Title

Evaluation of extreme sub-daily precipitation in high-resolution global climate model simulations

Journal
Philosophical Transactions of the Royal Society A Mathematical Physical and Engineering Sciences, 379(2195)

ISSN
1364-503X

Authors
Wehner, Michael
Lee, Jiwoo
Risser, Mark
et al.

Publication Date
2021-04-19

DOI
10.1098/rsta.2019.0545

Peer reviewed

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

MATHEMATICAL, PHYSICAL AND ENGINEERING SCIENCES

Evaluation of extreme subdaily precipitation in high-resolution global climate model simulations

Journal:	*Philosophical Transactions A*
Manuscript ID	Draft
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Wehner, Michael; Lawrence Berkeley National Laboratory Lee, Jiwoo; Lawrence Livermore National Laboratory Risser, Mark; Lawrence Berkeley National Laboratory Ullrich, Paul; University of California Davis; Lawrence Berkeley National Laboratory Gleckler, Peter; Lawrence Livermore National Laboratory Collins, William; Lawrence Berkeley National Laboratory, Climate Science; University of California Berkeley
Issue Code (this should have already been entered and appear below the blue box, but please contact the Editorial Office if it is not present):	DM02201
Subject:	Climatology < EARTH SCIENCES
Keywords:	extreme precipitation, high-resolution global climate models

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
All HighResMIP model output data is available via the Earth System Grid at https://esgf-node.llnl.gov/search/cmip6/
CAM5.1 model data is availabe via the C20C data portal at https://portal.nersc.gov/c20c/
The NCEP-EMC observed precipitation data is available at https://data.eol.ucar.edu/dataset/21.093

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

Authors' contributions

This paper has multiple authors and our individual contributions were as below

Statement (if applicable):
Wehner conceived and wrote the paper
Lee made the Taylor diagrams
Risser reviewed and added to the discussion
Glecker reviewed and added to the discussion
Ullrich provided remapping software and commentary
Collins downloaded the HighResMIP data and added to the discussion

Evaluation of extreme subdaily precipitation in high-resolution global climate model simulations

Michael Wehner[1*], Jiwoo Lee[2], Mark Risser[1], Paul Ullrich[3,1], Peter Gleckler[2], William D. Collins[1,4]

[1]Lawrence Berkeley National Laboratory
[2]Lawrence Livermore National Laboratory
[3]University of California at Davis
[4]University of California at Berkeley
* Corresponding author. mfwehner@lbl.gov

We examine the resolution dependence of errors in extreme subdaily precipitation in available high-resolution climate models. We find that simulated extreme precipitation increases as horizontal resolution increases but that appropriately constructed model skill metrics do not significantly change. We find little evidence that simulated extreme winter or summer storm processes significantly improve with resolution because the model performance changes identified are consistent with expectations from scale dependence arguments alone. We also discuss the implications of these scale dependent limitations on the interpretation of simulated extreme precipitation.

1) Introduction

Extreme precipitation at sub-daily scales can have significant flooding impacts in both urban and rural environments. Climate change is expected to increase the risk of such impacts as the magnitude of short term extreme precipitation will increase in many regions due to increases in available moisture and energy. Confidence in projection of these future increases as well as the attribution of current changes, if any, requires that climate models both simulate observed subdaily extreme precipitation statistics well and adequately represent the relevant physical processes causing severe storms. Climate models in recent coordinated international projects such as CMIP5 and CMIP6, the 5th and 6th generation of the Coupled Model Intercomparison Project (Eyring et al., 2016; Taylor et al., 2012) are typically configured at effective horizontal grid resolutions of 100km or coarser. For dynamical reasons alone, many properties of the severe storms responsible for extreme precipitation cannot be resolved at these grid spacings, no matter how good the subgrid scale physical parameterizations are (Reed and Jablonowski, 2012; Zarzycki et al., 2014)

Climate models at horizontal resolutions of ~20-50km have been shown to improve upon this situation (Wehner et al., 2014). In particular, the stronger gradients in moisture and temperature enabled at higher resolutions permit reasonable simulation of tropical cyclone properties (Reed et al., 2015; Shaevitz et al., 2014) and other severe storm statistics (Champion et al., 2011; Rhoades et al., 2020; Roberts et al., 2018; Walsh et al., 2015; Wehner et al., 2015). Advances in high performance computing technologies have progressed to the point where a limited number of multi-decadal simulations of climate models at these finer resolutions can now be performed. The multi-tiered HighResMIP subproject of the CMIP6 is the first attempt to intercompare the simulated past climate and projected future climate change of such models (Haarsma et al., 2016). The HighResMIP protocols specify that modeling groups perform simulations with both a coarse and fine resolution model configuration. In practice, the coarse grid configurations are generally the operational version of the model and the high resolution configuration an experimental version with grid spacings of 50km or finer. Hence, the physical parameterizations in the models are specified by the protocols to be the same across resolutions. However, for stability reasons some groups may have had to make minor parameter value adjustments, including time stepping controls. The HighResMIP protocols specify both fully coupled ocean-atmosphere model configurations as well as atmosphere only configurations forced by fixed surface ocean and sea-ice datasets. Simulations of the recent historical period from 1950 to 2014 and a near future period from 2015 to 2050 under the high emissions scenarios of RCP8.5 or SSP85 are called for in both the coupled and atmospheric-only configurations.

To the extent observations permit it, some aspects of subdaily simulated precipitation have been evaluated including the diurnal cycle (Dai et al., 2007). More recently, irregular subdaily fluctuations about the mean diurnal cycle or "intermittency" have been shown to be underestimated by models, even after taking into account the observational "error bars" implied by different space-time resolutions (Covey et al., 2018).

In this paper, we utilize standard practice model evaluation techniques (Gleckler et al., 2008; Lee et al., 2019) to analyze the quality of seasonal 3 hourly precipitation extremes produced by available HighResMIP models. Previous evaluations of simulated extreme precipitation has focused on daily or pentadal accumulations (Akinsanola et al., 2020; Bador et al., 2020; Sillmann et al., 2013; Srivastava et al., 2020; Wehner et al., 2020).

Model evaluation is only as good as the observational datasets used as a reference and quality observed sub-daily precipitation accumulations are even more limited than for daily accumulations (Trenberth et al., 2017). Furthermore, as shown by (Gervais et al., 2014) and explored in this paper, the order of operations in calculating gridded observational extreme subdaily precipitation metrics can affect their magnitude and the interpretation of model quality. In this first evaluation, we thus confine our analyses to the conterminous United States (CONUS) and the winter (DJF) and summer (JJA) seasons.

To date, six modeling groups have submitted both coarse and fine resolution 3 hourly precipitation data to the historical period atmosphere only (*highresSST-present*) experiment. Several of these groups have also submitted simulations to the fully coupled model simulations. Errors in simulated sea surface temperature can significantly affect the location and intensity of severe storms that would likely degrade the quality of simulated extreme precipitation statistics. Hence in this study we focus our model evaluation on the more complete *highresSST-present* experiment and defer analysis of the effect of ocean-atmosphere coupling on extreme precipitation. We also add a seventh model that is not part of the HighResMIP but was integrated under similar boundary conditions.

In section 2, we describe the merged radar and station observational dataset used as an evaluation standard and briefly describe the climate models with available fine and coarse resolution 3 hourly precipitation datasets. We also describe the effect of the order of gridding and extrema on the construction of a model evaluation standard in that section. In section 3, we present the model error metrics including bias maps for each model and summary Taylor diagrams (Taylor, 2001). In section 4, we discuss these errors and offer some interpretation of how model resolution affects the simulation quality of extreme sub-daily precipitation. We further discuss the limitations of simulated extreme precipitation and provide some context supplied by the expectations provided by the model evaluation standards. In section 5, we summarize our principal conclusions about the effect of refined horizontal resolution on simulated sub-daily precipitation quality.

2) Methods, observations and models.

Recognizing that the nature and magnitude of extreme storms in the mid and high latitudes is strongly seasonally dependent, we focus on the winter and summer seasonal extremes rather than on annual extremes. While long period return values of seasonal maxima would be relevant for impacts, we focus only on the average winter and summer maxima as uncertainties from the short observational record in fitted extreme value distributions would be large, even with non-stationary statistical models (Wehner et al 2020a). However, we note that a previous model evaluation of average annual daily maximum precipitation and associated long period return values (Wehner et al., 2020) found that although model performance degrades as rarity increases, the patterns of errors are similar.

Long records of observed subdaily precipitation data are a scarce resource and is available only over limited land regions from weather stations and/or radar. Sampling limitations currently make satellite-based products unsuitable reference data for our analysis. The HadISD (Dunn et al., 2016) is an available multi-variate station data set but precipitation is not one of the variables subjected to stringent quality control. The Global Sub-Daily Rainfall Dataset (GSDR),part of the INTENSE project (Lewis et al., 2019) is the first real attempt to collect and quality control station based sub-daily precipitation. Long, spatial dense records are mostly confined to the United States and some Western European countries. However, this dataset is not yet publicly available.

Operational weather radar provides a remote sensing alternative to ground-based observations. The National Centers for Environmental Prediction (NCEP) Environmental Modeling Center (EMC) has provided a merged ground based and radar derived hourly precipitation dataset from ~3000 weather stations and the 159 Doppler radars of the Next-Generation Radar (NEXRAD) on an approximately 4km polar stereographic grid spanning the CONUS region (Du et al. 2011) and is available at https://data.eol.ucar.edu/dataset/21.087. Most HighResMIP modeling groups provide 3 hourly precipitation accumulations so these hourly observations are similarly accumulated on the original stereographic grid from the raw downloaded data as the first step. We next used the data over the period June 1997 to February 2020 to calculate estimates of the average seasonal maximum 3 hourly precipitation accumulation in two ways discussed below. Note that there are substantial missing data throughout this period, especially prior to 2002.

As mentioned above the order of operations in the construction of the reference observations can introduce false estimation of model biases that are likely to be larger for subdaily than for daily extreme precipitation model performance metrics (Gervais 2013 et al). The observational extreme precipitation product most similar to model output is obtained by first gridding the raw high frequency data to the model grid then calculating block extrema, usually annually or seasonally. As precipitation is by definition a moisture flux, this procedure should be made conservative. In this paper, we refer to such model evaluation results as the “native grid” results since the observational extremes are calculated on each models’ native grid. In this case, a different reference set must be calculated for each model further adding to the complexity of the evaluation process.

