithm may split a precipitation feature into more than one rectangle by chopping off the edges of the curved feature. A postprocessing was conducted to overcome the issue where we reconstructed the split features–if any—into the original form. We used information from the topography field for that purpose. The qualitative analysis [*Yorgun and Rood*, 2014] gave information about the locations of each type of feature, therefore the terminal nodes, which are exactly over the peak of mountains are aggregated to form the small-scale features due to local evaporation and lee-side convergence. The other two types of features are formed in the same manner by measuring the node distances to mountain peaks and the locations relative to the mountain peaks (i.e., the Euclidian distance between the center of mass of the object within a node to the mountain peak). The similar distances in the windward sides of the mountains are merged to form the large-scale stable upslope features, whereas the similar distances in the leeward side (of the eastward mountain in double mountain case) are merged to form the features due to baroclinic waves. This flexibility of the algorithm also allows the usage of other information (e.g., moisture flux convergence, winds) if needed.

The detection and extraction of the features with the classification tree algorithm was applied to daily precipitation results for a total model run time of 30 days. Once the features were extracted, some simple statistical analyses (e.g., mean values, peak values, variances) were conducted and the results were compared for all four simulations, which are discussed in the next section.

4. Discussion of the Results

The classification tree algorithm is applied to the daily simulations of the double mountain and the realistic setups of the idealized simulations. Our comparison focuses on the large-scale upslope ascent features and the small-scale features due to local evaporation and lee-side convergence simulated by CAM FV 1° and 0.5°, and CAM Eulerian spectral T85 and T170 triangular truncation resolutions.

An example of how the algorithm partitions the precipitation field simulated by CAM FV 0.5° for the double mountain setup is given in Figure 4. The algorithm was able to identify and isolate the large-scale upslope ascent feature (Figure 4b) and the small-scale local evaporation and lee-side convergence feature (Figure 4c). The large-scale features in front of the windward mountain and the small-scale features on the leeward mountain are selected to be included in our evaluation. The analysis on the large-scale feature in front of the leeward mountain and the small-scale feature on the windward mountain produced similar results to those included in this study; therefore, they are excluded to avoid repetitive results (they are classified as the same with these two types of features). The algorithm was also applied to CAM FV 1° and the CAM Eulerian spectral T85 and T170 simulations of the double mountain case and the precipitation features were

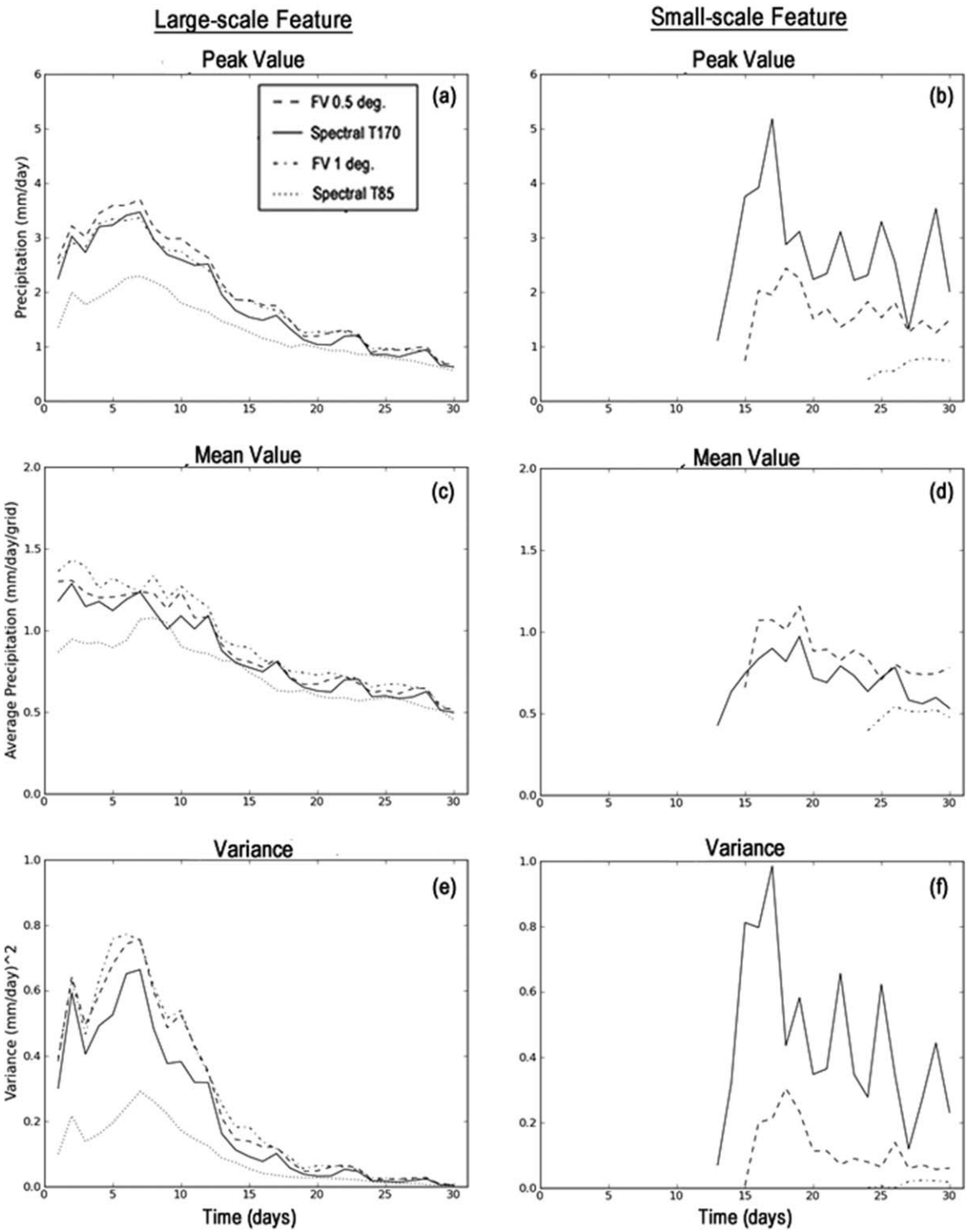

Figure 5. The peak value, mean value, and the variance for the large-scale stable upslope ascent feature on the windward mountain (a, c, e), and the small-scale local evaporation and lee-side convergence feature on the leeward mountain (b, d, f) for the double mountain setup, simulated by all four models.

successfully isolated. Then the peak value, the mean value, and the variance for the isolated features were calculated within the enclosed object boundaries for each day over 30 days of simulation time.

