A multiregion model evaluation and attribution study of historical changes in the area affected by temperature and precipitation extremes

Article

Published Version

A Multiregion Model Evaluation and Attribution Study of Historical Changes in the Area Affected by Temperature and Precipitation Extremes

ANDREA J. DITTUS AND DAVID J. KAROLY

School of Earth Sciences, and Australian Research Council Centre of Excellence for Climate System Science, University of Melbourne, Parkville, Victoria, Australia

SOPHIE C. LEWIS

Research School of Earth Sciences, and Australian Research Council Centre of Excellence for Climate System Science, Australian National University, Canberra, Australian Capital Territory, Australia

LISA V. ALEXANDER AND MARKUS G. DONAT

Climate Change Research Centre, and Australian Research Council Centre of Excellence for Climate System Science, University of New South Wales, Sydney, New South Wales, Australia

(Manuscript received 23 February 2016, in final form 15 August 2016)

ABSTRACT

The skill of eight climate models in simulating the variability and trends in the observed areal extent of daily temperature and precipitation extremes is evaluated across five large-scale regions, using the climate extremes index (CEI) framework. Focusing on Europe, North America, Asia, Australia, and the Northern Hemisphere, results show that overall the models are generally able to simulate the decadal variability and trends of the observed temperature and precipitation components over the period 1951–2005. Climate models are able to reproduce observed increasing trends in the area experiencing warm maximum and minimum temperature extremes, as well as, to a lesser extent, increasing trends in the areas experiencing an extreme contribution of heavy precipitation to total annual precipitation for the Northern Hemisphere regions. Using simulations performed under different radiative forcing scenarios, the causes of simulated and observed trends are investigated. A clear anthropogenic signal is found in the trends in the maximum and minimum temperature components for all regions. In North America, a strong anthropogenically forced trend in the maximum temperature component is simulated despite no significant trend in the gridded observations, although a trend is detected in a reanalysis product. A distinct anthropogenic influence is also found for trends in the area affected by a much-above-average contribution of heavy precipitation to annual precipitation totals for Europe in a majority of models and to varying degrees in other Northern Hemisphere regions. However, observed trends in the area experiencing extreme total annual precipitation and extreme number of wet and dry days are not reproduced by climate models under any forcing scenario.

1. Introduction

Many studies have documented trends in observed temperature and precipitation extremes that could be associated with severe socioeconomic impacts, both globally (Alexander et al. 2006; Donat et al. 2013b) and regionally [e.g., You et al. (2011) and Zhou and Ren (2011) for Asia; Moberg et al. (2006) for Europe; and Alexander and Arblaster (2009) and Gallant et al. (2007) for Australia]. Warm maximum and minimum temperature extremes have increased in most regions around the globe (exceptions occur, e.g., for maximum temperature extremes over eastern North America), while the frequency of cold maximum and minimum temperature extremes has decreased (Seneviratne et al. 2012; Donat et al. 2013b). While changes in precipitation extremes are more spatially heterogeneous, increases in the frequency of heavy precipitation events (e.g., exceeding the 95th percentile) occur in more regions than

Corresponding author address: Andrea J. Dittus, School of Earth, Atmosphere and Environment, Monash University, 9 Rainforest Walk, Clayton, 3800 VIC, Australia.
E-mail: andrea.dittus@monash.edu

DOI: 10.1175/JCLI-D-16-0164.1

decreases (Seneviratne et al. 2012). Much attention has therefore recently been given to understanding the causes of changes in temperature and precipitation extremes, and to identify any anthropogenic contribution to these changes. A number of studies have found a detectable anthropogenic signal for minimum and in some cases maximum temperature extremes globally, and in some regions (e.g., Christidis et al. 2005, 2011; Morak et al. 2011, 2013; Min et al. 2013). Human influence has also contributed to the intensification of annual precipitation extremes over Northern Hemisphere land areas (Min et al. 2011; Zhang et al. 2013).

In this study, a multimodel evaluation and attribution study of variations in the areal extent experiencing temperature and precipitation extremes for five regions across the globe is conducted. This is undertaken using an updated approach based on the climate extremes index (CEI) framework (Karl et al. 1996; Gleason et al. 2008). It consists of five components measuring the percentage area experiencing different types of ''much above or below average'' temperature and precipitation conditions. It was first introduced by Karl et al. (1996) for the United States and there have since been a number of updates (Gleason et al. 2008) and modifications applied to different regions (Gallant and Karoly 2010; Gallant et al. 2014). In a previous study (Dittus et al. 2015), we introduced a new method to calculate the areal extent of extremes, using extreme indices recommended by the Expert Team on Climate Change Detection and Indices, often referred to as ETCCDI indices (Zhang et al. 2011). Using this ETCCDI-based modified CEI (EmCEI), historical changes in the area experiencing temperature and precipitation extremes were analyzed for four continental regions and one hemispheric region: Europe (EURO), North America (NA), Asia, Australia (AUS), and the Northern Hemisphere (NH), where good observational data coverage is available. Increases in the area affected by warm temperature extremes were found for all regions, while increases in the area experiencing precipitation extremes were found for North America and Europe. In Asia, changes in the area affected by precipitation extremes were small but statistically significant.

The EmCEI and its components can be calculated in model simulations, providing an opportunity to better understand the causes for the observed trends. In the present study, we determine whether trends obtained under different radiative forcings allow the relative contributions of human and natural effects on the changes in the areal extent of extremes to be identified. The EmCEI and its components are analyzed in a suite of state-of-the art coupled climate models from phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) under different combinations of imposed forcings. The different sets of simulations include all forcings (historical), natural forcings only (historicalNat), and greenhouse gases only (historicalGHG). We first evaluate whether models are able to simulate the variations of the EmCEI and its components realistically, by investigating the variability in observational and model-simulated components over the period 1951–2005. Next, we compare historical model simulated trends under different forcing scenarios, including anthropogenic and natural forcing combinations, with the observed trends. Observed trends over the period 1951–2010 are also included to illustrate the importance of natural variability in the outcome of the attribution analysis. The results presented in this study thus provide an assessment of the relative contributions of human and natural forcings on the changes in the spatial extent of extremes, which cannot be achieved by investigating the EmCEI in observations alone.

2. Data and methods

a. *Definition of the EmCEI components*

Using definitions from Dittus et al. (2015), five temperature and precipitation components are analyzed in this study. Each component measures the fraction of area experiencing extreme conditions, calculated from seven indices recommended by the ETCCDI (Zhang et al. 2011). These indices, often referred to as ''extremes indices,'' measure the frequency, intensity, or duration of so-called moderate daily extremes, even though not all indices strictly represent extreme conditions. In this study, a subset of these indices is used as input data.

- Warm and cool days (TX90p and TX10p): Annual percentage of days where maximum temperature is above the 90th percentile or below the 10th percentile respectively.
- Warm and cool nights (TN90p and TN10p): Same definitions as their maximum temperature equivalent, only using minimum temperature.
- Total annual precipitation (PRCPTOT) on days where precipitation is equal to or exceeds 1 mm.
- The precipitation amount from heavy rain days (R95p): Total amount of rainfall per year that fell on days when the 95th percentile was exceeded.
- The simple daily intensity index (SDII): Average daily precipitation on a wet day (mm day^{-1}), where a wet day is classified as $\geq$1 mm.

These indices are used as input to calculate the five temperature and rainfall components of the EmCEI,

which measure extreme conditions in the annual indices of daily moderate extremes. Each component consists of two parts, corresponding to the upper and lower tail extremes.

1) Maximum temperature: Percentage area where the frequency of warm days (TX90p) is above the long-term 90th percentile minus the percentage area where the frequency of cool days (TX10p) is above the long-term 90th percentile.
2) Minimum temperature: Same definition as for maximum temperature above but for minimum temperature using warm (TN90p) and cool nights (TN10p).
3) Total precipitation: Percentage area where the annual precipitation anomaly (PRCPTOT) divided by its standard deviation exceeds the 90th percentile minus the percentage area where it is less than the 10th percentile.
4) Heavy rainfall: Percentage area where the proportion of annual rainfall due to heavy rain days (R95p/PRCPTOT) exceeds the 90th percentile, minus the percentage area where it is less than the 10th percentile.
5) Wet and dry days: Percentage area where the number of wet days (WD = PRCPTOT/SDII) exceeds the 90th percentile minus the percentage of area where the number of dry days (DD = 365 − WD)[1] exceeds the 90th percentile.