However, this order of operation is not always practical, especially for subdaily extremes due to the high computational cost of regridding and/or the availability of the high frequency observational data itself. For instance, high frequency station data may not be made available by the owners but block maxima or other extreme value indices are provided. In fact, this is the case for the daily extrema contained in the HadEX3 global land dataset (Dunn, 2020) where the stations’ extrema are gridded rather than the stations’ daily values themselves. In this paper, we refer to such model evaluation results as the “non-native grid” since the observational extremes are not calculated on the models’ native grid but are calculated either at individual stations or on a different grid. For precipitation, it is generally unlikely that the extrema at different locations within the same grid cell occur at the same time. Hence, observational estimates of gridded station extrema are generally larger than the extrema of gridded high frequency station data. This order of operations bias also extends to the case where the observations are on a much finer grid than the models, as is the case here with the NCEP-EMC hybrid radar station product. Figures 1 and 2 shows the observed average winter (1a) and summer (2a) maximum 3 hour precipitation accumulations calculated on the original 4km polar stereographic grid but regridded to a 4km latitude/longitude for plotting purposes. Also shown in figure 1 are “native grid” results at 25km (1b, 2b) and 100km (1c, 2c) and the “non-native grid” results at 25km (1d, 2d) and 100km (1e, 2e). Clearly, the non-native regridding shown in the bottom rows result in values close to the original 4km resolution. However, as Gervais et al., (2014) point out, the smaller values produced by the native mesh regridding shown in the top rows are what the models should be expected to produce. Figure S1 shows the percent differences between the native and non-native gridding results further revealing that the effect of the order of operations is larger for lower resolutions than higher resolutions. In the next section we show the effect of this order of regridding operations on model evaluation. Herein we use the conservative and consistent TempestRemap package for regridding operations (Ullrich et al., 2016; Ullrich and Taylor, 2015).

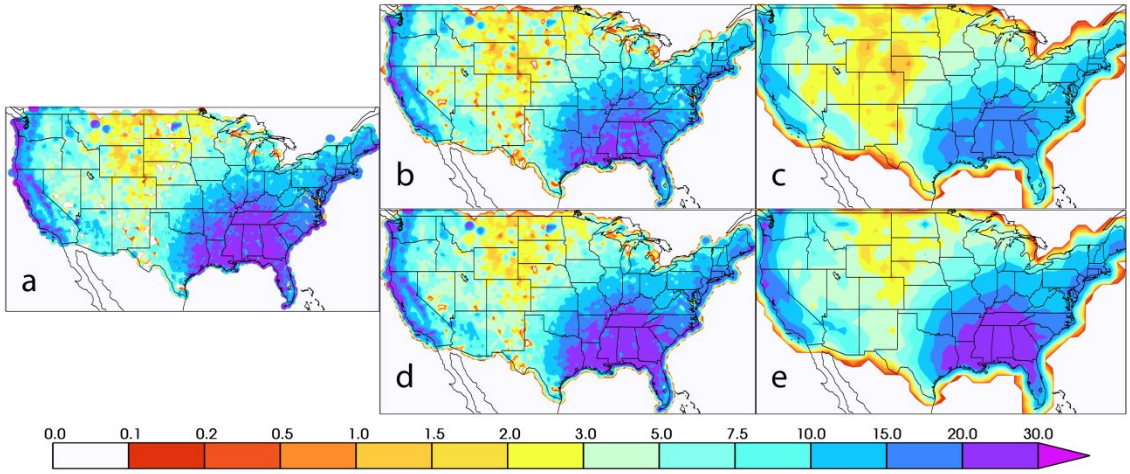

Figure 1. Average DJF maximum 3 hour precipitation accumulation. a) Maximum values calculated on the original 4km polar stereographic mesh and regridded to a 4km latitude-longitude mesh. b) Maximum values obtained by first regridding daily precipitation to a 25km mesh. c) Maximum values obtained by first regridding daily precipitation to a 100km mesh. d) Maximum values obtained by regridding 4km maxima to a 25km mesh. e) Maximum values obtained by regridding 4km maxima to a 100km mesh.

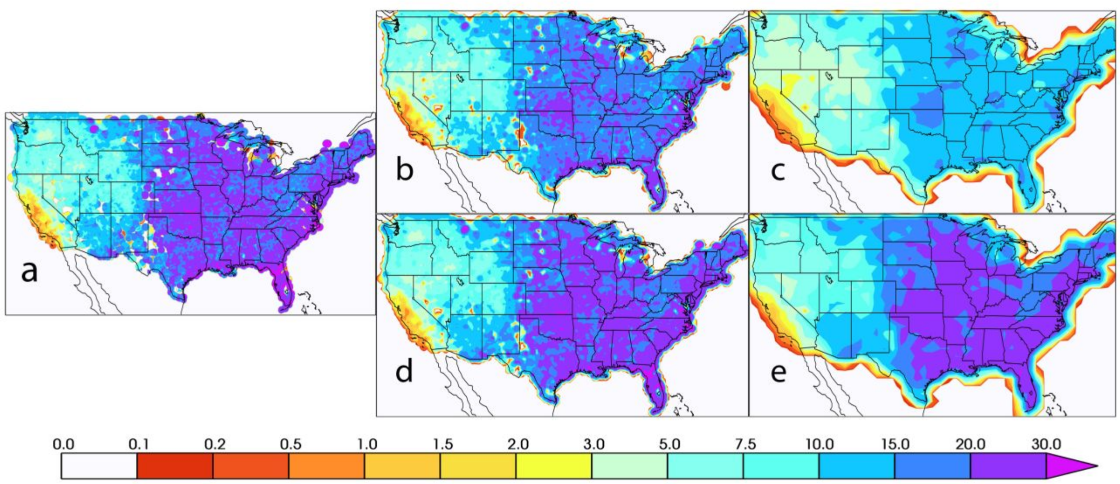

Figure 2. Average JJA maximum 3 hour precipitation accumulation. a) Maximum values calculated on the original 4km polar stereographic mesh and regridded to a 4km latitude-longitude mesh. b) Maximum values obtained by first regridding daily precipitation to a 25km mesh. c) Maximum values obtained by first regridding daily precipitation to a 100km mesh. d) Maximum values obtained by regridding 4km maxima to a 25km mesh. e) Maximum values obtained by regridding 4km maxima to a 100km mesh.

The CNRM-CM6-1 models were developed at the Centre National de Recherches Meteorologiques and the Centre Europeen de Recherche et de Formation Avancee en Calcul Scientifique in Toulouse, France (Voldoire et al., 2013). The EC-Earth3P models were developed by a consortium of European universities and laboratories (http://www.ec-earth.org) from Belgium, Denmark, Finland, Germany, Ireland, Italy, The Netherlands, Norway, Portugal, Spain, Sweden and the United Kingdom (Haarsma et al., 2016) and is based on the European Centre for Medium Range Forecasting IFS seasonal forecasting system (C. D. Roberts et al., 2018). The HadGEM3-GC3.1 is the current version of the United Kingdom's MetOffice Unified Model (Roberts et al., 2019) and results were supplied at three resolutions. The IPSL-CM6A models were developed at the Institut Pierre Simon Laplace in Paris, France (Boucher et al., 2019). The MRI-AGCM3-2 models were developed at the Max Planck Institute for Meteorology in Hamburg, Germany (Gutjahr et al., 2019). The NICAM16 models are based non-hydrostatic equations and were developed at multiple institutions in Yokohama, Tokyo and Tsukuba, Japan (Kodama et al., 2020). Additionally, we also evaluate the Community Atmospheric Model (CAM5.1), developed at the National Center for Atmospheric Research in Boulder, Colorado, United States (Bacmeister et al., 2014; Wehner et al., 2014). While this model was not submitted to the HighResMIP subproject, it was integrated under similar boundary conditions for the period evaluated. The model names and their provided latitude and longitude dimensions are listed in table 1. However, models' true native grids may not be based on a latitude-longitude coordinate system and submitted data is regridded according to CMIP6 protocols in such cases. Interested readers are directed to the cited model documentation.

The *highresSST-present* simulations nominally end in 2014 although some models end in 2015. For these simulations, we average the seasonal maxima over the 20 year period 1995 to 2014. The CAM5.1 model data is available only from 1996 to 2015, so we average over that 20 year period instead. While these periods are not identical to the observed period used here, the length of period is about the same when accounting for the missing observations. While any anthropogenic trend in extreme precipitation from 2014 to 2020 is negligible, we admit that some differences due to natural modes of sea surface temperature variability might not be. However, Risser et al. (2020) find that percentage of variance in extreme daily precipitation over the CONUS region explained by these modes is smaller than might be expected. Some models provide multiple realizations (table 1) and in these cases, the seasonal maxima are further averaged over realizations.

Model	latitude X longitude	# of realizations	DJF skill	JJA skill
CAM5-1-1degree	128x256	49	0.75	0.54
CAM5-1-2-025degree	**360x720**	**5**	**0.79**	**0.60**
CNRM-CM6-1	256x512	1	0.67	0.65
CNRM-CM6-1-HR	**512x1024**	**1**	**0.76**	**0.68**
EC-Earth3P	144x192	3	0.73	0.68
EC-Earth3P-HR	**324x432**	**3**	**0.74**	**0.69**
HadGEM3-GC31-LM	143x144	2	0.68	0.53
HadGEM3-GC31-MM	361x512	2	0.74	0.52
HadGEM3-GC31-HM	**768x1024**	**3**	**0.80**	**0.54**
IPSL-CM6A-LR	320x640	1	0.70	0.52
IPSL-CM6A-ATM-HR	**960x1920**	**1**	**0.73**	**0.60**
MRI-AGCM3-2-H	320x640	1	0.82	0.57
MRI-AGCM3-2-S	**640x1280**	**1**	**0.82**	**0.56**
NICAM16-7S	192x288	1	0.77	0.46
NICAM16-8S	**768x1152**	**1**	**0.74**	**0.32**

Table 1. Model resolution (column 2) and the number of realizations used in evaluation (column 3). Taylor's modified skill over the CONUS region for average maximum DJF and JJA 3 hourly precipitation using subdaily observations regridded to the models' resolutions is shown in columns 5 and 6. High resolution model versions are shown in bold font.

3) Model errors.

Figure 3 shows the percent error in simulated average DJF maximum 3 hour precipitation accumulation using the native grid observations. Consistent with the expectation shown in the top row of figure 1, the high resolution models produce larger maximum values than their lower resolution counterparts. Although there is little commonality between errors across the modeling groups, the pattern of native grid errors are remarkably similar across resolution within an individual modeling group. Figure S2 shows the percent error in simulated average DJF maximum 3 hour precipitation accumulation using the non-native grid observations and highlights the importance of the order of operations in the constructing the reference maxima. Non-native grid model errors are very different than native grid errors since the non-native reference values, shown in the bottom row of figure 1, are so much larger than the native reference values of the top row of figure 1. The patterns of non-native grid model errors in figure S2 are much less similar across resolutions than the native grid model errors in figure 3 and can even be of opposite sign as summarized in table S1. This difference in error pattern can lead to very difference conclusions about the effect of horizontal resolution on simulated extreme precipitation quality.