The results for the large-scale stable upslope ascent feature (Figures 5a, 5c, and 5e) show close agreement between CAM FV 1° and 0.5°, and the CAM Eulerian spectral T170 models; however, there is a disagreement

for CAM Eulerian spectral T85 results (the dotted line). CAM Eulerian spectral T85 simulated this large-scale feature with lower intensity, which is in agreement with our qualitative analysis in *Yorgun and Rood* [2014]. This lower intensity was directly related to the spectral filtering of the topography to control aliasing. Figure 5e shows that CAM Eulerian spectral T85 also produced lower values of variance whereas the other simulations exhibit high variances in the beginning and between day 5 and day 10. Note that the variance is calculated over the values within the boundaries of the objects determined by the classification tree algorithm; therefore, low values of variance indicate a smoother feature without a distinct peak of precipitation. A peak precipitation is expected for a stable upslope precipitation feature given the nature of such features [*Roe*, 2005], where a peak value is observed together with lower values around it. The variance results also show how both CAM FV resolutions (dashed lines) agree closely with each other yielding slightly higher values than that of CAM Eulerian spectral T170. The results start to become more similar for all models after day 15 for the large-scale stable upslope feature as it loses intensity and the amount of rain is reduced. These results suggest that the large-scale stable upslope precipitation is resolved by both schemes, with the spectral scheme revealing strong sensitivity to the topographical filtering required to control aliasing and noise [*Yorgun and Rood*, 2014].

The agreement of model simulations for the large-scale upslope ascent feature is not seen for the small-scale local evaporation and lee-side convergence feature (Figures 5b, 5d, and 5f). The onset time of this small-scale feature is different for all four simulations. CAM Eulerian spectral T170 starts simulating this object the earliest, with low peak and variance values (Figures 5b and 5f) followed by an abrupt increase afterward and an oscillatory behavior until the end of the simulation with significantly higher values compared to both CAM FV resolutions. On the other hand, CAM FV 1° and 0.5° resolution simulations (starting on day 24 and 15, respectively) exhibit relatively stable behavior, keeping similar peak, mean, and variance values throughout the simulations. CAM Eulerian spectral T85 did not simulate this feature within the 30 day simulation time; however, a 45 day run was made and it was observed that spectral T85 simulated this feature starting at day 37.

The mean values for the small-scale feature (Figure 5d) for CAM Eulerian spectral T170 are stable throughout the simulation time. The mean values are calculated such that the precipitation rate is summed over all grid points and divided by the number of grid points within the boundaries of the object. A qualitative examination of the isolated small-scale local evaporation and lee-side convergence features shows that CAM Eulerian spectral T170 simulates these features over a larger spatial extent, therefore the mean values of CAM Eulerian spectral T170 are comparable to CAM FV 0.5° even though CAM Eulerian spectral T170 produces higher peak precipitation rates. High variance values (i.e., high and low values enclosed within the object) for CAM Eulerian spectral T170 also support the agreement in the mean values between CAM Eulerian spectral and CAM FV. The agreement of mean values, but the disagreement of the variance between the models suggest that there is a level of agreement between models in terms of the amount of rain they produce within a given object, however there are significant differences in the variability of the simulated precipitation. It should also be noted that the evolution of the mean and variance for the whole field (as opposed to the partitioned field presented in this study) resembles the behavior of the large-scale features (Figures 5c and 5e). This shows the utility of the object-based technique used in this study, i.e., partitioning of the whole field led us understand the behavior of the small-scale features which would otherwise be lost.

The results shown in Figure 5 give a quantitative picture of how scales of precipitation features differ between the CAM FV and CAM Eulerian spectral dynamical cores. The results for the more realistic idealized setup with smaller scales in the horizontal provide the next level of complexity in analyzing how these dynamical cores simulate both large-scale and small-scale precipitation.

Figure 6 shows the day 20 total precipitation rate simulated by the CAM Eulerian spectral T85 and T170, CAM FV 1° and FV 0.5° for the realistic idealized setup. As discussed in *Yorgun and Rood* [2014], there are significant differences in the simulated precipitation between CAM Eulerian spectral and CAM FV models with the increased complexity of the underlying topography. We focus on the large-scale upslope ascent and the small-scale local evaporation and lee-side convergence features simulated on the leeward mountain by finer resolution simulations (CAM Eulerian spectral T170 and CAM FV 0.5°). As seen in Figures 6b and 6d, the precipitation features simulated on the leeward mountain are more complex than in the double mountain case. It is also worth noting that the much higher leeward mountain range does not behave like the single mountain experiment. That is, the lower windward range has a strong influence on the circulation and the