Note that the percentile thresholds of the ETCCDI indices are always calculated from the entire period available (i.e., 1951–2005 and 1951–2010 respectively). The maximum temperature component thus represents the difference between the percentage area experiencing an extreme number of warm days and the percentage area experiencing an extreme number of cold days. Likewise, the minimum temperature component represents the difference between the percentage area experiencing an extreme number of warm and cold nights. Since the area experiencing cold extremes is subtracted from the area experiencing warm extremes, a positive value represents the percentage area by which warm extremes exceed the percentage area affected by cold extremes, and vice versa. For the precipitation components, positive values correspond to larger areas affected by wet extremes and negative values to larger areas affected by dry extremes (or, in the case of the heavy precipitation component, a negative value indicates the absence of a wet extremes rather than a dry extreme). A limitation of the definitions above is that these components cannot measure changes in the variance of the underlying distribution. Hence, if increases in wet extremes were occurring concurrently with increases in dry extremes, this would not be indicated in the components as defined above. In Dittus et al. (2015) we showed changes in the upper and lower tail separately, and found no evidence of opposing trends occurring in any region, although at the interannual scale wet and dry extremes were found to sometimes occur concurrently over substantial fractions of area. Further details can be found in Dittus et al. (2015). To emphasize longer-term trends and forced responses, 5-yr running means and longer-term trends of the components listed above were used for analysis throughout this study.

[1] Or 366 days in a leap year for observations. In models this number may also be 360, 365, or 365/366 depending on the calendar used.

b. Data and processing

In this study, data from eight global coupled climate models participating in CMIP5 (Taylor et al. 2012) were used. The Global Historical Climatology Network–Daily (GHCND)-based gridded temperature and precipitation climate extremes indices (GHCNDEX; Donat et al. 2013a) complemented by an updated gridded land-based dataset of indices of temperature and precipitation extremes (HadEX2; Donat et al. 2013b) were used as observations and are described at the end of this section. To assess the role of natural variability and observational uncertainty, we also include results for the 1951–2010 period as in Dittus et al. (2015) and the ECMWF twentieth-century reanalysis (ERA-20C; Poli et al. 2016), respectively. Note that the reanalysis does not represent observed extremes. However, in the absence of other observational global datasets of extremes, and to get some idea of uncertainty across observation-based datasets, we here include extremes indices calculated from ERA-20C. The ERA-20C reanalysis is constrained by observations by assimilating surface pressure and wind observations, and using sea surface temperature and sea ice concentration as boundary conditions. Thereby ERA-20C should reproduce the synoptic-scale weather systems of the real world, but daily maximum and minimum temperatures and precipitation amounts are model-simulated variables. The individual indices used to calculate the EmCEI show reasonable agreement with gridded observations after 1950 (Donat et al. 2016a).

For the CMIP5 climate model simulations, the ETCCDI indices required to calculate the EmCEI components were obtained from the Canadian Centre for Climate Modelling and Analysis (Sillmann et al. 2013a,b; http://www.cccma.ec.gc.ca/data/climdex). All models for which indices were available for three or more ensemble members for the historical and historicalNat experiments were selected (Table 1). Since the focus of this paper is on assessing the potential role of

TABLE 1. List of model names, modeling groups, and corresponding number of ensemble members per scenario (historical, historicalNat, and historicalGHG, respectively) used in this study. (Expansions of acronyms are available online at http://www.ametsoc.org/PubsAcronymList.)

Model name	Modeling center	hist	histNat	histGHG
CCSM4	National Center for Atmospheric Research	3	4	3
CanESM2	Canadian Centre for Climate Modelling and Analysis	5	5	5
CSIRO Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation in collaboration with Queensland Climate Change Centre of Excellence	9	5	5
IPSL-CM5A-MR	L'Institut Pierre-Simon Laplace	3	3	0
CNRM-CM5	Centre National de Recherches Météorologiques–Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	9	6	6
GFDL CM3	NOAA/Geophysical Fluid Dynamics Laboratory	5	3	3
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	3	3	3
HadGEM2-ES	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	4	4	4

anthropogenic forcing through comparing historical and historicalNat simulations, the models were chosen based on availability of these scenarios. In addition, the historicalGHG simulations were also used where available (only seven models). Historical simulations include greenhouse gas forcings and anthropogenic aerosol forcings, as well as natural forcings including variations of solar irradiance and volcanic aerosols. The historicalNat experiments only include natural forcings (i.e., solar and volcanic). The historicalGHG experiments on the other hand include variations of greenhouse gas forcing only, excluding both natural and anthropogenic forcings other than greenhouse gases (e.g., aerosol forcing). These experiments span the period from 1850 to 2005; however, here we use data for the period from 1951 to 2005 to match the available observational period (1951–2010) as closely as possible. It should be noted that for two HadGEM2-ES simulations, the time period spanning 1951–2004 was used, as the annual extreme indices for 2005 were not available.

As in Dittus et al. (2015), the primary observations used in this study consist of GHCNDEX (Donat et al. 2013a) data complemented by HadEX2 (Donat et al. 2013b). GHCNDEX data were regridded to HadEX2 resolution (2.5° × 3.75°), using the same first-order conservative remapping procedure as for the climate models. Grid boxes that contain nonmissing data for at least 80% of the time period were used, as well as 80% in the first and last 10 years. HadEX2 data supplemented GHCNDEX where no data were available over the entire period or did not fulfil the completeness criteria. The methods described here are identical to those used in Dittus et al. (2015). We have included results for the two periods 1951–2005 and 1951–2010, to match the modeled time period available and the time period used in our previous study respectively. A recent study (Kim et al. 2016) found that attribution statements were sensitive to the time period used, which also allows us to assess the effect of natural variability. Figure 1 shows the observational coverage available over 1951–2005. The regions used in this study are the same as in Dittus et al. (2015) and are also shown in Fig. 1. All indices were masked to R95p coverage, as this index has the lowest observational coverage available. Note that the results for 1951–2010 were masked with the R95p observational coverage over that period; however, differences in coverage between the two periods are minimal.

The model and reanalysis data were regridded to HadEX2 resolution (2.5° × 3.75°) using a first-order conservative remapping procedure (Jones 1999) and masked to the observational coverage of R95p (Fig. 1). As in Dittus et al. (2015), prewhitened time series were used to account for autocorrelation for all trend calculations. Trends are calculated using the Theil–Sen slope estimator (Sen 1968). The multimodel or model mean trend corresponds to the average trend across all ensemble members, of all models or a single model respectively. Trends were calculated for each ensemble member prior to averaging. The variability in the EmCEI components is estimated from detrended 5-yr running mean time series. The multimodel mean variability is estimated as the square root of the average variance from all ensemble members. Note that the thresholds to determine extreme (i.e., top and bottom 10%) values are calculated relative to the climatology at each grid box for each model run individually. This effectively removes any model biases, as threshold exceedances are determined relative to the models' own threshold. This likely also reduces the effect of differences in spatial scales (point data for the gridded