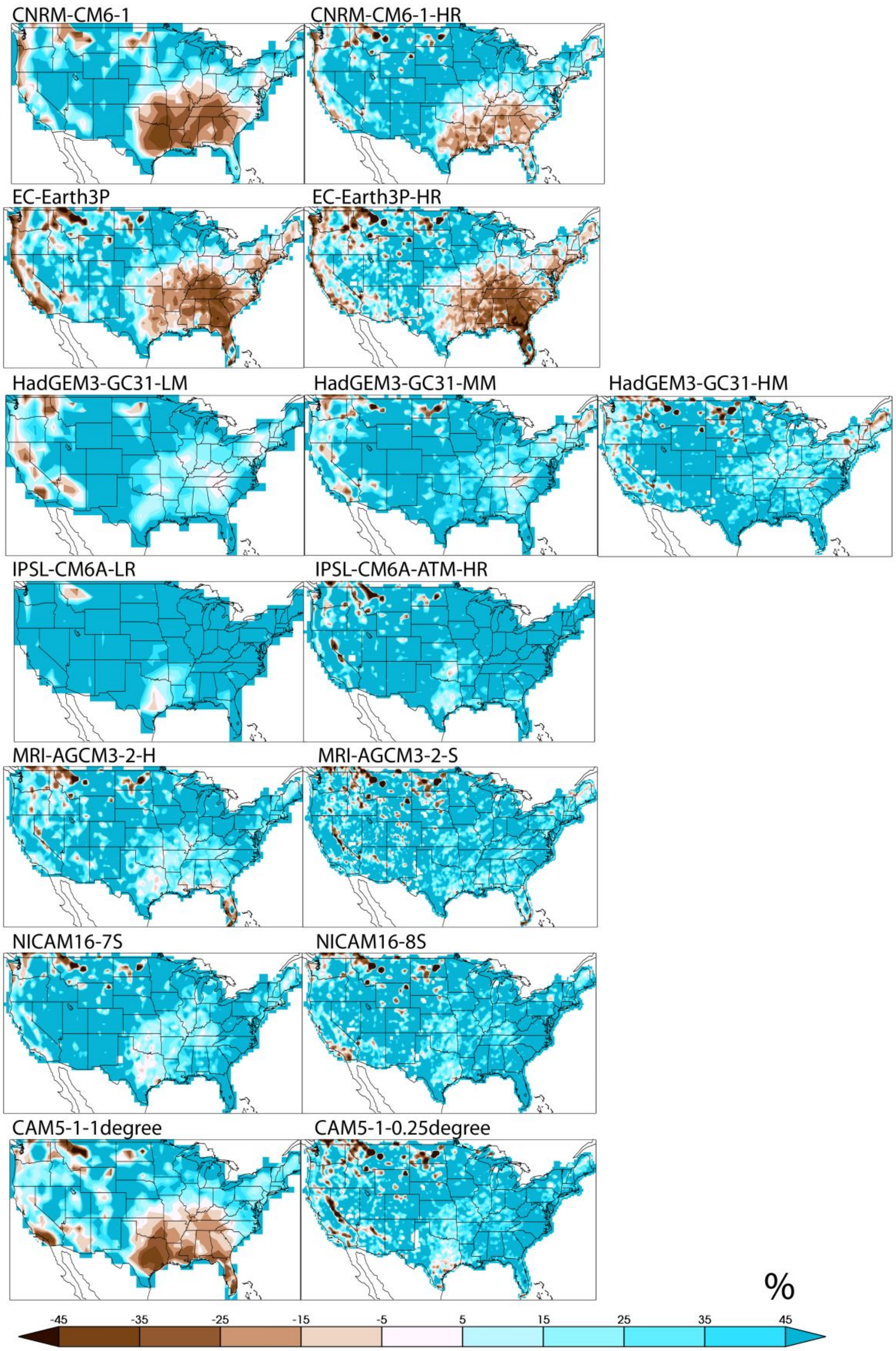

Figure 3. Percent native grid error in simulated average DJF maximum 3 hour precipitation accumulation. Models are arranged low to high horizontal resolution from left to right.

Figure 4 shows the percent error in simulated average JJA maximum 3 hour precipitation accumulation using the native grid observations. Summer errors are generally considerably larger than winter errors at any resolution. Error patterns are again very different across models and for some models similar across resolutions. The notable exception to similarity across resolutions is the CAM5.1 model in the southeastern US in both seasons. Figure S3 and Table S1 show that the non-native grid errors in summer are even more different from the native grid errors than in winter.

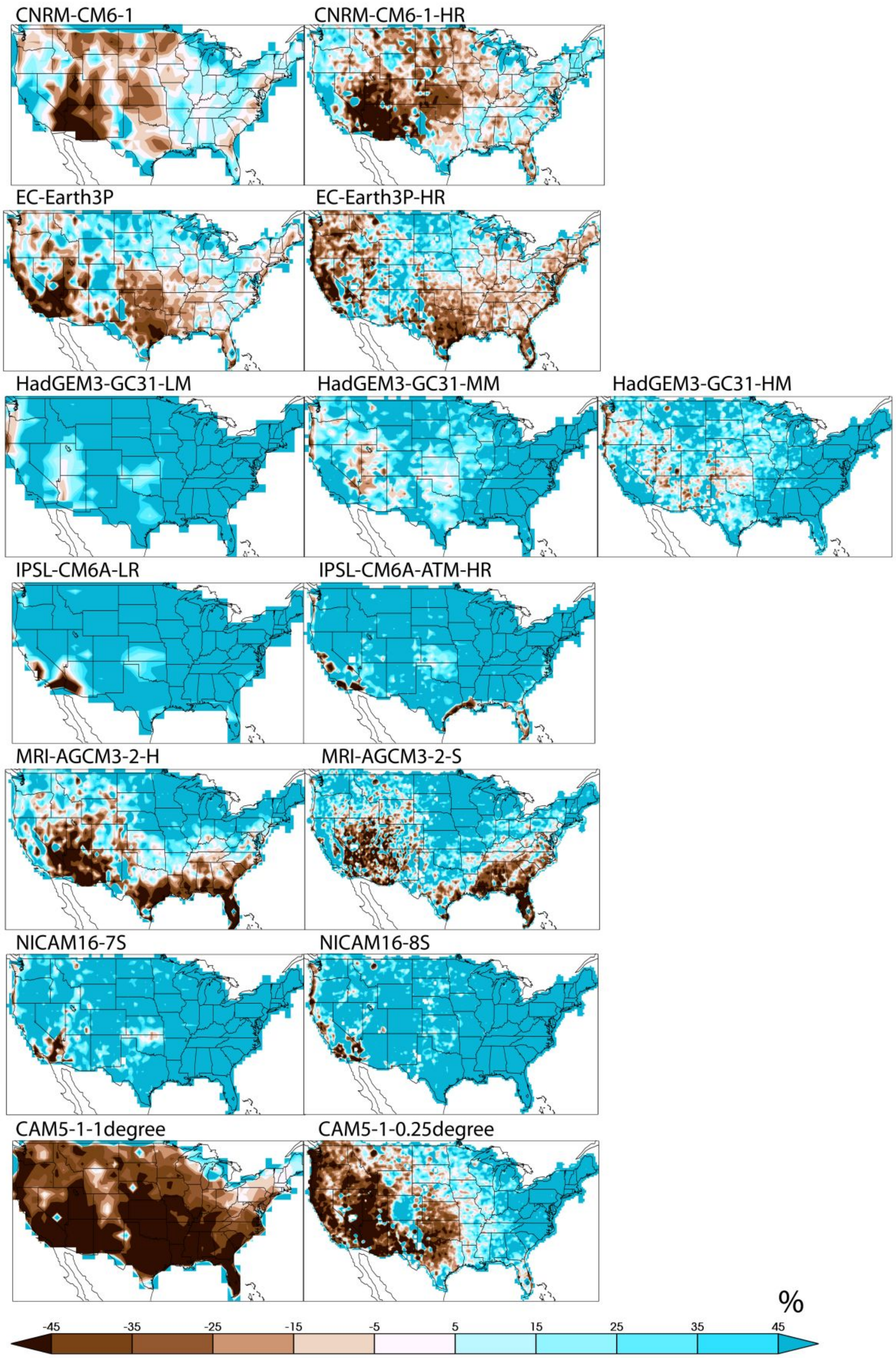

Figure 4. Percent native grid error in simulated average JJA maximum 3 hour precipitation accumulation. Models are arranged low to high horizontal resolution from left to right.

We use Taylor Diagrams to compare differences in selected centered statistics. Figure 5 shows Taylor Diagrams of the pattern correlation and normalized spatial standard deviation of the simulated seasonal maximum 3 hour precipitation accumulation for winter (left column) and summer (right column) using both the native (top row) and non-native (bottom row) grid observations. Traditionally, Taylor Diagrams show the different model results computed on a common grid, but here centered statistics are calculated on each individual model's grid to be consistent with the resolution dependent bias statistics. To facilitate comparison, we normalize the standard deviation of each model result by the corresponding value of the observed extremes. Consistent with the similarity in DJF native grid error structures across resolution shown in figure 3, there is little difference in the locations of symbols for the high (red) and low (blue) resolution simulations of a given modeling group shown in the upper left of figure 5. Despite the dissimilar winter error patterns across models shown in figure 3, points in the Taylor diagram are clustered in the angular dimension with centered pattern correlations between 0.6 and 0.8. However, there is considerable spread in the normalized spatial standard deviation indicating that the range of maximum values varies significantly across the CONUS region across models. Using the native grid observations, normalized Root Mean Square Error (RMSE) ranges from 0.6 to about 1.0 in the winter. Taylor's modified skill (Wehner, 2013) in winter using the native grid reference, shown in table 1, ranges from 0.67 to 0.82 with generally small differences between simulations from the same modeling group.

The Taylor diagram of JJA native grid errors (upper right of figure 5) also shows a tight cluster around pattern correlation values between 0.5 and 0.6 (the angular dimension of the diagram) but a much larger spread in the radial dimension indicating a wider spatial dynamic range across modeling groups. As in winter, there is little difference in the placement of symbols across simulations from a given model group. Summer normalized RMSE is larger than winter ranging from about 0.9 to 2.5 and Taylor's modified skill (Table 1) ranges from 0.32 to 0.69. Differences between simulations from the same modeling group again are small with the exception of the outlying NICAM16 models which exhibit exceptionally large simulated JJA 3 hour maximum precipitation accumulations.

Perhaps surprisingly, considering the difference in model biases between the native and non-native grid standards, there is little corresponding difference in the Taylor diagrams. Pattern correlation metrics are essentially the same and any differences come from the normalized standard deviation.