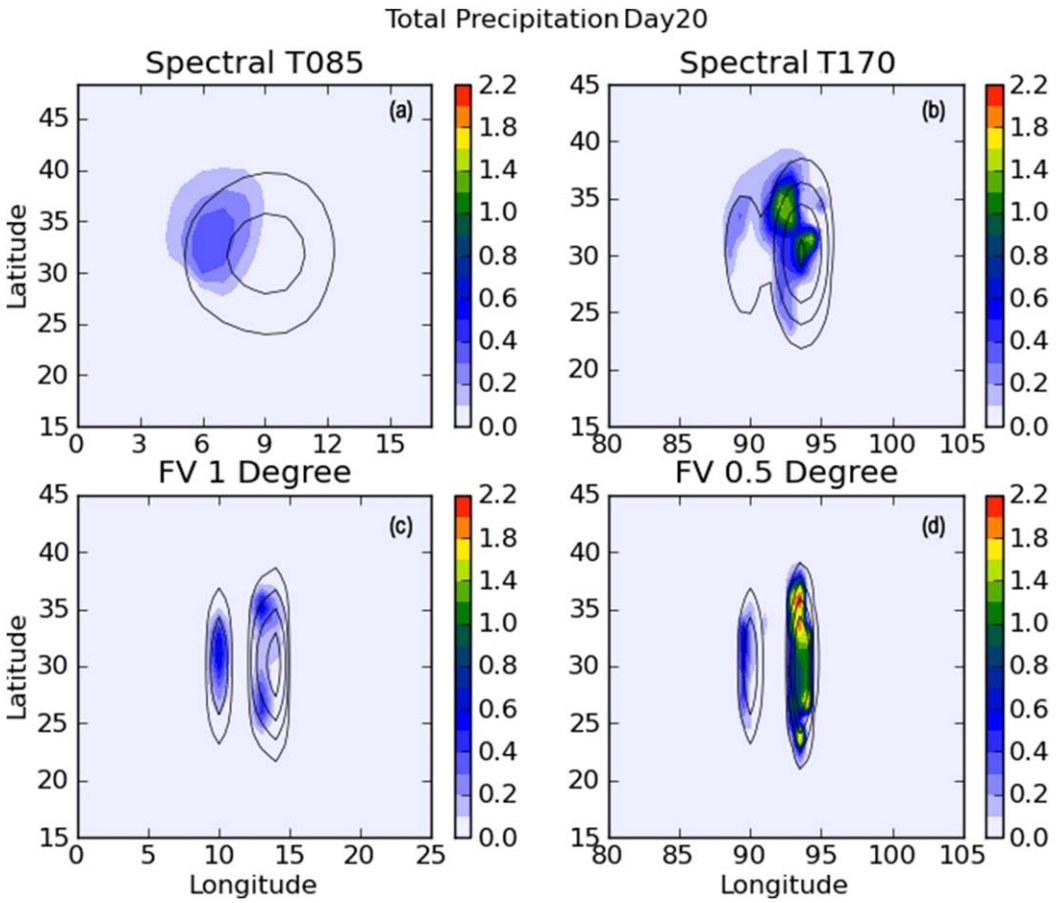

Figure 6. Day 20 total precipitation rate (mm/d), surface geopotential (topography) for the realistic case, simulated by (a) CAM Eulerian spectral T85, (b) CAM Eulerian spectral T170, (c) CAM FV 1°, and (d) CAM FV 0.5°. Surface geopotential contours are 500, 2000, 5000, and 11000 m^2s^{-2}, with the peak contour at the interior of the mountain.

precipitation. Both the high and low-resolution FV experiments have a region of no rain between the mountains. The lower resolution CAM Eulerian spectral (Figure 6a) does not even represent the two-mountain structure after the spectral filter has smoothed the topography to be consistent with the T85 resolution. The higher-resolution CAM Eulerian spectral (Figure 6d) joins the precipitation from two mountain ranges.

We note that for the realistic case the mountains are only separated by 3° longitude, as opposed to 7° longitude for the double mountain case. This offers special challenges for the global basis functions of the spectral method in terms of the ability to resolve the gap (valley) between the mountains. Though the FV dynamics do not formally resolve features of the spatial scale of the valley between the ranges, the local nature of the scheme allows the physics to function at a scale that increases the meteorological realism of the wet-dry contrast.

Focusing on the objects, and the possible differentiation of large-scale upslope and small-scale lee-side precipitation, the features are not obviously distinct over the leeward mountain in Figures 6b and 6d. However, the classification tree algorithm is able to isolate each feature successfully by the precipitation value patterns, without the need of additional information from other fields such as winds or moisture flux convergence.

As we have shown in *Yorgun and Rood* [2014], the positive moisture flux convergence (MFC) is a good indication of precipitation occurring on smaller scales. Figure 7 shows the MFC plots (Figures 7a and 7b) and precipitation (Figures 7c and 7d). The small-scale, leeside precipitation features were objectively isolated from the whole precipitation field (Figures 6b and 6d) by the classification tree algorithm for CAM FV 0.5° and CAM Eulerian spectral T170, respectively. The isolated precipitation overlaps with the positive MFC areas. The surface evaporation leads to an influx of moisture, which combined with the small-scale