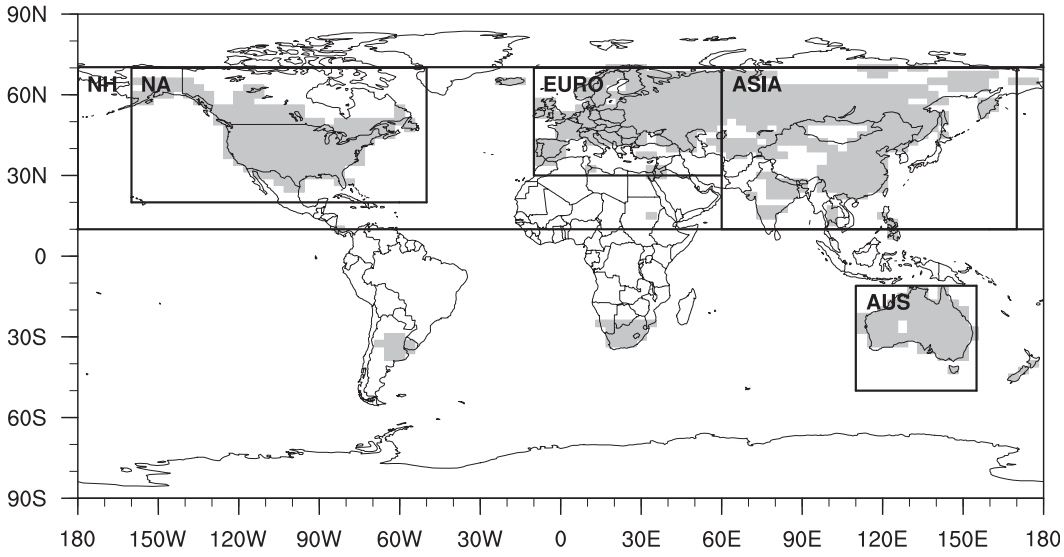

FIG. 1. Merged GHCNDEX and HadEX2 coverage, adapted from Dittus et al. (2015) for the 1951–2005 base period used here. The boxes define the regions used in this study. As in Dittus et al. (2015), GHCNDEX data were regridded to HadEX2 resolution. HadEX2 data were used to complement GHCNDEX data where no GHNCDEX data were available over the entire period.

observations and area averages for reanalysis and models) between the different datasets, which is known to be a major source of observational uncertainty (Herold et al. 2016) and an obstacle to adequately compare observed and modeled precipitation extremes (e.g., Chen and Knutson 2008). Therefore, this allows examination of forced responses irrespective of existing model biases. However, it is important to bear in mind that model agreement would likely be lower if absolute thresholds were used. The same logic also applies to the different observational datasets.

3. Evaluating simulated EmCEI and components

Time series of the 5-yr smoothed EmCEI components as simulated by CMIP5 models (using Europe as an example) are shown in Fig. 2. The orange plume in Fig. 2 corresponds to the historical (all forcings) simulations. The blue plume corresponds to the historicalNat scenario and will be discussed in section 4. For both temperature components, the historical simulations show a clear increasing trend in the area affected by a much above average number of warm days and nights, corresponding also to a decrease in the area affected by a much above average number of cold days and nights. The range of historical simulations includes the observations and reanalysis and is therefore consistent with observed variations in these components. For the minimum temperature component, the historical simulations appear to slightly underestimate the observed increase in area affected by warm extremes. However, as the plumes represent the 5th–95th percentile range, the observations are likely still within the simulated range. Note that because the area affected by extremes is calculated using individual percentile thresholds for each realization, differences between models as well as differences between models and observations are likely to be reduced, as any systematic model biases are removed by comparing the changes relative to individual percentile thresholds, not common thresholds. For the heavy precipitation component (Fig. 2, middle right), the model simulations are consistent with observations; however, the reanalysis lies slightly outside the modeled range. For this component, there is little to no correlation between interannual variations of observations and reanalysis. For the total precipitation and wet and dry day components, the models are unable to capture the full range of observations and reanalysis. In this case, while there are larger differences between observations and reanalysis than for the temperature components, both products agree reasonably well on decadal variations in these components. The models' inability to capture the full range of observations and reanalysis could be due to an inability to capture the long-term trends in these components, or as a result of underestimating the observed decadal variability in these components, or both.

To assess the models' ability to simulate the observed EmCEI and its components, a Kolmogorov–Smirnov (KS) test was performed to determine whether there is a statistical difference between the distributions of

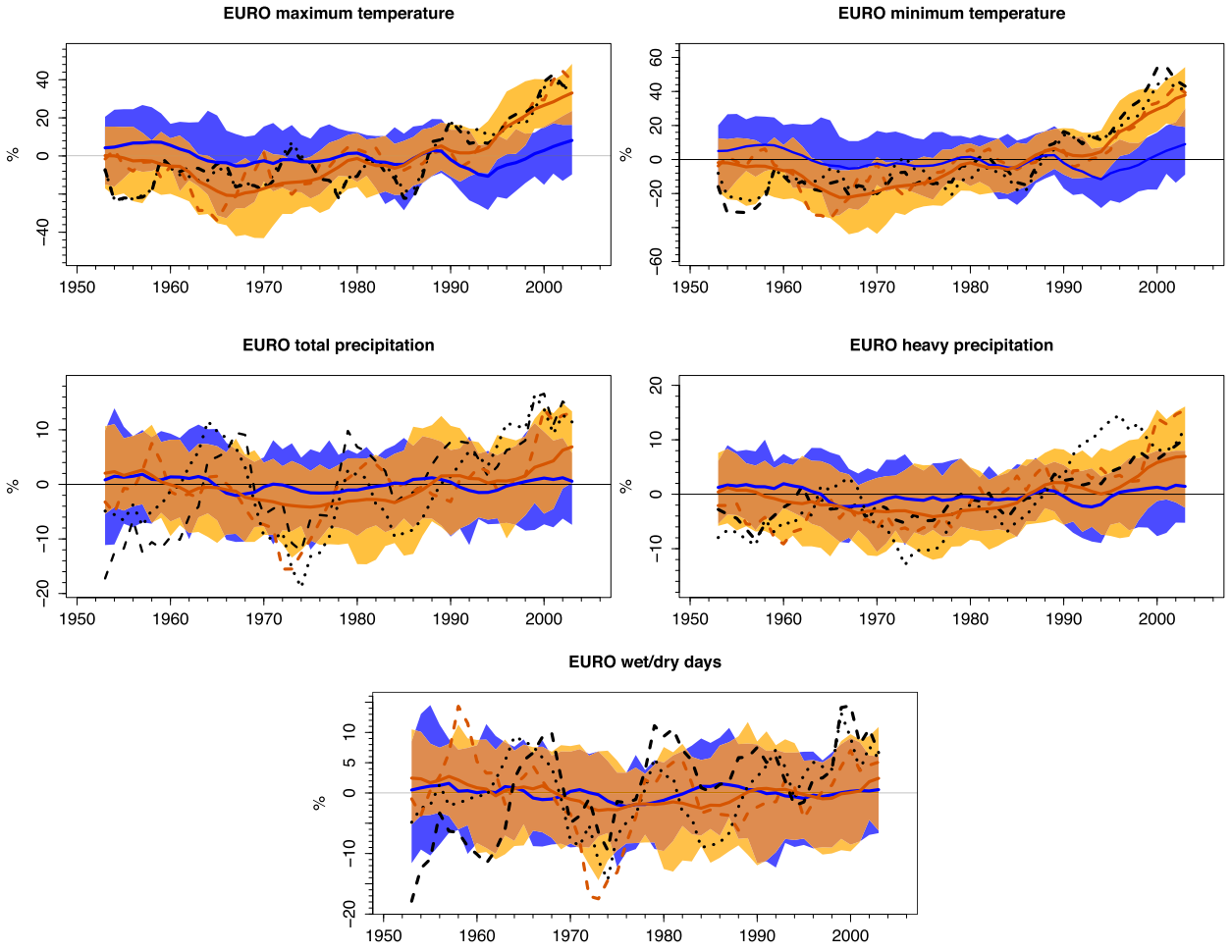

FIG. 2. The simulated 5-yr smoothed EmCEI and its components for Europe. The plumes were obtained from eight multimember CMIP5 models (41 ensemble members for the historical plume and 33 for the historicalNat plume). The range of values in the plume represents the 5th–95th percentile of all values. The orange plume shows the historical simulations, and the blue plume shows the historicalNat results. The solid lines show the multimodel mean for each set of experiments respectively. The black dashed line corresponds to the gridded observations (GHCNDEX + HadEX2), the dotted line corresponds to the reanalysis (ERA-20C), and the orange dashed line corresponds to a single historical model realization, for illustration purposes.

observed and simulated annual values. This test was conducted for all detrended runs across the eight models used in this study (41 in total for the historical simulations) and for each of the detrended components for both the gridded observations and reanalysis. In all regions and components, the observed and modeled distributions of values are indistinguishable in a statistical sense, at the 5% level. Some exceptions occur as shown Table 2, perhaps slightly more so for the precipitation components. Using a statistical test at the 5% level, the expected number of random independent cases that would be expected to fail such a test even if they were drawn from the same distribution would be 1 in 20 on average.