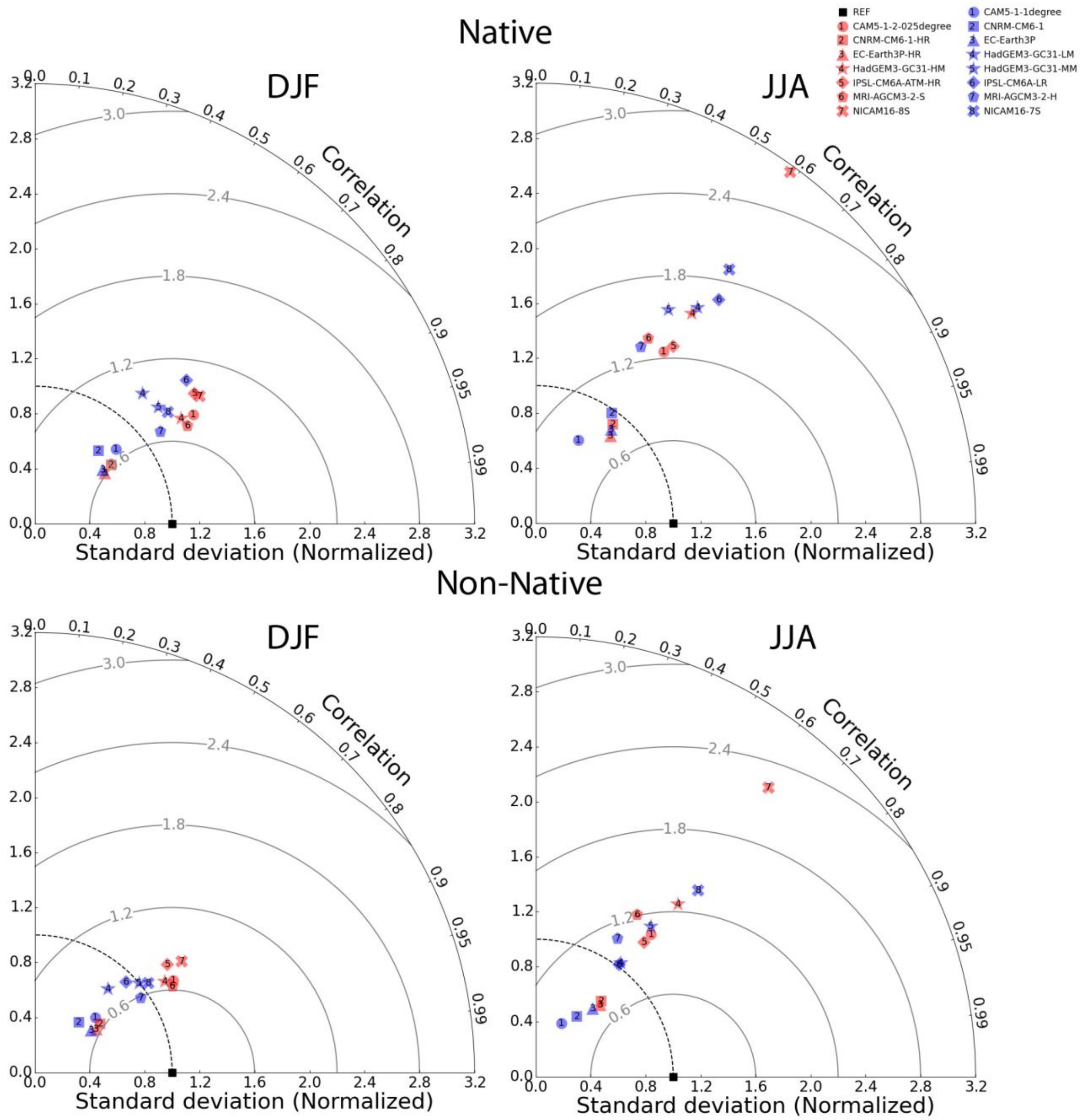

Figure 5. Taylor diagram of average DJF (right) and JJA (left) maximum simulated 3 hour precipitation accumulation from native (top) and non-native (bottom) grid errors. Blue symbols are low resolution models. Red symbols are high resolution models. Symbol shapes are the same for models from the same modeling group. The concentric semi-circles are isolines of normalized root mean square error. The dashed circle represents a normalized standard deviation of unity.

4) Discussion

Part of the motivation for increasing climate models' horizontal resolution to a few 10's of kilometers is to more realistically simulate the severe storms responsible for extreme precipitation. And indeed simulated seasonal maximum sub-daily precipitation accumulations increase with refined computational grids. The extrema based on appropriately coarsened sub-daily observations provide us the appropriate standard reference for model evaluation (upper panels of Figures 1 and 2). Based on that standard, we find little improvement with grid refinement in simulated 3 hour extreme precipitation accumulations when held to that expectation, at least over the CONUS region (figures 3-5).

The difference in the resolution dependent standards in the upper panels of Figures 1 and 2 provide an expectation of the increase in simulated extreme precipitation with resolution. Figure 6 shows the expected percent change in simulated average seasonal maximum 3 hour precipitation accumulation for a change in model horizontal resolution from ~100km to ~25km. In this figure, the ~100km standard was conservatively remapped to ~25km and is used in the denominator. This expectation, based on simple scaling arguments, is mostly of an increase. Decreases are mostly localized and confined to dry regions in areas of high orography. The shortness of record, combined with high variability in these regions is the most likely explanation for these decreases, rather than deficiency in the scaling argument.

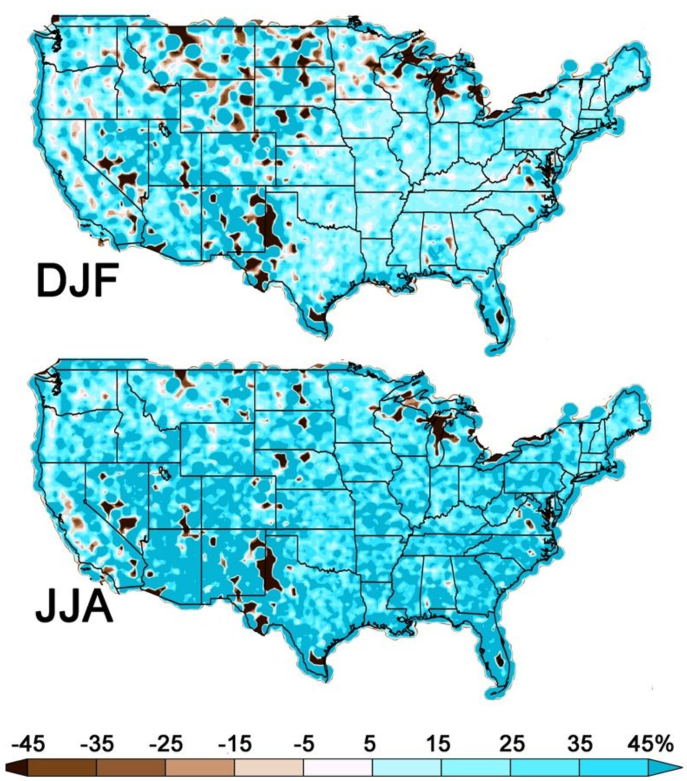

Figure 6: The expected percent change in simulated average seasonal maximum 3 hour precipitation accumulation for a change in horizontal resolution from ~100km to ~25km.

We find that high resolution models generally exhibit similar patterns in percent errors to their low resolution counterparts. The similarity in sub-daily precipitation errors across resolution suggests that large scale circulation errors are not affected much by resolution. It also suggests that the locations and frequency of winter and summer extreme storms resulting from the simulated large scale circulation are also not greatly affected by resolution although that aspect of the HighResMIP simulations has not yet been evaluated. While the magnitude of extreme storms are substantially larger and hence more realistic at high resolution, the "native grid" method of defining an extreme precipitation standard accounts for this and little resolution dependence can be robustly identified in percent error magnitude when models are evaluated against that standard. Further evidence that resolution has little effect on extreme precipitation beyond what is expected by figures 1 and 2 is provided by the normalized error metrics of the Taylor diagrams (figure 5). The distance between points representing models of different resolution from the same modeling groups is small and both normalized RMSE and Taylor's modified skill (table 1) exhibit only minor improvements at high resolution. In fact, skill values of extreme 3 hour precipitation accumulations over the CONUS region for the HighResMIP models are quite similar to global land skill values for extreme daily precipitation accumulations for the CMIP5 and CMIP6 models (Wehner et al., 2020). Indeed, as extreme daily precipitation errors are highly correlated to mean precipitation errors (Wehner et al., 2020) it is not surprising that sub-daily errors would be closely related to daily errors.

There is no systematic error pattern across all modeling groups, for all but one pair of models with an ancestral relationship. The CNRM-CM6-1 models and EC-Earth3P models both descend from versions of the ECMWF IFS atmospheric model and have very similar winter error patterns although differ in the summer. The other models, except CAM5.1, are largely biased high in the winter but mixed in the sign of summer errors.

Model evaluation is only as good as the reference data available and for sub-daily precipitation there are many obstacles to constructing them from in-situ and remote observations for sub-daily precipitation. While the NCEP-EMC hybrid station and radar data set is relatively short at about 20 years, the differences between it and the model simulations are likely much larger than natural variability, even for extreme sub-daily precipitation accumulations. Other data sets covering larger regions and longer time periods constructed by gridding station extrema are now becoming available (Dunn et al., 2020; Lewis et al., 2019). While gridded station extrema are useful for assessing the actual risk of extreme precipitation, they are inappropriate for evaluation of simulated extreme precipitation bias due to fundamental discrepancies in their definition relative to model representation. Climate model precipitation within a grid is best thought of as a moisture flux and is a conserved quantity in a well-constructed climate model. While available sub-daily *in situ* station or radar measurements within a computational grid cell may be sparse, placing them on a grid at the same frequency as sampled by the model and subsequently calculating maxima most closely resembles what models simulate. Clearly, gridded station maxima is a different quantity than the maxima of gridded high frequency precipitation and has no conservative properties. This same statement holds true for remapping very finely gridded observational maxima to a coarser grid. This is most clear by recognizing that within a grid cell, not all locations will experience the maximum precipitation accumulation at the same time. Hence, the gridded maxima is always larger than the maxima of gridded high frequency precipitation. This effect is exacerbated as grids coarsen as shown in figures 1, 2 and S1.

This inconsistency between gridded maxima and what climate models actually simulate presents a challenge to comprehensive model evaluation. However, as figure 5 shows, a standard based on gridded maxima does provide useful information about the patterns of errors. Normalized RMSE and Taylor's modified skill from such a standard are biased but not as much as might be expected from the biases in error magnitude. This behavior will prove useful in a limited model evaluation over a larger fraction of the planet when observed gridded maxima products such as from the INTENSE project become available.

We must point out that gridded extrema are indeed useful for other purposes, if not for the evaluation of the magnitude of model bias. For if one requires the risk of extreme precipitation at a point, long period return values calculated from some variant of gridded extrema, preferably borrowing strength using spatial statistics, is the most credible estimate (Risser et al., 2018) as spatial smoothing damps some of the sampling variability.

There are important ramifications for the interpretation of simulated extreme precipitation from the reduced expectations in the upper panels of figures 1 and 2. First and foremost, return values or periods as calculated from climate models are not to be interpreted as representing the probability at a point of a specified extreme value. Although beyond the scope of this paper, they may be ways to utilize the top and bottom rows of figure 1 and 2 to bias correct the grid effect. Whether these errors cancel when inferring changes in the future probability of extreme precipitation from simulated return periods differences (Collins et al., 2013), remains an open question.

If the models simulate extreme precipitation statistics at values close to those from station or radar data (i.e. the lower panels of figures 1 and 2), then they are actually biased high. Also in extreme event attribution studies (e.g. van Oldenborgh et al., 2017), models are often queried about the probability of a rare event of a given observed magnitude. However, comparison of climate model precipitation return values to the station or radar values describing a rare event leads to an overestimation of event probability in an unbiased model, even if the observations are placed on the same grid.