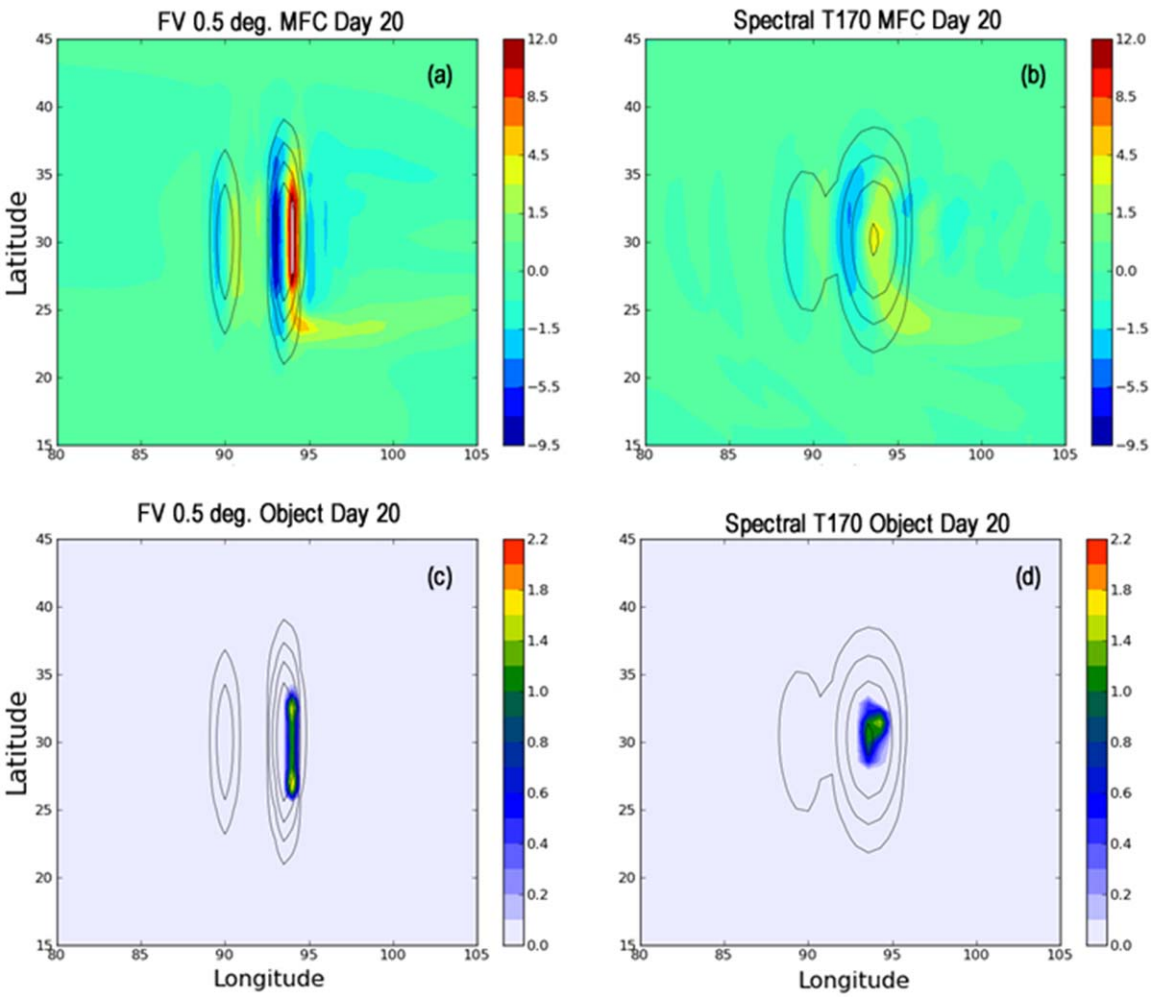

Figure 7. Day 20 moisture flux convergence (10^{-4} kg m^{-2}s^{-1}) by (a) CAM FV 0.5°, (b) CAM Eulerian spectral T170 and their related precipitation (mm/d) objects (c, d) isolated by the classification tree algorithm. Surface geopotential contours are 500, 2000, 5000, and 11000 m^{2}s^{-2}, with the peak contour at the interior of the mountain.

circulation caused by the complex topography, leads to a convergence of moisture leading to precipitation. Therefore, these features are not related to stable upslope ascent but rather the local evaporation and lee-side convergence features the classification tree algorithm isolated from the whole precipitation field.

The peak and the mean values as well as the variances are calculated for these isolated features (Figure 8). In this case, there is better agreement between the CAM FV 0.5° and CAM Eulerian spectral T170 simulated small-scale local evaporation and lee-side convergence features compared to their analogs in the double mountain setup (Figures 5b, 5d, and 5f). In the realistic case, the onset of the small-scale feature is earlier for CAM FV 0.5°, which was not the case for double mountain setup. The mean values for CAM FV 0.5° are higher than that of CAM Eulerian spectral T170 (Figure 8b), which is also not the case for the double mountain setup (Figure 5d). The mean values and the variance for CAM Eulerian spectral T170 exhibit an abrupt increase right after the onset of precipitation. The statistical analysis given in Figure 8 shows that, unlike in the double mountain case, both CAM Eulerian spectral T170 and CAM FV 0.5° simulated similar small-scale local evaporation and lee-side convergence features.

The spatial structure is, however, quite different. The FV simulation is aligned with the topography, and more confined to the leeward side of the mountain. The largest quantitative difference is the spatial extent, where CAM Eulerian spectral T170 bulges due to the shape of topography (Figure 7d). Though, statistically similar in this case, the relation of the spatial structure to the topography suggests different behavior of the two dynamical cores.

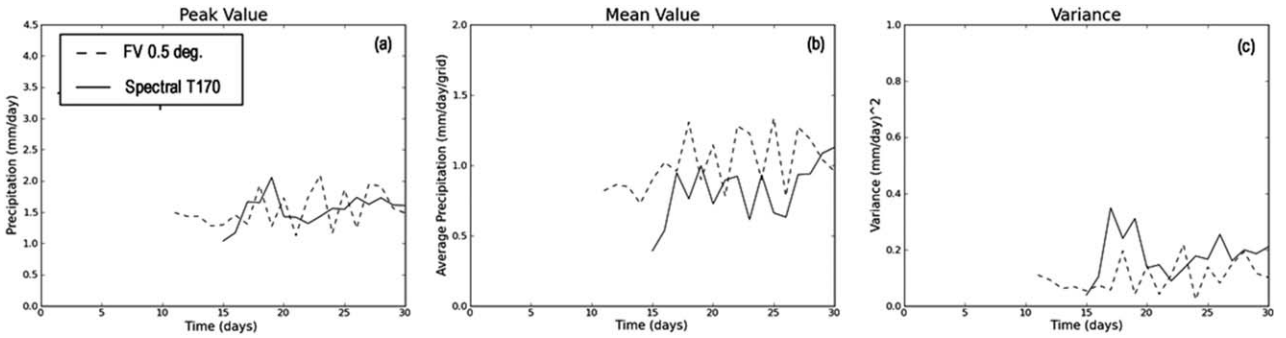

Figure 8. (a) The peak value, (b) mean value, and (c) the variance for the small-scale local evaporation and lee-side convergence features for the realistic setup shown in Figure 7, simulated by CAM FV 0.5° and CAM Eulerian spectral T170.

The results, thus far, establish that when well resolved, with comparable lift, the two dynamical cores are consistent in their representation of stable, upslope precipitation. The issue of topographical lift is not trivial, as the filtering to control aliasing and noise in the spectral method leads to large differences at different resolutions. The smaller-scale precipitation, which might be framed as that associated with internal as opposed to forced dynamics, is quite different between the two dynamical cores. The spatial structure is also quite different. The mechanisms of the smaller scale precipitation can be categorized as two types. The first is associated with dynamical structure, e.g., waves, that are formed after the flow encounters the mountains. The second type is numerical; that is, a numerical artifact causes condensation and precipitation. The first type is expected to show consistency with the moisture flux and its convergence. The second type is not consistently related to the dynamical features.

4.1. Sensitivity to Sea Surface Temperature (SST)

The statistics and spatial structure of the smaller-scale precipitation objects suggest different behavior between the two dynamical cores. In *Yorgun and Rood* [2014], we noted the small-scale features were highly sensitive to the specification of the sea-surface temperature (SST). Numerical experiments show that when increasing the SST above 288 K the small-scale precipitation features become dominant not just over the mountains, but, also, away from the mountains. As a sensitivity experiment, therefore, we decrease the surface temperature of the experimental configuration by 1°K (i.e., 287 K instead of 288 K) for the double mountain configuration. The statistics for the large-scale and the small-scale objects are shown in Figure 9, calculated as *differences* from the control experiment with the original surface temperature. For the stable, upslope simulation, the differences between the two experiments are small for both dynamical cores (Figures 9a and 9c). For the small-scale precipitation, the 1°K difference of sea-surface temperature causes much larger changes in the spectral simulation than in the FV simulation (Figures 9b and 9d).