The interannual (not shown) and decadal variability in the models has also been evaluated against the variability in the gridded observations and reanalysis (Fig. 3). Overall, the models have comparable decadal variability to the observations and reanalysis, with some differences between the regions. For the heavy precipitation component (component 4), the models tend to overestimate the observed variability, but the model variability is still consistent with observations except for Europe. The variability in reanalysis lies well within the modeled range. In the case of the total precipitation and wet and dry day components, the simulated variability is generally smaller than for the observations for most runs over North America, and the average variability across all models is substantially lower than observed. It is interesting to note that both simulated and observed variability of the total precipitation and wet and dry day components are much larger across Australia than any other region. This is likely due to the influence of El

TABLE 2. Number of cases where the null hypothesis (observed and simulated EmCEI components are taken from the same distribution) is rejected, out of 41 detrended historical simulations from eight models. The left number corresponds to the test being performed with the gridded observations, and the second number corresponds to test results with ERA-20C. Significance was assessed at the 5% level.

	EURO	NA	ASIA	AUS	NH
Max temp (C1)	0/0	0/0	0/0	0/1	1/1
Min temp (C2)	0/0	0/2	0/0	0/0	0/0
Total precipitation (C3)	0/0	0/0	0/0	0/0	0/0
Heavy precipitation (C4)	3/0	0/0	0/1	0/0	0/0
Wet and dry days (C5)	1/0	2/0	0/0	0/0	0/0

Niño–Southern Oscillation (ENSO) on large-scale variations in rainfall across Australia [see also Dittus et al. (2015), where the observed time series showed large wet and dry variations in phase with ENSO].

4. Modeled trends and causes

With few exceptions discussed in the previous section, the models are able to capture the variability in the individual components. Hence, the trends in each component as simulated under different forcing scenarios are investigated here. Comparing the trends under the different forcing scenarios allows identification of the effect of individual forcings on components of the EmCEI. In particular, the importance of anthropogenic effects is assessed using historical simulations (including all forcings) and historicalNat simulations (including only natural influences). We first discuss the temperature attribution results, followed by a detailed analysis of the precipitation components across the different regions.

a. Maximum and minimum temperature components

Figures 4 and 5 provide an overview of the trends simulated for the different scenarios across regions, for the maximum and minimum temperature components respectively. Each vertical line represents trends from one individual climate model, with models shown in the same order as listed in Table 1. Each dot on these vertical lines represents an individual ensemble member, allowing an estimate of internal variability by providing a range of trends within one climate model. Positive trends in the temperature components are found across all models for all regions (Figs. 4 and 5). The multimodel mean trend for the historical simulations is very similar across regions, and the multimodel simulated trends are consistent with observations in most cases. Exceptions occur for the model-simulated trends in the maximum

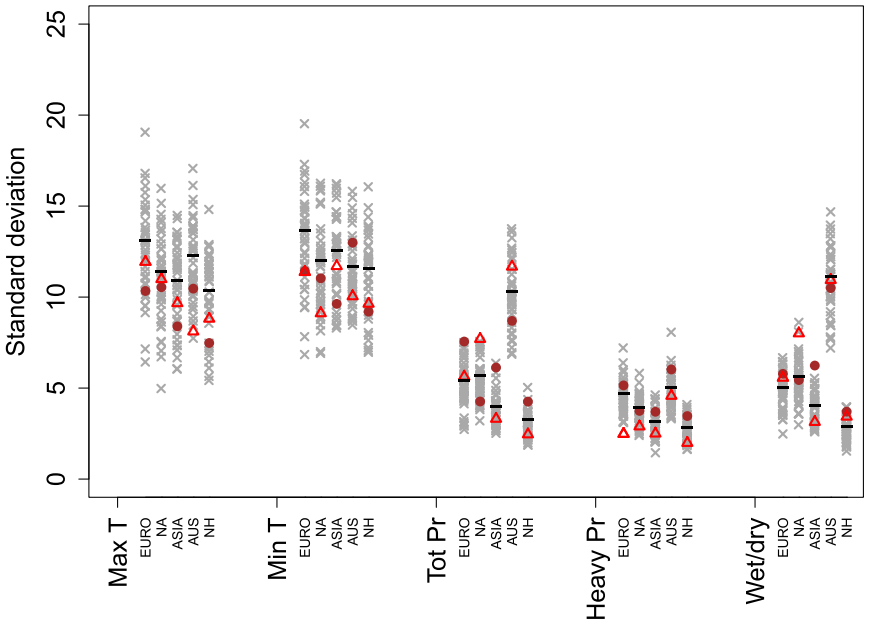

FIG. 3. Decadal variability of the EmCEI components, across all models and regions. Decadal variability is represented by the standard deviation of the low-pass-filtered components (5-yr running mean). The interannual time series were detrended using the Theil–Sen slope estimator prior to applying the running mean and calculating the standard deviation. Each gray symbol represents one ensemble member; the black horizontal line corresponds to the square root of the multimodel mean variance from all ensemble members. The red triangles represent the observations based on the merged GHCNDEX and HadEX2 dataset, and the brown circles represent the variability in ERA-20C.

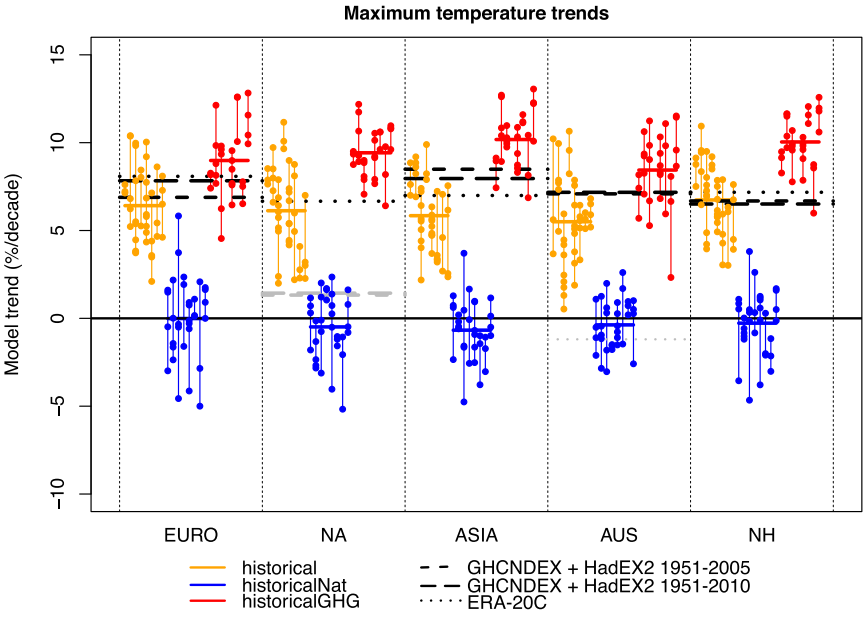

FIG. 4. Trends in the maximum temperature component for all regions and for all scenarios (historical, historicalNat, and historicalGHG). Each model is represented by one vertical line, and dots indicate the individual ensemble members. The models are in the order listed in Table 1. The colored horizontal lines correspond to the multimodel mean. The dashed horizontal lines correspond to the trends in gridded observations for two different time periods, and the dotted line corresponds to the trends in the ERA-20C reanalysis. Trends that are statistically significantly different from zero are shown in black, and gray indicates that the trends are not significantly different from zero. The significance of these trends is assessed using the Mann–Kendall trend test at the 5% level.