Trends in average and extreme precipitation are usually presented either as absolute or percent changes from a reference period. Maps of absolute changes tend to highlight wet areas, while maps of percent changes tend to highlight dry areas. Because of these reduced expectations, absolute changes in extreme precipitation of a given return period obtained from climate models would also be low in an unbiased model, if that return period is to be interpreted as a probability at a given point or region. The magnitude of the reduced expectations from the high frequency gridding is likely a function of the rarity of the extreme precipitation considered. This would also introduce biases in point-wise probability changes interpreted from simulated percent changes in long period return values, although the magnitude of these errors would depend on how strong a function of rarity the reduced expectations are.

Likewise, similar caveats should be recognized in formal Detection and Attribution (D&A) analyses of observed trends in extreme precipitation (Min et al., 2011; Zhang et al., 2013). “Scaling factors” are a ratio of the observed to simulated trends and are tested against zero to infer causality in many D&A approaches. If observations are based on gridded extrema (lower panels of figures 1 and 2) and the climate models are unbiased, lower bounds of scaling factors of absolute extreme precipitation trends would be overestimated, possibly leading to erroneous causal inference.

5) Conclusion

Increasing global atmospheric model horizontal resolution increases the magnitude of simulated extreme sub-daily precipitation. In that sense, resolution increases are an important step towards more realistic estimation of their behavior. However, the expected magnitude of simulated extreme precipitation from scaling arguments is a strong function of resolution and when held against this standard, we find improvements in simulation quality to be nominal. In principle, horizontal resolution increases should improve the representation of extreme storms, and in actual practice, they do with tropical cyclones being a well-studied case in point. Hence, the lack of substantial improvement in the quality of simulated extreme sub-daily precipitation is puzzling, at least in winter and summer. Model errors in summer are larger than in winter, suggesting that parameterization of cumulus convection plays a role in these errors. However, even at high resolutions, most of the models examined herein are significantly too wet in the winter, suggesting that moisture transport errors also play an important role.

Acknowledgements

The research performed at Lawrence Berkeley National Laboratory was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract No. DE340AC02-05CH11231 and its ICoM and CASCADE Scientific Focus Areas. The research performed at Lawrence Livermore National Laboratory was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract No. DE-AC52-07NA27344. Support from the Regional and Global Climate Modeling program is gratefully acknowledged. This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

This work used resources at the National Energy Research Supercomputer Center (NERSC) at the Lawrence Berkeley National Laboratory. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP5 and CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

Bibliography

Supplement

Figure S1 shows the percent difference between native and non-native constructions of NCEP-EMC average seasonal maximum 3 hour precipitation accumulation revealing the strong dependence of resolution on the order of operations in constructing extrema. Figures S2 and S3 show the non-native grid percent error for simulated average winter and summer maximum 3 hour precipitation accumulation. Conclusions about changes in model performance would be considerably different than those drawn from the errors constructed on the native grids in figures 3 and 4.

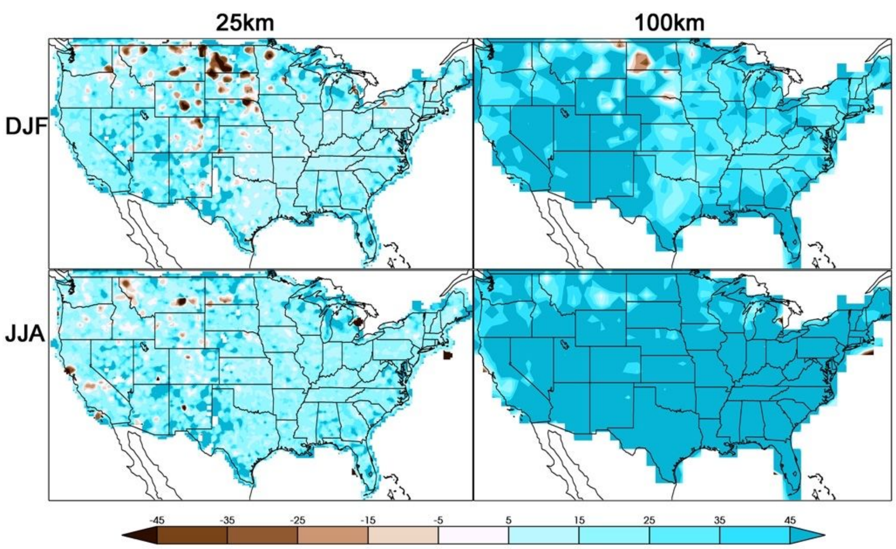

Figure S1. Percent difference between native and non-native constructions of NCEP-EMC average seasonal maximum 3 hour precipitation accumulation. Top row: DJF. Bottom row: JJA. Left column: ~25km grid. Right Column: ~100km grid.

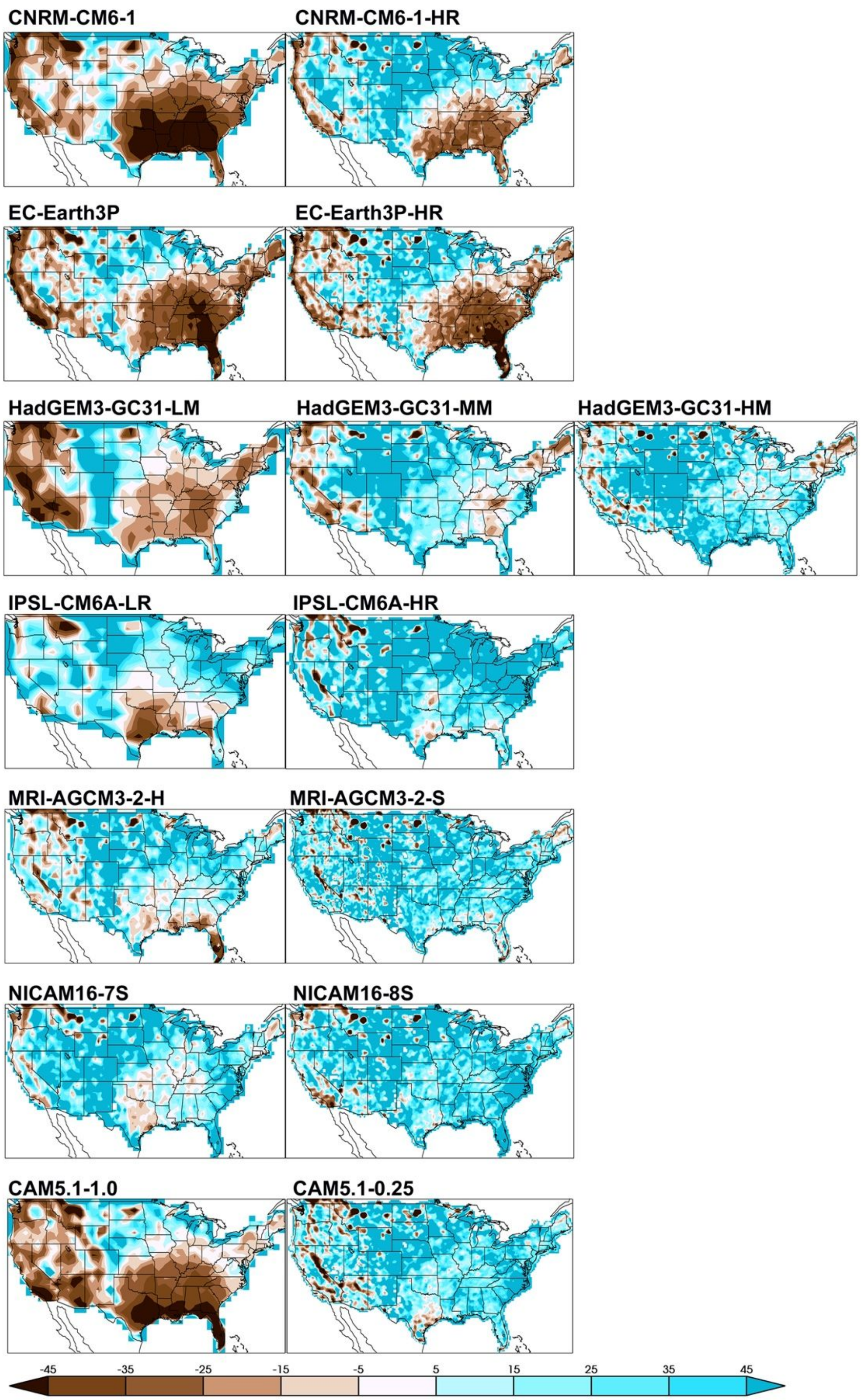

Figure S2. Percent non-native grid error in simulated average DJF maximum 3 hour precipitation accumulation. Models are arranged low to high horizontal resolution from left to right.