4.2. Grid-Scale Noise Structure

It is reasonable to postulate that the grid-scale noise structure of the two dynamical cores has an important impact on precipitation. Notionally, dispersion errors, especially in water vapor transport [*Lin and Rood*, 1996; *Rood*, 1987], can cause numerical initiation of precipitation. However, to determine cause and effect in simulations of this complexity is more difficult. The condensation of moisture is a discontinuous process in time, which introduces sharp gradients that are especially problematic for spectral models. These processes are candidates to produce the Gibbs phenomenon; thereby reducing the realism of spectral model simulations [*Geil and Zeng*, 2015; *Navarra et al*., 1994].

It is natural in model studies of precipitation to focus on where it is raining and the characteristics of the rain events. However, to understand the sensitivity of precipitation to a dynamical core, it is proposed that examination of grid-scale structure, which can initiate precipitation, is better founded. To suggest the potential value of this approach, the whole extent, rather than the boundaries determined by the classification tree algorithm, of the features on the western (nominally, windward) mountain for day 16 is given in Figure 10. The color scheme of Figure 10 (0–0.1 mm/d) is adjusted to show how the grid-scale precipitation differs between FV 0.5° and CAM Eulerian spectral T170. Not only are there scalloped edges of the rain border, but the structure of the gradient of increasing precipitation on the western side of the mountain has

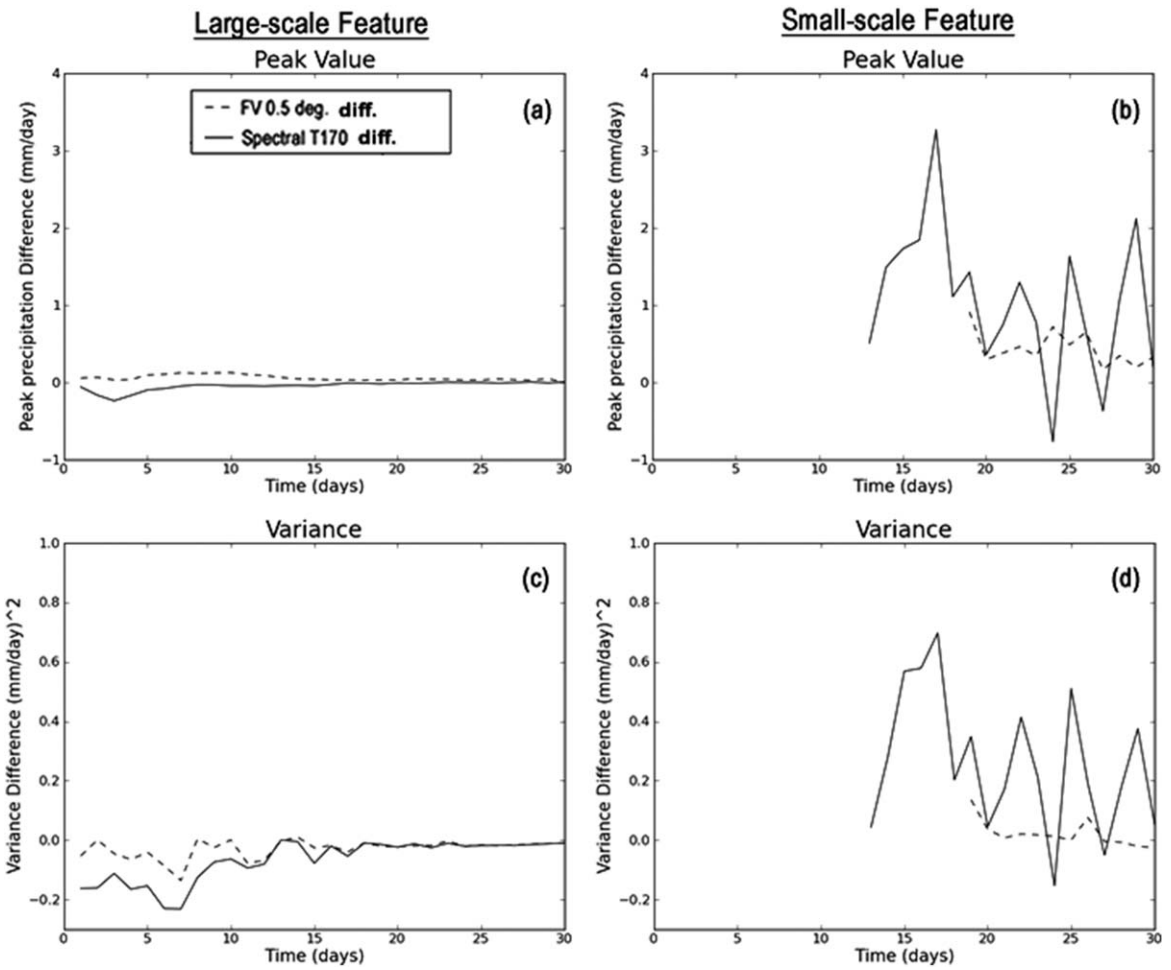

Figure 9. The peak value, and the variance *difference* (288 K – 287 K SST runs) for (a, c) the large-scale stable upslope ascent feature on the windward mountain, and (b, d) the small-scale local evaporation and lee-side convergence feature on the leeward mountain for the double mountain setup, simulated by CAM Eulerian spectral T170 and CAM FV 0.5°.

lobes of higher and lower precipitation. This suggests a strategy of removing the objects from the field in order to expose the grid-scale information, which will be the subject of future studies.

5. Conclusions

Daily simulations of CAM Eulerian spectral T170 and T85, CAM FV 1° and 0.5° resolutions for the double mountain and the realistic setups were analyzed in an object-based framework. The treatment of the topography as well as the physics/dynamics time-steps are different between CAM Eulerian and CAM FV models used in this study (as indicated in section 2) in order to keep the standard model configurations. A classification tree algorithm that allows flexible thresholding was implemented to detect and isolate the small-scale and large-scale precipitation features. The isolated features were compared by simple statistics across all simulations.