temperature component in North America. In this region, the trend is small and not statistically significant in the gridded observations but is statistically significant in ERA-20C. The model-simulated trends are of similar magnitude to the other regions and thus consistent with the reanalysis, but not the gridded observations. The lack of trend in the observations is likely due to the so-called warming hole in the southeastern United

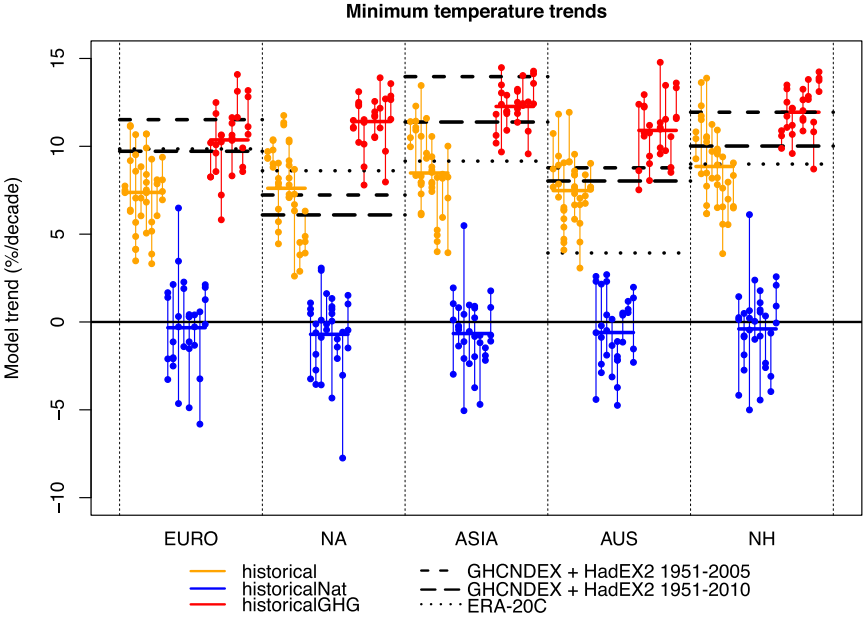

FIG. 5. As in Fig. 4, but for the minimum temperature component.

States (e.g., Portmann et al. 2009). Studies have found that coupled models cannot simulate a warming hole in the second half of the twentieth century robustly (e.g., Kunkel et al. 2006; Meehl et al. 2012). Some studies have suggested that the observed warming hole is a consequence of internally generated decadal variability in the tropical Pacific (e.g., Meehl et al. 2012, 2015) or associated variations in the North Atlantic (e.g., Kunkel et al. 2006), in which case the models' inability to simulate a warming hole is to be expected. Other studies have also suggested altered hydrologic feedback (Pan et al. 2004, 2013) or anthropogenic aerosols (e.g., Yu et al. 2014) as causes for the warming hole. It is unclear why the reanalysis does not appear to produce a warming hole. A recent study did identify a warming hole in ERA-20C (Donat et al. 2016a), but small differences in the timing and extent of the warming hole could have a significant influence on whether or not it is captured by our definition of the spatial extent of extremes.

The simulated historical trends in the minimum temperature component are consistent with observations and reanalysis in all regions, although for Europe and Asia the trends in the 1951–2005 gridded observations are slightly larger and just outside the modeled range. It is interesting to note that while the multimodel ensemble simulated trends are comparable with observations, many individual models are not. The range of historical forced trends is distinct from the range of trends from the historicalNat simulations and in most cases inconsistent with the historicalNat simulations on a model by model basis. The historicalNat simulations are inconsistent with the observations, indicating that the long-term trends in the minimum temperature components are caused by anthropogenic forcing. However, the difference in the magnitude of observed trends between the two time periods and datasets is quite substantial in some regions. It is important to note that our results pertain to the long-term trends in the temperature components, and do not preclude an influence of natural forcings on variability at decadal and longer time scales.

The historicalGHG simulations generally simulate stronger trends than the historical simulations. This difference is likely due to anthropogenic sulfate aerosol forcing offsetting part of the greenhouse gas warming, as this forcing is included in the historical but not the historicalGHG simulations. An anthropogenic contribution to the increasing area affected by maximum and minimum temperature extremes is thus found for all regions, in particular from greenhouse gas forcing. For Europe, this can also be seen in the top panels of Fig. 2, where the increase in the maximum and minimum temperature components in the historical simulations is not reproduced by the natural simulations (blue plume). For the maximum temperature component over North America, an anthropogenic contribution is found in the model simulations, but this is not consistent with the gridded observations.

b. *Precipitation components*

The observed increasing trends in the area affected by extreme total precipitation amounts are statistically significantly different from zero at the 5% level for Europe, North America, Asia, and the Northern Hemisphere (Fig. 6). The reanalysis shows substantially different results in some cases, in Asia for instance where the sign of the trend is negative, albeit small and nonsignificant. The simulated trends under the historical and historicalNat scenario are not consistent with the observed trends for any model over Europe, North America, and the Northern Hemisphere, although they are consistent with ERA-20C in all of these regions except North America (Fig. 7). It is thus difficult to assess the models' performance with confidence given the differences between observations and reanalysis. These results suggest that models (and reanalysis) potentially underestimate the magnitude of the forced signal, or underestimate the 50-yr variability in this component. However, the differences between gridded observations and reanalysis also are substantial, and hence add additional uncertainty. The multimodel mean trend of the historical simulations is larger than in the historicalNat simulations. However, as the range of trends between both scenarios overlap and the magnitude of the simulated historical trend is substantially too low, attribution of changes in the area affected by total precipitation extremes is not possible. Over Asia, the historical trends are consistent with the observed trend. Rather than the models performing better over this region, it is likely that the better agreement may be due to coincidence, as the observed trend is smaller over Asia compared to other regions and thus closer to the simulated trends. The historicalNat simulations do not include the observed trend, but there is substantial overlap between the range of historical and historicalNat trends. Over Australia, the observed trend was not statistically significant and trends from all scenarios include the observed trend. Note that the range of simulated trends for Australia is approximately twice that in any other region. This is consistent with the large decadal variability in the observed and simulated rainfall components over Australia (Fig. 3), associated with the importance of ENSO over this continent. The models are thus able to simulate changes in the area affected by total rainfall extremes over Australia well,

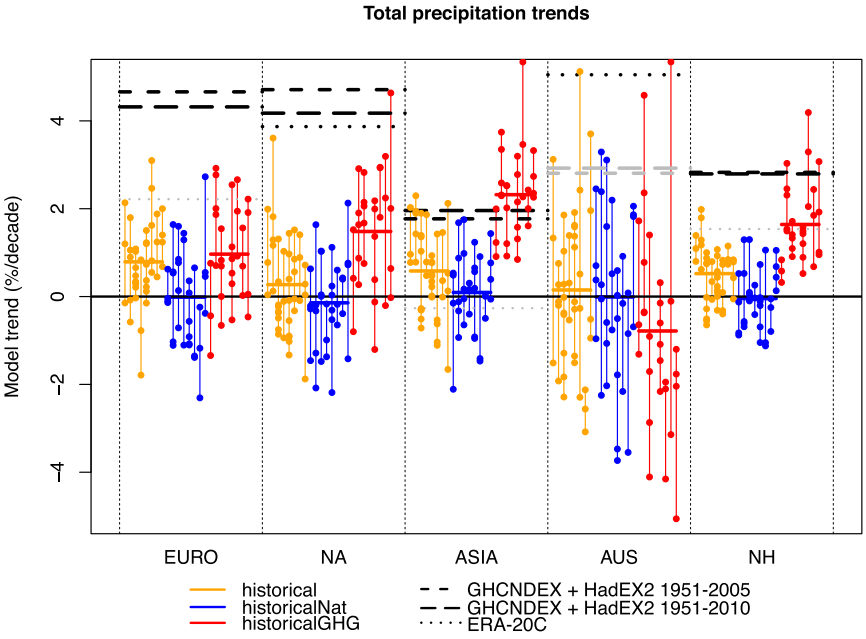

FIG. 6. As in Fig. 4, but for the total precipitation component.

even though natural variability is dominant for this component over Australia.