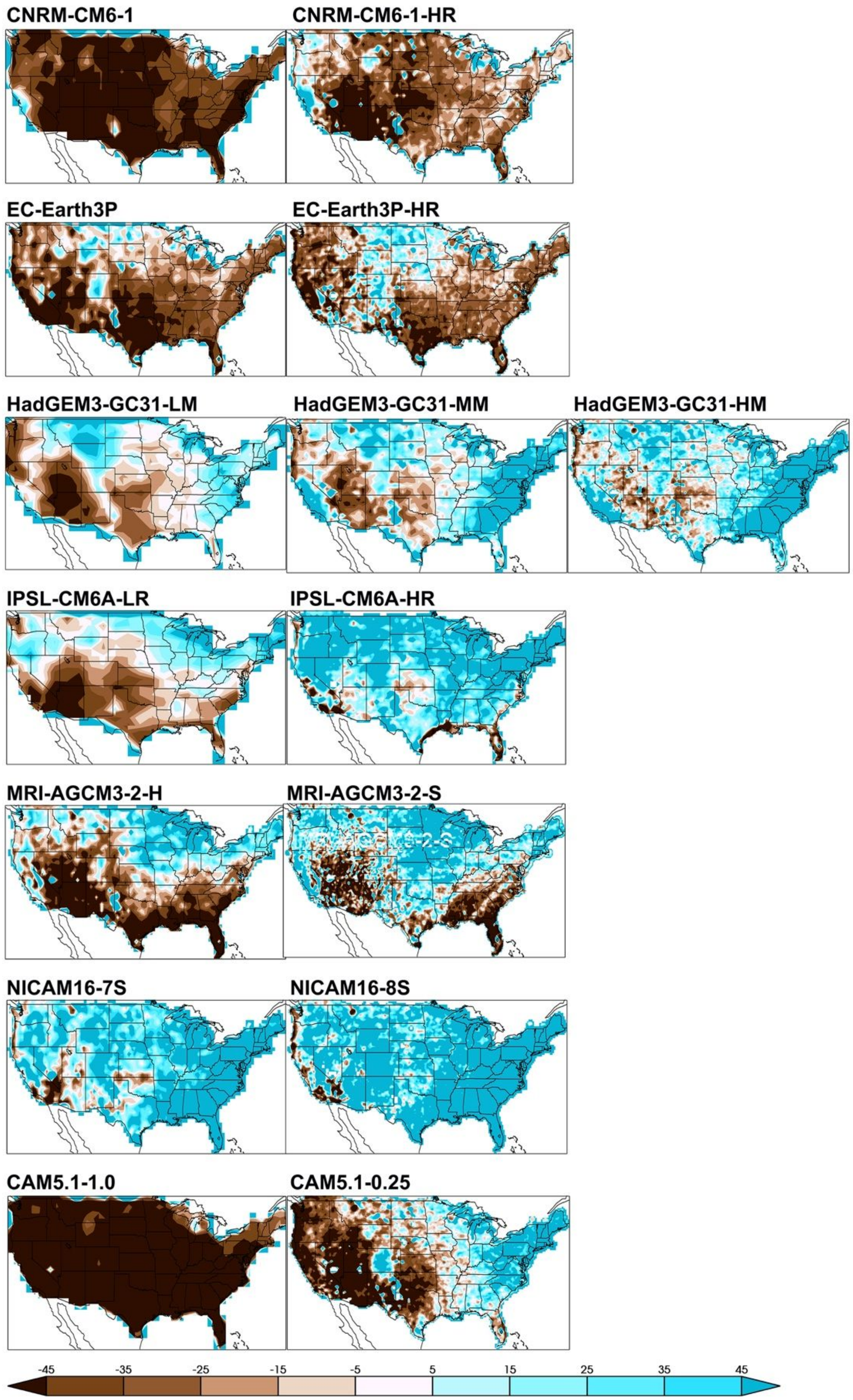

Figure S3. Percent non-native grid error in simulated average JJA maximum 3 hour precipitation accumulation. Models are arranged low to high horizontal resolution from left to right.

Model	latitude X longitude	DJF Non-native	DJF Native	JJA Non-native	JJA Native
CAM5-1-1degree	128x256	-18%	16%	-59%	-32%
CAM5-1-2-025degree	**360x720**	**29%**	**50%**	**4%**	**25%**
CNRM-CM6-1	256x512	-20%	20%	-39%	13%
CNRM-CM6-1-HR	**512x1024**	**2%**	**23%**	**-23%**	**1%**
EC-Earth3P	144x192	-19%	4%	-25%	6%
EC-Earth3P-HR	**324x432**	**-14%**	**1%**	**-19%**	**0%**
HadGEM3-GC31-LM	143x144	1%	58%	13%	118%
HadGEM3-GC31-MM	361x512	22%	57%	28%	82%
HadGEM3-GC31-HM	**768x1024**	**30%**	**49%**	**39%**	**68%**
IPSL-CM6A-LR	320x640	23%	105%	6%	126%
IPSL-CM6A-ATM-HR	**960x1920**	**46%**	**80%**	**44%**	**94%**
MRI-AGCM3-2-H	320x640	16%	43%	-6%	26%
MRI-AGCM3-2-S	**640x1280**	**32%**	**49%**	**16%**	**37%**
NICAM16-7S	192x288	27%	56%	63%	118%
NICAM16-8S	**768x1152**	**40%**	**61%**	**109%**	**152%**

Table S1. Spatially averaged percent error over the CONUS region for average maximum DJF and JJA 3 hourly precipitation using observed maxima regridded to the models' resolution (3rd and 5th columns labeled "non-native") and using subdaily observations regridded to the models' resolutions (4th and 6th columns labeled "native"). High resolution model versions are shown in bold font.

Akinsanola, A.A., Kooperman, G.J., Pendergrass, A.G., Hannah, W.M., Reed, K.A., 2020. Seasonal representation of extreme precipitation indices over the United States in CMIP6 present-day simulations. Environ. Res. Lett. 15, 94003. https://doi.org/10.1088/1748-9326/ab92c1

Bacmeister, J.T., Wehner, M.F., Neale, R.B., Gettelman, A., Hannay, C., Lauritzen, P.H., Caron, J.M., Truesdale, J.E., 2014. Exploratory high-resolution climate simulations using the community atmosphere model (CAM). J. Clim. 27. https://doi.org/10.1175/JCLI-D-13-00387.1

Bador, M., Boé, J., Terray, L., Alexander, L. V, Baker, A., Bellucci, A., Haarsma, R., Koenigk, T., Moine, M.-P., Lohmann, K., Putrasahan, D.A., Roberts, C., Roberts, M., Scoccimarro, E., Schiemann, R., Seddon, J., Senan, R., Valcke, S., Vanniere, B., 2020. Impact of higher spatial atmospheric resolution on precipitation extremes over land in global climate models. J. Geophys. Res. Atmos. n/a, e2019JD032184. https://doi.org/10.1029/2019JD032184

Boucher, O., Denvil, S., Caubel, A., Foujols, M.A., 2019. IPSL IPSL-CM6A-ATM-HR model output prepared for CMIP6 HighResMIP. https://doi.org/10.22033/ESGF/CMIP6.2361

Champion, A.J., HODGES, K.I., BENGTSSON, L.O., KEENLYSIDE, N.S., ESCH, M., 2011. Impact of increasing resolution and a warmer climate on extreme weather from Northern Hemisphere extratropical cyclones. Tellus A 63, 893–906. https://doi.org/10.1111/j.1600-0870.2011.00538.x

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W.J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A.J., Wehner, M., 2013. Long-term Climate Change: Projections, Commitments and Irreversibility, in: Intergovernmental Panel on Climate Change (Ed.), Climate Change 2013 - The Physical Science Basis. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1029–1136. https://doi.org/10.1017/CBO9781107415324.024

Covey, C., Doutriaux, C., Gleckler, P.J., Taylor, K.E., Trenberth, K.E., Zhang, Y., 2018. High-Frequency Intermittency in Observed and Model-Simulated Precipitation. Geophys. Res. Lett. 45, 12,512-514,522. https://doi.org/10.1029/2018GL078926

Dai, A., Lin, X., Hsu, K.-L., 2007. The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low- and mid-latitudes. Clim. Dyn. 29, 727–744. https://doi.org/10.1007/s00382-007-0260-y

Du, J. 2011. GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data. Version 1.0. UCAR/NCAR - Earth Observing Laboratory. https://doi.org/10.5065/D6PG1QDD. Accessed 21 Aug 2020.

Dunn, R., 2020. Development of an updated global land in-situ-based dataset of temperature and precipitation extremes: HadEX3. J. Atmos. Sci.

Dunn, R.J.H., Alexander, L. V, Donat, M.G., Zhang, X., Bador, M., Herold, N., Lippmann, T., Allan, R., Aguilar, E., Barry, A.A., Brunet, M., Caesar, J., Chagnaud, G., Cheng, V., Cinco, T., Durre, I., de Guzman, R., Htay, T.M., Wan Ibadullah, W.M., Bin Ibrahim, M.K.I., Khoshkam, M., Kruger, A., Kubota, H., Leng, T.W., Lim, G., Li-Sha, L., Marengo, J., Mbatha, S., McGree, S., Menne, M., de los Milagros Skansi, M., Ngwenya, S., Nkrumah, F., Oonariya, C., Pabon-Caicedo, J.D., Panthou, G., Pham, C., Rahimzadeh, F., Ramos, A., Salgado, E., Salinger, J., Sané, Y., Sopaheluwakan, A., Srivastava, A., Sun, Y., Timbal, B., Trachow, N., Trewin, B., van der Schrier, G., Vazquez-Aguirre, J., Vasquez, R., Villarroel, C., Vincent, L., Vischel, T., Vose, R., Bin Hj Yussof, M.N., 2020. Development of an Updated Global Land In Situ-Based Data Set of Temperature and Precipitation Extremes: HadEX3. J. Geophys. Res. Atmos. 125, e2019JD032263. https://doi.org/10.1029/2019JD032263

Dunn, R.J.H., Willett, K.M., Parker, D.E., Mitchell, L., 2016. Expanding HadISD: quality-controlled, sub-daily station data from 1931. Geosci. Instrumentation, Methods Data Syst. 5, 473–491. https://doi.org/10.5194/gi-5-473-2016

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. 9, 1937–

1958. https://doi.org/10.5194/gmd-9-1937-2016

Gervais, M., Tremblay, L.B., Gyakum, J.R., Atallah, E., 2014. Representing Extremes in a Daily Gridded Precipitation Analysis over the United States: Impacts of Station Density, Resolution, and Gridding Methods. J. Clim. 27, 5201–5218. https://doi.org/10.1175/JCLI-D-13-00319.1

Gleckler, P.J., Taylor, K.E., Doutriaux, C., 2008. Performance metrics for climate models. J. Geophys. Res. Atmos. 113. https://doi.org/10.1029/2007JD008972

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J.H., von Storch, J.-S., Brüggemann, N., Haak, H., Stössel, A., 2019. Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP). Geosci. Model Dev. 12, 3241–3281. https://doi.org/10.5194/gmd-12-3241-2019

Haarsma, R.J., Roberts, M.J., Vidale, P.L., Senior, C.A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N.S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L.R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M.S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., von Storch, J.-S., 2016. High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. Geosci. Model Dev. 9, 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Kodama, C., Ohno, T., Seiki, T., Yashiro, H., Noda, A.T., Nakano, M., Yamada, Y., Roh, W., Satoh, M., Nitta, T., Goto, D., Miura, H., Nasuno, T., Miyakawa, T., Chen, Y.-W., Sugi, M., 2020. The non-hydrostatic global atmospheric model for CMIP6 HighResMIP simulations (NICAM16-S): Experimental design, model description, and sensitivity experiments. Geosci. Model Dev. Discuss. 2020, 1–50. https://doi.org/10.5194/gmd-2019-369

Lee, J., Sperber, K.R., Gleckler, P.J., Bonfils, C.J.W., Taylor, K.E., 2019. Quantifying the agreement between observed and simulated extratropical modes of interannual variability. Clim. Dyn. 52, 4057–4089. https://doi.org/10.1007/s00382-018-4355-4

Lewis, E., Fowler, H., Alexander, L., Dunn, R., McClean, F., Barbero, R., Guerreiro, S., Li, X.-F., Blenkinsop, S., 2019. GSDR: A Global Sub-Daily Rainfall Dataset. J. Clim. 32, 4715–4729. https://doi.org/10.1175/JCLI-D-18-0143.1

Min, S.-K., Zhang, X., Zwiers, F.W., Hegerl, G.C., 2011. Human contribution to more-intense precipitation extremes. Nature 470, 378.

Reed, K.A., Bacmeister, J.T., Rosenbloom, N.A., Wehner, M.F., Bates, S.C., Lauritzen, P.H., Truesdale, J.E., Hannay, C., 2015. Impact of the dynamical core on the direct simulation of tropical cyclones in a high-resolution global model. Geophys. Res. Lett. 42. https://doi.org/10.1002/2015GL063974

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

Rhoades, A.M., Jones, A.D., O'Brien, T.A., O'Brien, J.P., Ullrich, P.A., Zarzycki, C.M., 2020. Influences of North Pacific Ocean Domain Extent on the Western U.S. Winter Hydroclimatology in Variable-Resolution CESM. J. Geophys. Res. Atmos. 125, e2019JD031977. https://doi.org/10.1029/2019JD031977

Risser, M.D., Paciorek, C.J., Wehner, M.F., O'Brien, T.A., Collins, W.D., 2018. A probabilistic gridded product for daily precipitation extremes over the United States. Clim. Dyn.