Though these idealized simulations have simple, smooth topography, complexity in the precipitation fields develops quickly. This complexity is due to both internally computed dynamics and numerical artifacts. The classification tree algorithm using objective thresholding proved successful in sensible isolation of precipitation features even as the complexity of the precipitation field increased. The isolation of like objects in different simulations let us focus on local phenomena and quantify the differences between CAM Eulerian spectral and CAM FV models in simulation of orographic precipitation. We documented higher sensitivity of the spectral simulations, compared to the FV simulations, to resolution as well as to changes in the surface temperature of the aquaplanet. The studies of *Kent et al.* [2014], *Whitehead et al.* [2015], and *Scott et al.*

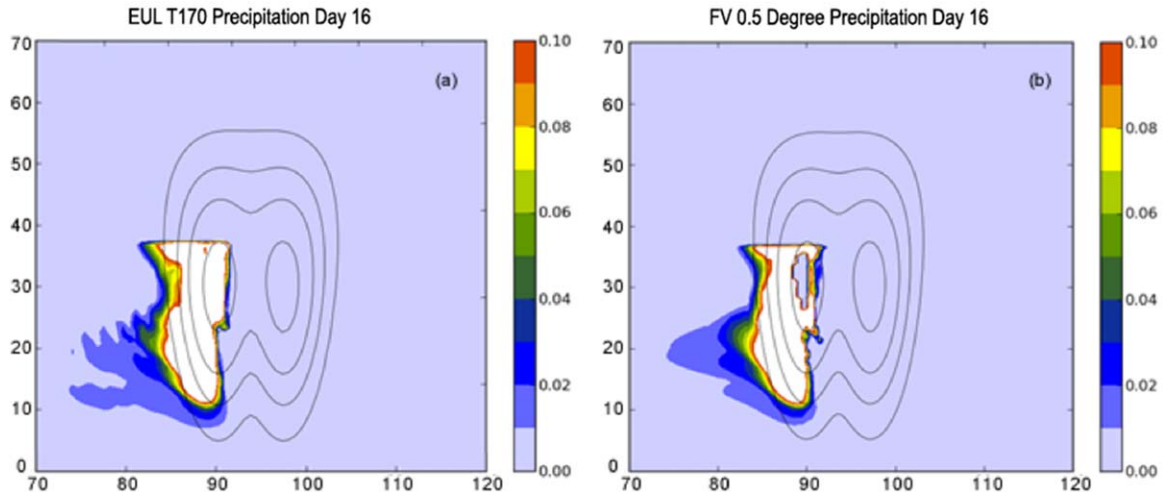

Figure 10. Day 16 total precipitation rate for the small-scale features on the windward mountain, surface geopotential (topography) for double mountain case, simulated (a) by FV 0.5° and (b) by CAM Eulerian spectral T170.

[2015] investigate the diffusivity of CAM FV and comparison to diffusivity in spectral methods. Generally, it is sought to minimize the diffusion needed to assure accurate diffusion. At comparable levels of diffusion to manage numerical dispersion errors and grid-scale noise, the spectral methods continue to show signatures of spectral ringing, oscillations, which leave signatures in the precipitation fields.

The results of this study suggest that even for a climate model, investigation of precipitation events on daily of even hourly output can provide information about model biases. We were able to quantify the differences between the FV and the spectral models (e.g., the erratic behavior of spectral models, especially at the onset of the small-scale precipitation features) by looking at the daily evolution of simulated precipitation features. Averaging in the case of monthly-mean analysis would lose this information. Small-scale features such as the ones identified due to local-evaporation carry significant information about the bias originated by their corresponding physical parameterizations and numerical schemes. This is also important for analyzing "local" scales where significant bias can manifest in GCM simulations, which cannot be quantified with summary statistics due to averaging out.

Our study showed that analyzing local features rather than a whole precipitation field gives more accurate insight about the nature of bias. Two different types of features over the mountains of the same model run (double mountain case) show significant difference in their evolution over time. These differences are large enough to contribute to biases in the means, partitioning in the scales of precipitation, and, especially in complex topography, provision of precipitation to watersheds. The small-scale feature has a completely different mechanism of manifestation (i.e., local evaporation and lee-side convergence) than the large-scale, upslope precipitation, and isolating these features gives insight about the bias related to that mechanism, how it is parameterized, and how the parameterization is coupled with dynamics.

One prominent conclusion of this study is that the resolvable spatial scales of precipitation features to be simulated play a crucial role in how they will be simulated by GCMs. This is naturally related to the scales of the mountains involved in the simulation. Throughout the discussion of the results, it has been clear that the simulation of large-scale features due to stable upslope ascent was in agreement between CAM Eulerian spectral and CAM FV model with the exception of CAM Eulerian spectral T85. However, the quantitative characteristics of the simulated small-scale features by CAM FV and CAM Eulerian spectral models were not in agreement. An important reason for this difference, or in other words, the reason why the CAM Eulerian spectral models failed to simulate precipitation features in agreement with CAM FV models and observations as seen in the AMIP runs [*Yorgun and Rood*, 2014] is the bias introduced in small-scale phenomena due to the spectral transform method and the spectral filtering applied to topography in the CAM Eulerian spectral as discussed in *Yorgun and Rood* [2014]. Our results suggest that the spectral scheme exhibits strong sensitivity to the topographical filtering required to control aliasing and noise. In *Bala et al*. [2008a], it was discovered that the default configuration of the CAM applied the spectral-filtered topography to

both the FV and EUL dynamical cores. A filter consistent with the FV numerical formulation was introduced; it was strongly focused on grid-scale (twice the grid length) waves. In global simulations, comparing the two filters, surface pressure patterns and upper air temperatures, for example, were quite similar. Differences in precipitation near topography, for example, the summertime monsoonal flow in Mexico, the Andes, and the West Coast of North America were immediately discernible. The spectral-filtered topography was far smoother than required for the FV numerical attributes, and it was concluded that a specification of topography that was consistent with the dispersion characteristics of the dynamical core was scientifically correct and an integral part of the dynamical core.