For all regions except Australia, the multimodel mean trend for the total precipitation component in the historicalGHG simulations is larger than for the historical simulations. For Australia, the multimodel mean trend for the greenhouse gas (GHG) simulations is negative, but there are a broad range of trends both positive and negative for different models. In Asia as well as the Northern Hemisphere, positive trends for all the GHG simulations are found, which are generally consistent with observations. These results indicate that GHG forcing is needed to explain the observed trends. However, as the historical simulations cannot reproduce observed trends, the strength of different forcings in the model is not well captured and an attribution statement

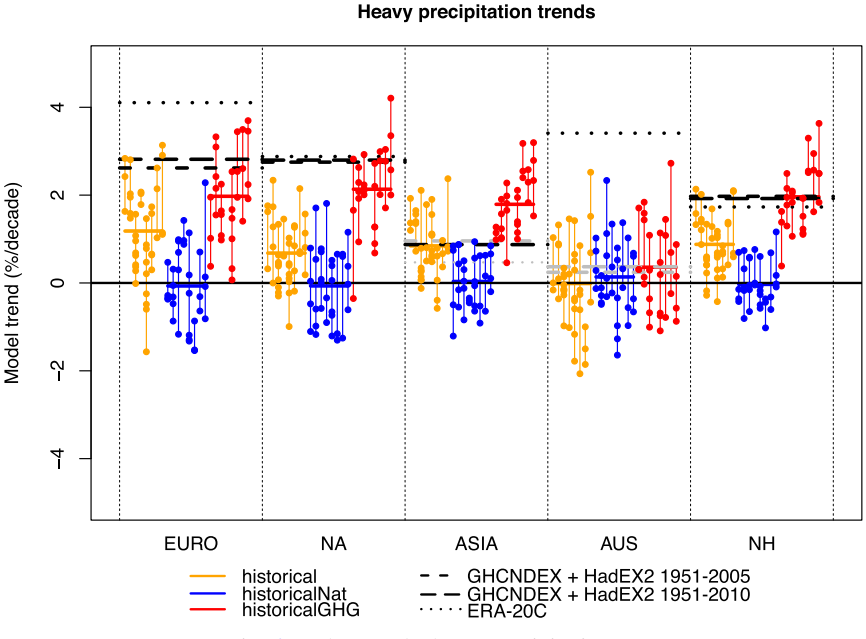

FIG. 7. As in Fig. 4, but for the heavy precipitation component.

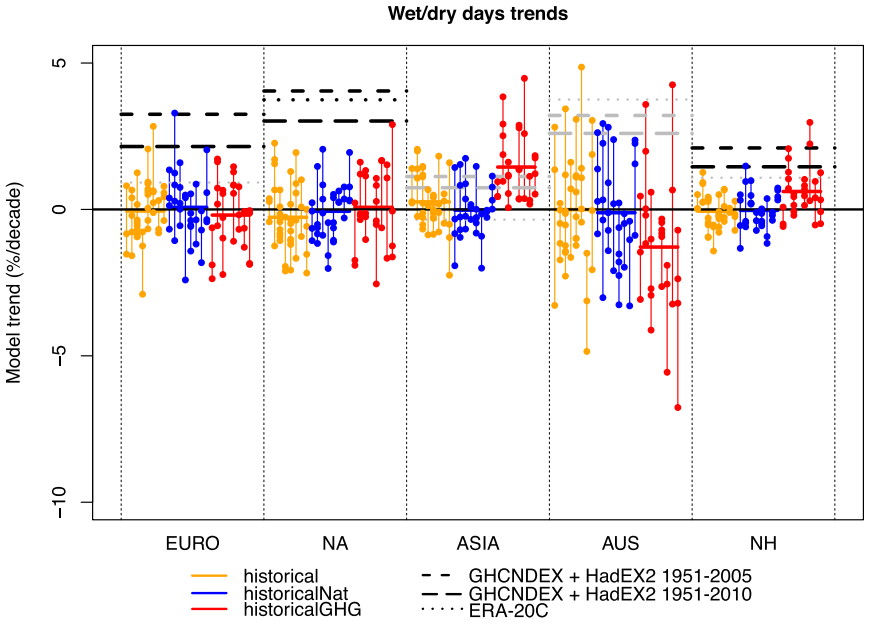

FIG. 8. As in Fig. 4, but for the wet and dry day component.

cannot be made. Overall, the models are not able to reproduce the observed trends in this component in Europe, North America, and the Northern Hemisphere. For Europe, this can also be seen in Fig. 2, where the plumes overlap throughout the entire period. For Asia and Australia, the simulated trends are consistent with observations. In Asia, this is likely due to the small observed trend occurring in this region. In Australia, there is no significant trend in the observations or the models, suggesting that natural variability is dominant over this continent. Very similar results are found for the wet and dry day components (Fig. 8). It is well known that climate models generally produce precipitation too frequently with too small intensities (e.g., Sun et al. 2006; Dai 2006; Stephens et al. 2010). It is thus perhaps not surprising that trends in this component based on the number of wet days are less well captured than for the heavy precipitation component, for example.

The results are more interesting for the heavy precipitation component (Fig. 7), particularly for Europe. For Europe, six out of eight models are simulating positive trends in this component, of which four have trend magnitudes comparable with the observed trend, and for five models the ranges of trends between historical and historicalNat simulations do not overlap. In Fig. 2, this can be seen in the last 15 years approximately, where the plumes overlap but are increasingly diverging. The trends in the historicalGHG simulations are slightly larger than in the historical simulations. These results show that an anthropogenic contribution to the heavy precipitation component over Europe is found for a majority of models analyzed. However, these results also illustrate the importance of considering multiple models for attribution studies of precipitation extremes, as different attribution statements may be reached depending on the choice of model. Moreover, the choice of reference observational dataset is also key when assessing model performance and agreement with observations, as illustrated by the difference between gridded observations and reanalysis in Fig. 7. A majority of simulations also show increasing trends in the area experiencing above average contribution of heavy precipitation to total precipitation for Asia, and to a lesser extent for North America. This is also reflected in the Northern Hemisphere as a whole. Furthermore, as for Europe, more than half of the models exhibit trend magnitudes consistent with observations in Asia. The models underestimate the magnitude of the trends over North America; however, there is a clear difference emerging between the different scenarios, indicating a likely anthropogenic influence on this component that is underestimated in the models. Over Australia, the simulations do not exhibit any trend, consistent with no trend in the observations. In all Northern Hemisphere regions, stronger trends are found for the historicalGHG simulations than the historical (all-forcing) simulations. These results for the heavy precipitation component are more similar to those for the temperature components than for the other precipitation components. This is perhaps unsurprising, since thermodynamic arguments suggest that globally averaged changes in daily precipitation extremes should increase

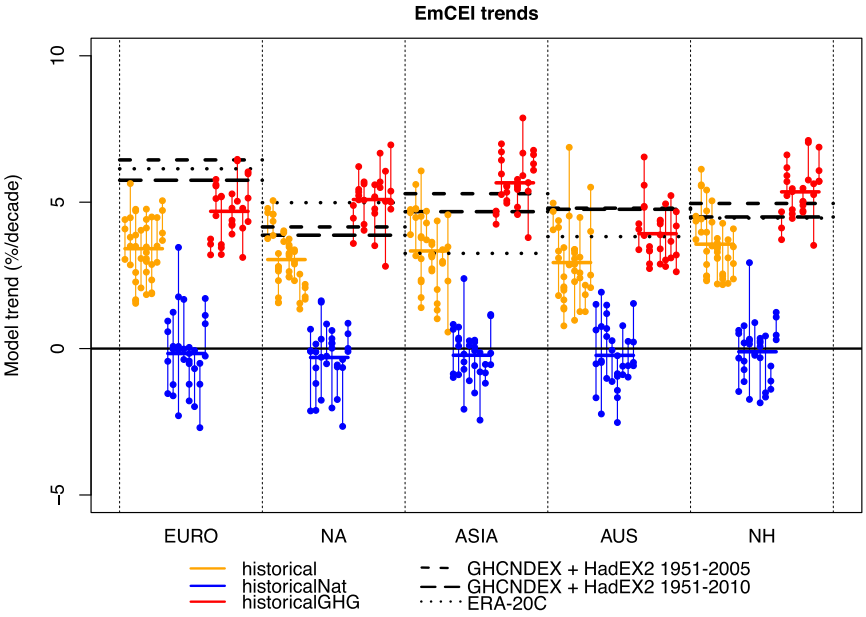

FIG. 9. As in Fig. 4, but for the EmCEI.

at a faster rate than increases in total precipitation, which are energetically constrained (e.g., Westra et al. 2013; Kharin et al. 2013), although there are regional differences (e.g., Donat et al. 2016b). It is therefore expected that the heavy precipitation component, which is derived from daily precipitation extremes, should increase faster than extremes in total annual precipitation or wet and dry days.