Roberts, C.D., Senan, R., Molteni, F., Boussetta, S., Mayer, M., Keeley, S.P.E., 2018. Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP. Geosci. Model Dev. 11, 3681–3712. https://doi.org/10.5194/gmd-11-3681-2018

Roberts, M.J., Baker, A., Blockley, E.W., Calvert, D., Coward, A., Hewitt, H.T., Jackson, L.C., Kuhlbrodt, T., Mathiot, P., Roberts, C.D., Schiemann, R., Seddon, J., Vannière, B., Vidale, P.L., 2019. Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments. Geosci. Model Dev. Discuss. 2019, 1–47. https://doi.org/10.5194/gmd-2019-148

Roberts, M.J., Vidale, P.L., Senior, C., Hewitt, H.T., Bates, C., Berthou, S., Chang, P., Christensen, H.M., Danilov, S., Demory, M.-E., Griffies, S.M., Haarsma, R., Jung, T., Martin, G., Minobe, S., Ringler, T., Satoh, M., Schiemann, R., Scoccimarro, E., Stephens, G., Wehner, M.F., 2018. The Benefits of Global High Resolution for Climate Simulation: Process Understanding and the Enabling of Stakeholder Decisions at the Regional Scale. Bull. Am. Meteorol. Soc. 99, 2341–2359. https://doi.org/10.1175/BAMS-D-15-00320.1

Shaevitz, D.A., Camargo, S.J., Sobel, A.H., Jonas, J.A., Kim, D., Kumar, A., Larow, T.E., Lim, Y.-K., Murakami, H., Reed, K.A., Roberts, M.J., Scoccimarro, E., Vidale, P.L., Wang, H., Wehner, M.F., Zhao, M., Henderson, N., 2014. Characteristics of tropical cyclones in high-resolution models in the present climate. J. Adv. Model. Earth Syst. 6. https://doi.org/10.1002/2014MS000372

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

Srivastava, A., Grotjahn, R., Ullrich, P.A., 2020. Evaluation of historical CMIP6 model simulations of extreme precipitation over contiguous US regions. Weather Clim. Extrem. 29, 100268. https://doi.org/https://doi.org/10.1016/j.wace.2020.100268

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An Overview of CMIP5 and the Experiment Design. Bull. Am. Meteorol. Soc. 93, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Trenberth, K.E., Zhang, Y., Gehne, M., 2017. Intermittency in Precipitation: Duration, Frequency, Intensity, and Amounts Using Hourly Data. J. Hydrometeorol. 18, 1393–1412. https://doi.org/10.1175/JHM-D-16-0263.1

Ullrich, P.A., Devendran, D., Johansen, H., 2016. Arbitrary-Order Conservative and Consistent Remapping and a Theory of Linear Maps: Part II. Mon. Weather Rev. 144, 1529–1549. https://doi.org/10.1175/MWR-D-15-0301.1

Ullrich, P.A., Taylor, M.A., 2015. Arbitrary-Order Conservative and Consistent Remapping and a Theory of Linear Maps: Part I. Mon. Weather Rev. 143, 2419–

2440. https://doi.org/10.1175/MWR-D-14-00343.1

van Oldenborgh, G.J., van der Wiel, K., Sebastian, A., Singh, R., Arrighi, J., Otto, F., Haustein, K., Li, S., Vecchi, G., Cullen, H., 2017. Attribution of extreme rainfall from Hurricane Harvey, August 2017. Environ. Res. Lett. 12, 124009.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., Chauvin, F., 2013. The CNRM-CM5.1 global climate model: description and basic evaluation. Clim. Dyn. 40, 2091–2121. https://doi.org/10.1007/s00382-011-1259-y

Walsh, K.J.E.E., Camargo, S.J., Vecchi, G.A., Daloz, A.S., Elsner, J., Emanuel, K., Horn, M., Lim, Y.K., Roberts, M., Patricola, C., Scoccimarro, E., Sobel, A.H., Strazzo, S., Villarini, G., Wehner, M., Zhao, M., Kossin, J.P., La Row, T., Oouchi, K., Schubert, S., Wang, H., Bacmeister, J., Chang, P., Chauvin, F., Jablonowski, C., Kumar, A., Murakami, H., Ose, T., Reed, K.A., Saravanan, R., Yamada, Y., Zarzycki, C.M., Luigi Vidale, P., Jonas, J.A., Henderson, N., 2015. Hurricanes and climate: The U.S. Clivar working group on hurricanes. Bull. Am. Meteorol. Soc. 96, 997–1017. https://doi.org/10.1175/BAMS-D-13-00242.1

Wehner, M., Gleckler, P., Lee, J., 2020. Characterization of long period return values of extreme daily temperature and precipitation in the CMIP6 models: Part 1, model evaluation. Weather Clim. Extrem. 100283. https://doi.org/https://doi.org/10.1016/j.wace.2020.100283

Wehner, M., Prabhat, Reed, K.A., Stone, D., Collins, W.D., Bacmeister, J., 2015. Resolution Dependence of Future Tropical Cyclone Projections of CAM5.1 in the U.S. CLIVAR Hurricane Working Group Idealized Configurations. J. Clim. 28, 3905–3925. https://doi.org/10.1175/JCLI-D-14-00311.1

Wehner, M.F., 2013. Very extreme seasonal precipitation in the NARCCAP ensemble: Model performance and projections. Clim. Dyn. 40. https://doi.org/10.1007/s00382-012-1393-1

Wehner, M.F., Reed, K.A., Li, F., Prabhat, Bacmeister, J., Chen, C.-T., Paciorek, C., Gleckler, P.J., Sperber, K.R., Collins, W.D., Gettelman, A., Jablonowski, C., 2014. The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1. J. Adv. Model. Earth Syst. 6. https://doi.org/10.1002/2013MS000276

Zarzycki, C.M., Levy, M.N., Jablonowski, C., Overfelt, J.R., Taylor, M.A., Ullrich, P.A., 2014. Aquaplanet Experiments Using CAM's Variable-Resolution Dynamical Core. J. Clim. 27, 5481–5503. https://doi.org/10.1175/JCLI-D-14-00004.1

Zhang, X., Wan, H., Zwiers, F.W., Hegerl, G.C., Min, S.-K., 2013. Attributing intensification of precipitation extremes to human influence. Geophys. Res. Lett. 40, 5252–5257. https://doi.org/10.1002/grl.51010

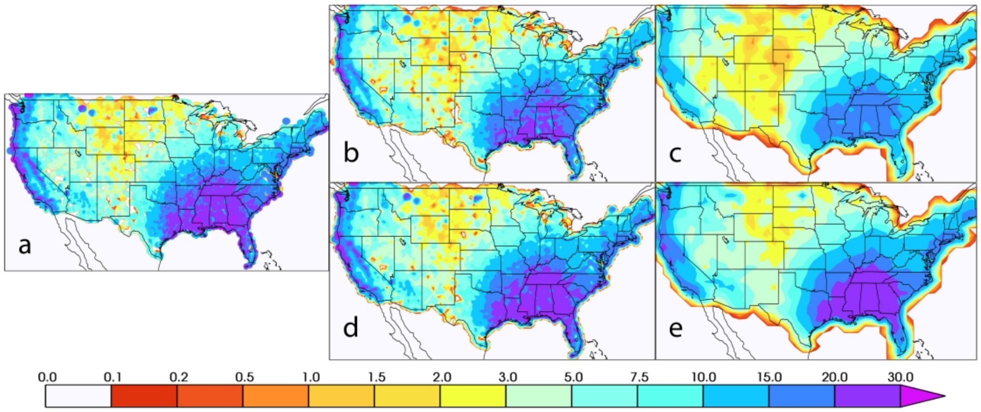

Figure 1. Average DJF maximum 3 hour precipitation accumulation. a) Maximum values calculated on the original 4km polar stereographic mesh and regridded to a 4km latitude-longitude mesh. b) Maximum values obtained by first regridding daily precipitation to a 25km mesh. c) Maximum values obtained by first regridding daily precipitation to a 100km mesh. d) Maximum values obtained by regridding 4km maxima to a 25km mesh. e) Maximum values obtained by regridding 4km maxima to a 100km mesh.

254x111mm (300 x 300 DPI)

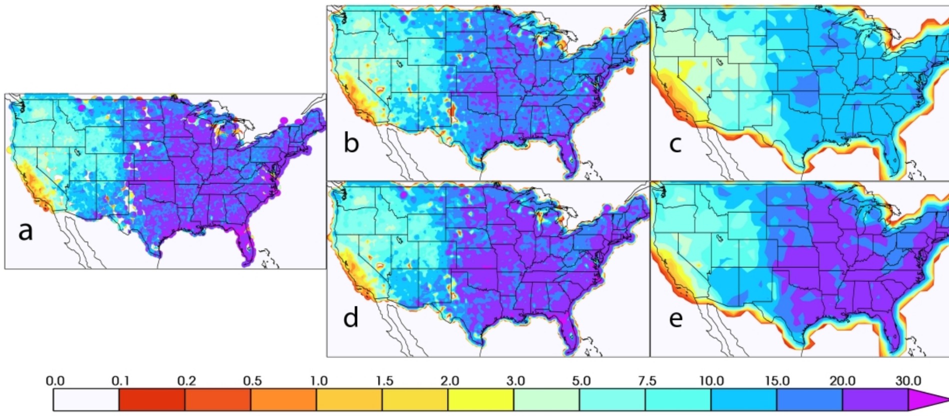

Figure 2. Average JJA maximum 3 hour precipitation accumulation. a) Maximum values calculated on the original 4km polar stereographic mesh and regridded to a 4km latitude-longitude mesh. b) Maximum values obtained by first regridding daily precipitation to a 25km mesh. c) Maximum values obtained by first regridding daily precipitation to a 100km mesh. d) Maximum values obtained by regridding 4km maxima to a 25km mesh. e) Maximum values obtained by regridding 4km maxima to a 100km mesh.