In the original specification of the mountain wave test, it was suggested that the test be run at 1° resolution as it was determined to be converged at this resolution (C. Jablonowski, 2015, personal communication). With the introduction of the double mountain and the realistic topography, smaller spatial scales are introduced to the problem. Furthermore, we have moved our attention from the convergence of resolved fluid dynamical parameters to parameters dependent on grid-scale physics. As discussed in *Yorgun and Rood* [2014], for the double mountain case the T85 has effectively 5 grid points between the mountain peaks and the T170 has 10 grid points. For the realistic topography, T85 merges the peaks and T170 has only five grid points. Our results suggest that with the spatial scale of the topography at five grids points the spectral simulation degrades badly. As the spatial scale of the topography approaches 10 grid points, the FV and EUL solutions start to converge. *Bala et al*. [2008b] and *Rood* [1987] discuss that because the FV core has a highly localized computational stencil, as compared to the global stencil of the spectral basis functions, FV has benefit near steep topography for many variables associated with the model physics.

The procedure for the coupling of dynamics and physics in CAM 5.0 plays an important role in the manifestation of this type of bias in spectral dynamical core. In CAM 5.0, there are two types of coupling strategies, namely time split and process split. In the time split coupling, dynamics and physics are calculated sequentially, each based on the state produced by other. Whereas in the process split coupling the calculations of dynamics and physics are both based on the same past state [*Neale et al*., 2010]. Between the two coupling strategies, process split is more suitable for spectral transform models (and it is the default in CAM 5.0) since the time split requires extra spectral transforms to convert the momentum variables provided by the physical parameterizations. The formulation of the CAM spectral model is such that every physical process is calculated and updated on the Gaussian grid, and then their contribution to the dynamics is transformed to spectral space. This procedure creates problems when dealing with discontinuous physical processes that occur in small-scale such as the ones highlighted in this study (small-scale features in Figure 1).

As indicated throughout this paper and quantified in Figure 9, the manifestation of the small-scale features due to local evaporation and lee-side convergence have significant contribution from the surface flux parameterization of the simple physics suite [*Reed and Jablonowski*, 2012] and a key component of this evaporation is the SST value of the aquaplanet setting. Once the Gaussian grid precipitation value is updated by surface flux parameterizations, it introduces a relatively high value of precipitation on the grid level and causes a "jump" between the previous and current time states. This jump is not handled accurately with the spectral transform method and numerical noise is introduced. This noise is clearly observed in Figures 5b and 5f where we quantified the initial oscillatory behavior of double mountain test case small-scale feature simulated by the CAM Eulerian spectral model immediately after its genesis. As the simulation proceeds, the oscillations become less severe since the grid values become similar between the previous and current time states and consequently the spectral transform method produces lesser numerical noise. The CAM Eulerian spectral model also simulated higher precipitation values compared to FV throughout the simulation which shows the overrepresentation of precipitation by spectral models shown by *Williamson and Rasch* [1994].

However, we also showed that such noise is not as prominent for the same type of feature in the realistic test case (Figure 8). This is also related to the interaction between physical parameterizations and dynamics. The planetary boundary layer (PBL) diffusion in simple-physics suite is parameterized such that the physical processes below 850 hPa are affected by it [*Reed and Jablonowski*, 2012]. In the double mountain case for the spectral model, the peak mountain height is below 850 hPa (the smoothing of the topography by filtering also contributes to that) therefore the small-scale scale features simulated over the mountain are

affected by PBL parameterization. Whereas the leeward mountain in the realistic case (where the small-scale features occur) is sufficiently higher so that the PBL parameterization has lesser effect. Therefore, there is more contribution from physics in Gaussian grid level in the double mountain case, which results in considerable amount of bias when the values are transformed to the spectral space. Our analysis also suggests that when the simulations are physical with less numerical noise (as in the case of the realistic setup in this study), it is observed that there is a correlation between the resolution and the onset time of the small-scale features, i.e., the higher the resolution, the earlier the onset.

Acknowledgments
The authors thank Christiane Jablonowski, Derek Posselt, Kevin Reed, James Kent, and Jared Whitehead for their help in running idealized test cases and for their valuable input. We also would like to acknowledge the computing support by National Center for Atmospheric Research (NCAR) (especially Lawrence Buja) for the model runs. The CAM EUL model data were provided by Julie Caron (jcaron@ucar.edu) from NCAR and the CAM FV model data were provided by Gerald Potter (Gerald.potter@nasa.gov) from the Lawrence Livermore National Lab archives. This study was funded by the National Aeronautics and Space Administration (NASA) (Award NNX08AF77G) and the U.S. Department of Energy (DOE) (Award DE-SC0006684), and by NOAA Climate Program Office (award number NA10OAR4310213), which supports the Great Lakes regional Integrated Sciences and Assessments Center (GLISA).

References

Bala, G., et al. (2008a), Evaluation of a CCSM3 simulation with a finite volume dynamical core for the atmosphere at 1° latitude × 1.25° longitude resolution, *J. Clim.*, *21*, 1467–1486.

Bala, G., R. B. Rood, D. Bader, A. Mirin, D. Ivanova, and C. Drui (2008b), Simulated climate near steep topography: Sensitivity to numerical methods for atmospheric transport, *Geophys. Res. Lett.*, *35*, L14807, doi:10.1029/2008GL033204.

Beck, C., J. Grieser, and B. Rudolf (2005), A New Monthly Precipitation Climatology for the Global Land Areas for the Period 1951 to 2000 published in Climate Status Report 2004, pp. 181–190, German Weather Service, Offenbach, Germany.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone (1984), *Classification and Regression Trees*, CRC Press, Boca Raton, Fla.

Byna, S., Prabhat, M. F. Wehner, and K. Wu (2011), Detecting atmospheric rivers in large climate datasets, in *Proceedings of the 2nd International Workshop on Petascal Data Analytics: Challenges And Opportunities*, pp. 7–14, Association for Computing Machinery, N. Y.

Davis, C., B. Brown, and R. Bullock (2006), Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas, *Mon. Weather Rev.*, *134*, 1772–1784.

Douglass, B. P. (2000), *Real-Time UML: Developing Efficient Objects for Embedded Systems*, Addison-esley, 328 pp., Addison-Wesley Longman Ltd., Essex, U. K.

Ebert, E. E., and J. L. McBride (2000), Verification of precipitation in weather systems: Determination of systematic errors, *J. Hydrol.*, *239*, 179–202.