As discussed in Dittus et al. (2015), the different components can be combined to form the EmCEI. This is achieved by averaging the five components. However, as noted in Dittus et al. (2015), the interpretation of the EmCEI can be somewhat difficult as a result of combining five different components. Here we briefly discuss the attribution results for the EmCEI.

There are very few differences in the simulated EmCEI trend magnitudes between the different regions. A significant positive trend is found for the historical simulations in all regions, largely stemming from the temperature components (Fig. 9). The simulated historical trends underestimate the observed trends in all regions in the multimodel mean, although individual models are consistent with observations in all regions except Europe. Given this is the average of the five components, the trends are likely underestimated because of the models being unable to accurately reproduce the measured trends in the total precipitation and wet and dry day components. There are no trends in the natural runs or the observations, and substantially larger trends in the GHG simulations than in the historical simulations. We expect that the differences between the historical and historicalGHG simulations are primarily due to the sulfate aerosol forcing included in the historical simulations, although natural and other anthropogenic forcings may also contribute to these differences. In Dittus et al. (2015), we argued that the combination of components may be useful for detection and attribution because of an enhanced trend-to-noise ratio compared to each individual component. This conclusion that the combined EmCEI has a higher signal-to-noise ratio than the individual components is true for most regions in observations. However, because of the models' underestimate of the observed trends in two precipitation components, this conclusion does not hold in general for the models.

5. Concluding remarks

In this study, we use the EmCEI introduced in Dittus et al. (2015) to assess the respective roles of natural and anthropogenic forcings in driving changes in the area affected by temperature and precipitation extremes across four continental and one hemispheric region. This method is computationally efficient, and through investigating fractions of area above percentile thresholds determined separately for observations and each model, the effect of model biases is reduced. Furthermore, studies have shown advantages to adopting a spatial perspective when analyzing extremes; for example, spatial aggregation reduces the effects of gridbox-level noise and leads to more robust projections of extremes (e.g., Fischer et al. 2013). Here, we show

that a clear anthropogenic signal is present in the temperature components in all regions. For the maximum temperature component, the model-simulated historical trends are consistent with observed trends in all regions except North America, and inconsistent with the natural simulations. Over North America as a whole, there is no trend in the gridded observations for this component although a significant trend is found using the ERA-20C reanalysis. The trends in the historical simulations are of similar magnitude to the reanalysis and inconsistent with the gridded observations. For the minimum temperature component, the historical simulations are consistent with observations in North America and Australia, while for Europe and Asia this is dependent on the time period considered. Over the period 1951–2005, whereas there is a positive trend in all historical simulations, the magnitude of the trend is too low compared to the observations. This emphasizes the sensitivity of model evaluation and attribution statements to decadal variability in the observed record. The influence of modes of variability such as the Pacific decadal oscillation and Arctic Oscillation and associated choice of reference period on attribution statements was pointed out by Kim et al. (2016). In these regions (Europe and Asia), the magnitude of the trends in the historicalGHG simulations is closer to the observations than for the historical simulations, suggesting that the models underestimate the anthropogenic forced response. The models are unable to simulate the observed trends in the total precipitation and wet and dry day components, although simulated trends are comparable with observations in regions with no or low trends in observations (e.g., Asia and Australia). There is evidence that climate models are at least partially able to replicate the observed increase in the areas experiencing ''much above average'' contribution of heavy precipitation to total precipitation across the Northern Hemisphere, and that part of this increase in area is due to anthropogenic effects. Our results are consistent with previous studies that find an anthropogenic contribution to temperature extremes globally (e.g., Christidis et al. 2005; Min et al. 2013) and precipitation extremes in Northern Hemisphere land areas (Min et al. 2011; Zhang et al. 2013). Our study has been able to extend the attribution analysis to measures of the spatial extent for different types of temperature and precipitation extremes and to most of the continents with sufficient observational data. The areas for which we find an anthropogenic contribution further correspond to the areas where King et al. (2015) find that the time of anthropogenic emergence for temperature and precipitation extremes has occurred or will occur in the near future. Note that our analysis focused on the causes of long-term trends, and did not investigate the causes of decadal variability, for which natural forcing may be important (e.g., Christidis et al. 2011; Kim et al. 2016).

In conclusion, the models are able to simulate the magnitude of decadal variability in the EmCEI and its components. The trends in the modeled temperature components indicate an increase in area experiencing a much above average number of warm days and nights, and a decrease in the area experiencing a much above average number of cold days. This increase cannot be reproduced by the natural simulations, indicating the role of anthropogenic forcing on these components. The simulated trends in the area experiencing a much above average annual precipitation amount and number of wet and dry days cannot reproduce the observed trends. However, a majority of models are able to reproduce the increasing area experiencing an extreme proportion of total rainfall from heavy rainfall in historical simulations for the Northern Hemisphere regions. This work provides the model evaluation basis for possible further work investigating future projections using the EmCEI framework.

Acknowledgments. Authors are supported by Australian Research Council Grant CE110001028. AJD also acknowledges support from the David Lachlan Hay Memorial Fund. SCL is also supported by Australian Research Council DECRA DE160100092. MGD was supported by Australian Research Council Grant DE150100456. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups listed in Table 1 of this paper for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. ECMWF ERA-20C data used in this study have been obtained from the ECMWF Data Server. We thank Dr. Dáithí Stone and an anonymous reviewer for their useful comments, which helped improve this manuscript.

REFERENCES

Alexander, L. V., and J. M. Arblaster, 2009: Assessing trends in observed and modelled climate extremes over Australia in relation to future projections. *Int. J. Climatol.*, **29**, 417–435, doi:10.1002/joc.1730.

——, and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, **111**, D05109, doi:10.1029/2005JD006290.

Chen, C.-T., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *J. Climate*, **21**, 1605–1621, doi:10.1175/2007JCLI1494.1.

Christidis, N., P. A. Stott, S. Brown, G. C. Hegerl, and J. Caesar, 2005: Detection of changes in temperature extremes during the second half of the 20th century. *Geophys. Res. Lett.*, **32**, L20716, doi:10.1029/2005GL023885.

——, ——, and S. J. Brown, 2011: The role of human activity in the recent warming of extremely warm daytime temperatures. *J. Climate*, **24**, 1922–1930, doi:10.1175/2011JCLI4150.1.

Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate*, **19**, 4605–4630, doi:10.1175/JCLI3884.1.

Dittus, A. J., D. J. Karoly, S. C. Lewis, and L. V. Alexander, 2015: A multiregion assessment of observed changes in the areal extent of temperature and precipitation extremes. *J. Climate*, **28**, 9206–9220, doi:10.1175/JCLI-D-14-00753.1.

Donat, M. G., L. V. Alexander, H. Yang, I. Durre, R. Vose, and J. Caesar, 2013a: Global land-based datasets for monitoring climatic extremes. *Bull. Amer. Meteor. Soc.*, **94**, 997–1006, doi:10.1175/BAMS-D-12-00109.1.