254x119mm (300 x 300 DPI)

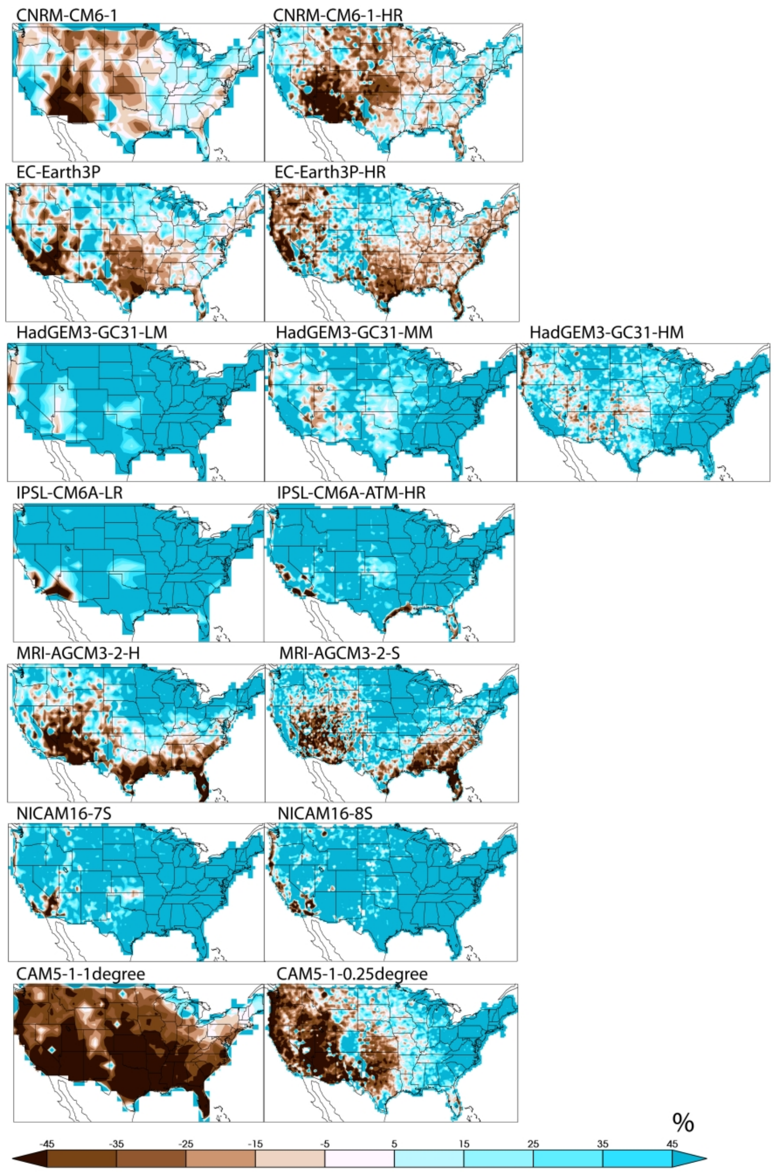

Figure 4. Percent native grid error in simulated average JJA maximum 3 hour precipitation accumulation. Models are arranged low to high horizontal resolution from left to right.

291x446mm (300 x 300 DPI)

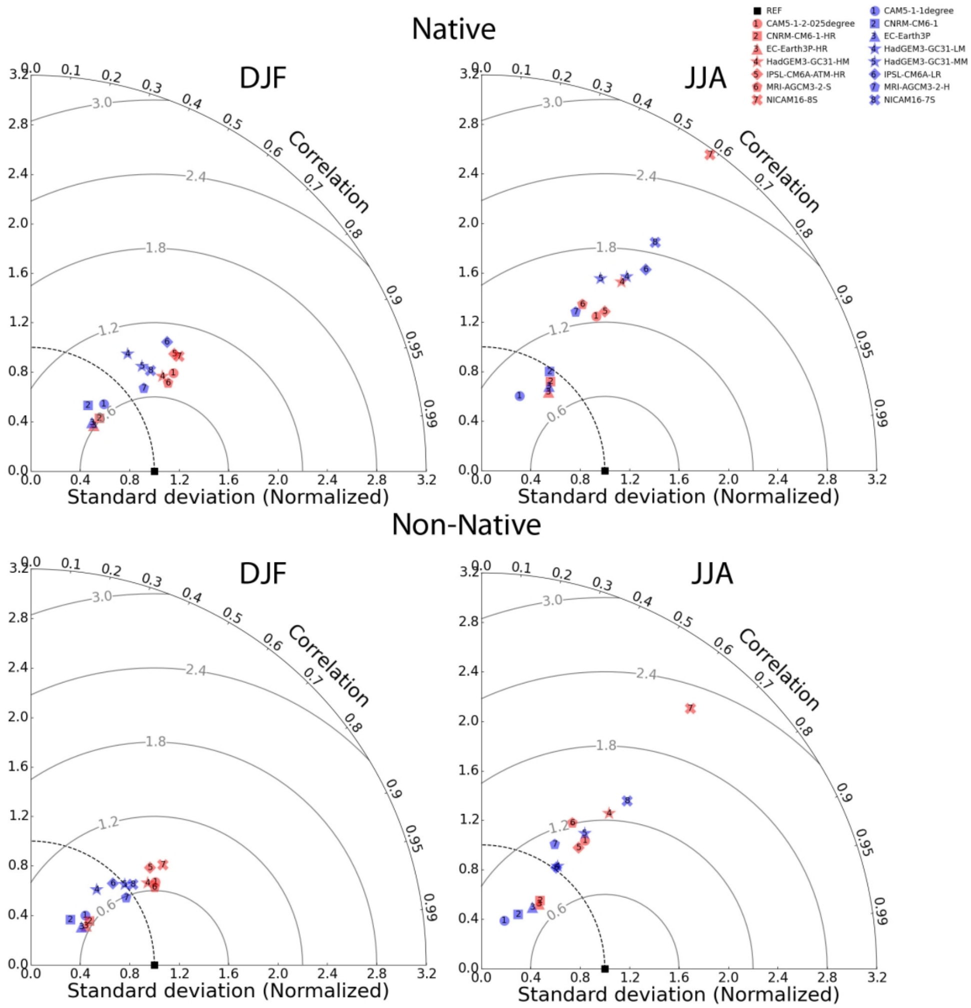

Figure 5. Taylor diagram of average DJF (right) and JJA (left) maximum simulated 3 hour precipitation accumulation from native (top) and non-native (bottom) grid errors. Blue symbols are low resolution models. Red symbols are high resolution models. Symbol shapes are the same for models from the same modeling group. The concentric semi-circles are isolines of normalized root mean square error. The dashed circle represents a normalized standard deviation of unity.

533x558mm (100 x 100 DPI)

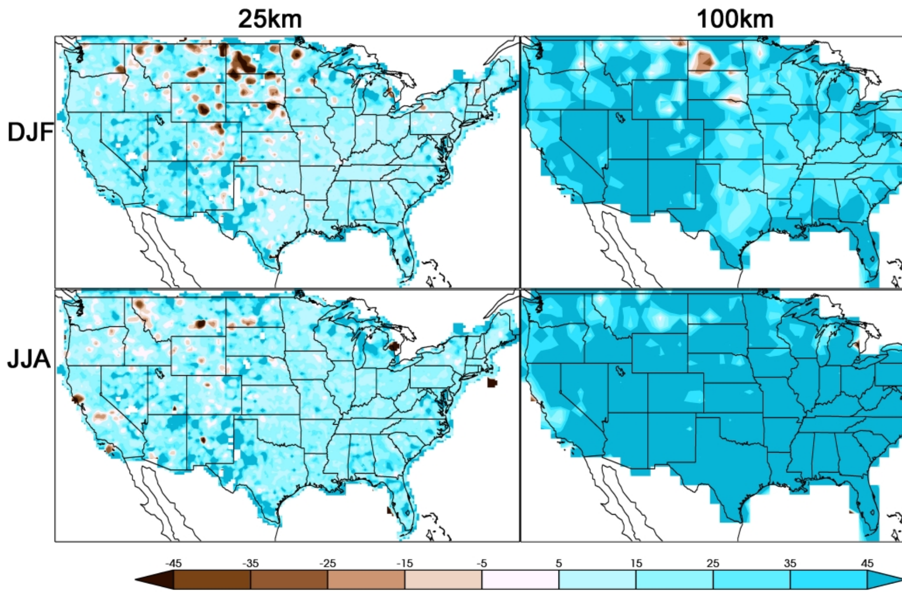

Figure S1. Percent difference between native and non-native constructions of NCEP-EMC average seasonal maximum 3 hour precipitation accumulation. Top row: DJF. Bottom row: JJA. Left column: ~25km grid. Right Column: ~100km grid.

335x212mm (300 x 300 DPI)

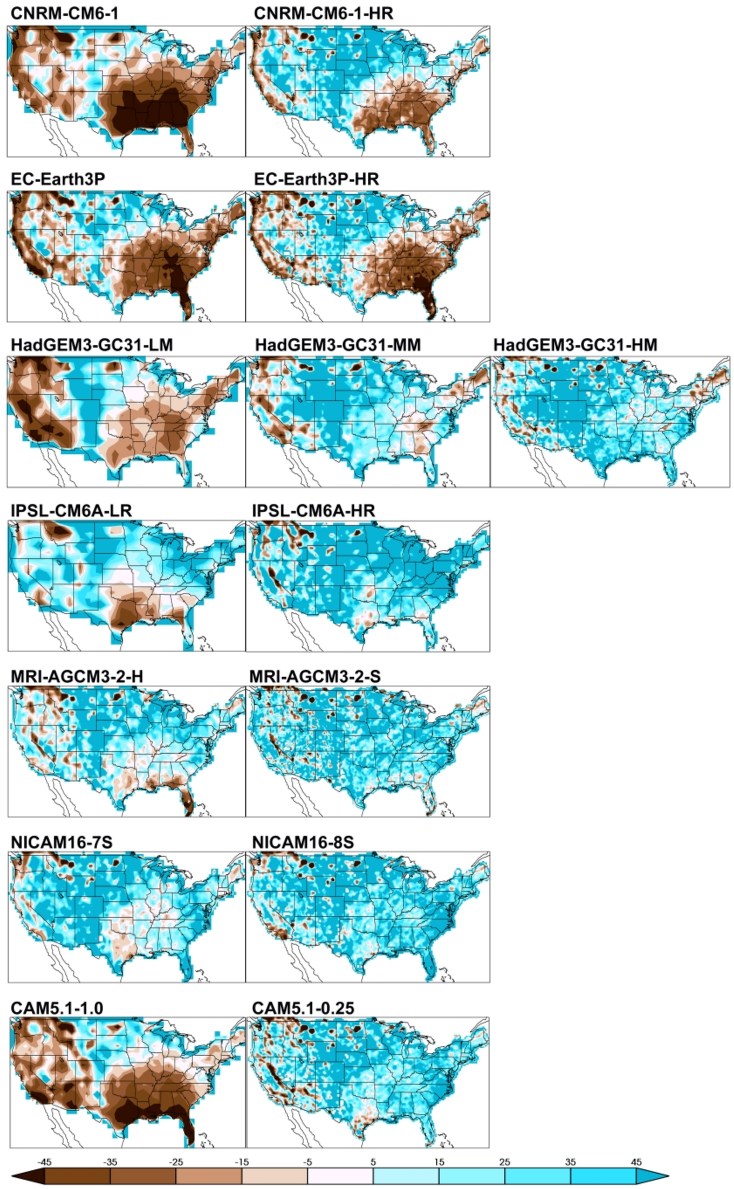

Figure S2. Percent non-native grid error in simulated average DJF maximum 3 hour precipitation accumulation. Models are arranged low to high horizontal resolution from left to right.

293x470mm (300 x 300 DPI)