Endris, H. S., et al. (2013), Assessment of the performance of CORDEX regional climate models in simulating East African Rainfall, *J. Clim.*, *26*, 8453–8475.

Geil, K. L., and X. Zeng (2015), Quantitative characterization of spurious numerical oscillations in 48 CMIP5 models, *Geophys. Res. Lett.*, *42*, 5066–5073, doi:10.1002/2015GL063931.

Haslinger, K., I. Anders, and M. Hofstatter (2013), Regional climate modelling over complex terrain: An evaluation study of COSMO-CLM hindcast model runs for the Greater Alpine Region, *Clim. Dyn.*, *40*, 511–529.

Kent, J., J. P. Whitehead, C. Jablonowski, and R. B. Rood (2014), Determining the effective resolution of advection schemes. Part I: Dispersion analysis, *J. Comput. Phys.*, *278*, 485–496.

Landgren, O. A., J. E. Haugen, and E. J. Forland (2014), Evaluation of regional climate model temperature and precipitation outputs over Scandinavia, *Clim. Res.*, *60*, 249–264.

Lin, S. J., and R. B. Rood (1996), Multidimensional flux-form semi-Lagrangian transport schemes, *Mon. Weather Rev.*, *124*, 2046–2070.

Lin, S. J., and R. B. Rood (1997), An explicit flux-form semi-Lagrangian shallow-water model on the sphere, *Q. J. R. Meteorol. Soc.*, *123*, 2477–2498.

Loh, W.-Y. (2011), Classification and regression trees, *in Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, pp. 14–23, John Wiley & Sons, Inc., Hoboken, N. J.

Martynov, A., R. Laprise, L. Sushama, K. Winger, L. Separovic, and B. Dugas (2013), Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian Regional Climate Model, version 5: Model performance evaluation, *Clim. Dyn.*, *41*, 2973–3005.

Micheas, A. C., I. F. Neil, S. A. Lack, and C. K. Wikle (2007), Cell identification and verification of QPF ensembles using shape analysis techniques, *J. Hydrol.*, *343*, 105–116.

Navarra, A., W. F. Stern, and K. Miyakoda (1994), Reduction of the Gibbs oscillation in spectral model simulations, *J. Clim.*, *7*, 1169–1183.

Neale, R. B., et al. (2010), Description of the NCAR Community Atmosphere Model (*CAM 5.0*), NCAR Technical, *NCAR Tech. Note NCAR/TN-486+STR*, National Center for Atmospheric Research, Boulder, Colo.

Posselt, D. J., A. R. Jongeward, C. Y. Hsu, and G. L. Potter (2012), Object-based evaluation of MERRA cloud physical properties and radiative fluxes during the 1998 El Nino-La Nina transition, *J. Clim.*, *25*, 7313–7327.

Reed, K. A., and C. Jablonowski (2012), Idealized tropical cyclone simulations of intermediate complexity: A test case for AGCMs, *J. Adv. Model. Earth Syst.*, *4*, M04001, doi:10.1029/2011MS000099.

Roe, G. H., and M. B. Baker (2006), Microphysical and geometrical controls on the pattern of orographic precipitation, *J. Atmos. Sci.*, *63*, 861–880, doi:10.1175/JAS3619.1.

Rood, R.B. (1987), Numerical advection algorithms and their role in atmospheric transport and chemistry, *Rev. Geophys.*, *25*, 71–100.

Rood, R. B. (2011), A perspective on the role of the dynamical core in the development of weather and climate models, in *Numerical Techniques for Global Atmospheric Models*, edited by P. Lauritzen et al., pp. 513–537, Springer, Berlin.

Scott, R. K., L. M. Harris, and L. M. Polvani (2015), A test case for the inviscid shallow-water equations on the sphere, *Q. J. R. Meteorol. Soc.*, *142*, 488–495.

Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh (2013), Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, *J. Geophys. Res. Atmos.*, *118*, 1716–1733, doi:10.1002/jgrd.50203.

Skok, G., J. Bacmeister, and J. Tribbia (2013), Analysis of tropical cyclone precipitation using an object-based algorithm, *J. Clim.*, *26*, 2563–2579.

Sylla, M. B., F. Giorgi, E. Coppola, and L. Mariotti (2013), Uncertainties in daily rainfall over Africa: Assessment of gridded observation products and evaluation of a regional climate model simulation, *Int. J. Climatol.*, *33*, 1805–1817.

Wernli, H., M. Paulat, M. Hagen, and C. Frei (2008), SAL-A novel quality measure for the verification of quantitative precipitation forecasts, *Mon. Weather Rev.*, *136*, 4470–4487.

Whitehead J. P., C. Jablonowski, J. Kent, and R. B. Rood (2015), Potential vorticity: Measuring consistency between GCM dynamical cores and tracer advection schemes, *Q. J. R. Meteorol. Soc.*, 688, 739–751.

Wick, G. A., P. J. Neiman, F. M. Ralph, and T. M. Hamill (2013), Evaluation of forecasts of the water vapor signature of atmospheric rivers in operational numerical weather prediction models, *Weather Forecast.*, *28*, 1337–1352.

Williamson, D. L. (2007), The evolution of dynamical cores for global atmospheric models, *J. Meteorol. Soc. Jpn.*, *85B*, 241–269.

Williamson, D. L., and P. J. Rasch (1994), Water-vapor transport in the NCAR-CCM2, *Tellus Ser. A*, *46*, 34–51.

Xu, K. M. (2009), Evaluation of cloud physical properties of ECMWF Analysis and Re-Analysis (ERA) against CERES tropical deep convective cloud object observations, *Mon. Weather Rev.*, *137*, 207–223.

Xu, K. M., T. M. Wong, B. A. Wielicki, L. Parker, and Z. A. Eitzen (2005), Statistical analyses of satellite cloud object data from CERES. Part I: Methodology and preliminary results of the 1998 El nino/2000 La Nina, *J. Clim.*, *18*, 2497–2514.

Yorgun, M. S., and R. B. Rood (2014), An object-based approach for quantification of GCM biases of the simulation of orographic precipitation. Part I: Idealized simulations, *J. Clim.*, *27*, 9139–9154.

Yorgun, M. S., and R. B. Rood (2015), An object-based approach for quantification of GCM biases of the simulation of orographic precipitation. Part II: Quantitative analysis, *J. Clim.*, *28*, 4863–4876.