——, and Coauthors, 2013b: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *J. Geophys. Res. Atmos.*, **118**, 2098–2118, doi:10.1002/jgrd.50150.

——, L. V. Alexander, N. Herold, and A. J. Dittus, 2016a: Temperature and precipitation extremes in century-long gridded observations, reanalyses, and atmospheric model simulations. *J. Geophys. Res. Atmos.*, **121**, 11 174–11 189, doi:10.1002/2016JD025480.

——, A. L. Lowry, L. V. Alexander, P. A. O'Gorman, and N. Maher, 2016b: More extreme precipitation in the world's dry and wet regions. *Nat. Climate Change*, **6**, 508–513, doi:10.1038/nclimate2941.

Fischer, E. M., U. Beyerle, and R. Knutti, 2013: Robust spatially aggregated projections of climate extremes. *Nat. Climate Change*, **3**, 1033–1038, doi:10.1038/nclimate2051.

Gallant, A. J. E., and D. J. Karoly, 2010: A combined climate extremes index for the Australian region. *J. Climate*, **23**, 6153–6165, doi:10.1175/2010JCLI3791.1.

——, K. J. Hennessy, and J. Risbey, 2007: Trends in rainfall indices for six Australian regions: 1910–2005. *Aust. Meteor. Mag.*, **56**, 223–241.

——, D. J. Karoly, and K. L. Gleason, 2014: Consistent trends in a modified climate extremes index in the United States, Europe, and Australia. *J. Climate*, **27**, 1379–1394, doi:10.1175/JCLI-D-12-00783.1.

Gleason, K. L., J. H. Lawrimore, D. H. Levinson, T. R. Karl, and D. J. Karoly, 2008: A revised U.S. climate extremes index. *J. Climate*, **21**, 2124–2137, doi:10.1175/2007JCLI1883.1.

Herold, N., L. V. Alexander, M. G. Donat, S. Contractor, and A. Becker, 2016: How much does it rain over land? *Geophys. Res. Lett.*, **43**, 341–348, doi:10.1002/2015GL066615.

Jones, P. W., 1999: First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon. Wea. Rev.*, **127**, 2204–2210, doi:10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2.

Karl, T. R., R. W. Knight, D. R. Easterling, and R. G. Quayle, 1996: Indices of climate change for the United States. *Bull. Amer. Meteor. Soc.*, **77**, 279–292, doi:10.1175/1520-0477(1996)077<0279:IOCCFT>2.0.CO;2.

Kharin, V. V., F. W. Zwiers, X. Zhang, and M. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345–357, doi:10.1007/s10584-013-0705-8.

Kim, Y.-H., S.-K. Min, X. Zhang, F. Zwiers, L. V. Alexander, M. G. Donat, and Y.-S. Tung, 2016: Attribution of extreme temperature changes during 1951–2010. *Climate Dyn.*, **46**, 1769–1782, doi:10.1007/s00382-015-2674-2.

King, A. D., and Coauthors, 2015: The timing of anthropogenic emergence in simulated climate extremes. *Environ. Res. Lett.*, **10**, 094015, doi:10.1088/1748-9326/10/9/094015.

Kunkel, K. E., X.-Z. Liang, J. Zhu, and Y. Lin, 2006: Can CGCMs simulate the twentieth-century "warming hole" in the central United States? *J. Climate*, **19**, 4137–4153, doi:10.1175/JCLI3848.1.

Meehl, G. A., J. M. Arblaster, and G. Branstator, 2012: Mechanisms contributing to the warming hole and the consequent U.S. east–west differential of heat extremes. *J. Climate*, **25**, 6394–6408, doi:10.1175/JCLI-D-11-00655.1.

——, ——, and C. T. Y. Chung, 2015: Disappearance of the southeast U.S. "warming hole" with the late 1990s transition of the interdecadal Pacific Oscillation. *Geophys. Res. Lett.*, **42**, 5564–5570, doi:10.1002/2015GL064586.

Min, S.-K., X. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, **470**, 378–381, doi:10.1038/nature09763.

——, ——, ——, H. Shiogama, Y.-S. Tung, and M. Wehner, 2013: Multimodel detection and attribution of extreme temperature changes. *J. Climate*, **26**, 7430–7451, doi:10.1175/JCLI-D-12-00551.1.

Moberg, A., and Coauthors, 2006: Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000. *J. Geophys. Res.*, **111**, D22106, doi:10.1029/2006JD007103.

Morak, S., G. C. Hegerl, and J. Kenyon, 2011: Detectable regional changes in the number of warm nights. *Geophys. Res. Lett.*, **38**, L17703, doi:10.1029/2011GL048531.

——, ——, and N. Christidis, 2013: Detectable changes in the frequency of temperature extremes. *J. Climate*, **26**, 1561–1574, doi:10.1175/JCLI-D-11-00678.1.

Pan, Z., R. W. Arritt, E. S. Takle, W. J. Gutowski Jr., C. J. Anderson, and M. Segal, 2004: Altered hydrologic feedback in a warming climate introduces a "warming hole." *Geophys. Res. Lett.*, **31**, L17109, doi:10.1029/2004GL020528.

——, X. Liu, S. Kumar, Z. Gao, and J. Kinter, 2013: Intermodel variability and mechanism attribution of central and southeastern U.S. anomalous cooling in the twentieth century as simulated by CMIP5 models. *J. Climate*, **26**, 6215–6237, doi:10.1175/JCLI-D-12-00559.1.

Poli, P., and Coauthors, 2016: ERA-20C: An atmospheric reanalysis of the 20th century. *J. Climate*, **29**, 4083–4097, doi:10.1175/JCLI-D-15-0556.1.

Portmann, R. W., S. Solomon, and G. C. Hegerl, 2009: Spatial and seasonal patterns in climate change, temperatures, and precipitation across the United States. *Proc. Natl. Acad. Sci. USA*, **106**, 7324–7329, doi:10.1073/pnas.0808533106.

Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau. *J. Amer. Stat. Assoc.*, **63**, 1379–1389, doi:10.1080/01621459.1968.10480934.

Seneviratne, S. I., and Coauthors, 2012: Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, C. B. Field et al., Eds., Cambridge University Press, 109–230.

Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh, 2013a: 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.

——, ——, F. W. Zwiers, X. Zhang, and D. Bronaugh, 2013b: Climate extreme indices in the CMIP5 multi-model ensemble.

Part 2: Future climate projections. *J. Geophys. Res. Atmos.*, **118**, 2473–2493, doi:10.1002/jgrd.50188.

Stephens, G. L., and Coauthors, 2010: Dreary state of precipitation in global models. *J. Geophys. Res.*, **115**, D24211, doi:10.1029/2010JD014532.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2006: How often does it rain? *J. Climate*, **19**, 916–934, doi:10.1175/JCLI3672.1.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Westra, S., L. V. Alexander, and F. W. Zwiers, 2013: Global increasing trends in annual maximum daily precipitation. *J. Climate*, **26**, 3904–3918, doi:10.1175/JCLI-D-12-00502.1.

You, Q., and Coauthors, 2011: Changes in daily climate extremes in China and their connection to the large scale atmospheric circulation during 1961–2003. *Climate Dyn.*, **36**, 2399–2417, doi:10.1007/s00382-009-0735-0.

Yu, S., and Coauthors, 2014: Attribution of the United States "warming hole": Aerosol indirect effect and precipitable water vapor. *Sci. Rep.*, **4**, 6929, doi:10.1038/srep06929.

Zhang, X., L. Alexander, G. C. Hegerl, P. Jones, A. Klein Tank, T. C. Peterson, B. Trewin, and F. W. Zwiers, 2011: Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 851–870, doi:10.1002/wcc.147.

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

Zhou, Y., and G. Ren, 2011: Change in extreme temperature event frequency over mainland China, 1961–2008. *Climate Res.*, **50**, 125–139, doi:10.3354/cr01053.