Geophysical Research Letters

RESEARCH LETTER
10.1029/2020GL091626

Quantifying the Occurrence of Record Hot Years Through Normalized Warming Trends

Xubin Zeng[1], J. E. Jack Reeves Eyre[1,2], Ross D. Dixon[1,3], and Jorge Arevalo[1,4]

[1]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA, [2]Cooperative Institute for Climate, Ocean and Ecosystem Studies, University of Washington, Seattle, WA, USA, [3]Department of Earth and Atmospheric Sciences, University of Nebraska—Lincoln, Lincoln, NE, USA, [4]Departamento de Meteorología, Universidad de Valparaíso, Valparaíso, Chile

Key Points:

- Occurrence of record hot years in different latitudes is better correlated with normalized, rather than raw, temperature trends
- Compared with the raw trends showing Arctic amplification, the normalized trends show a tropical amplification over land
- Earth system models' correlations between normalized trends and record-breaking events are as high over land as over ocean, unlike in observations

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

X. Zeng,
xubin@arizona.edu

Citation:

Zeng, X., Reeves Eyre, J. E. J., Dixon, R. D., & Arevalo, J. (2021). Quantifying the occurrence of record hot years through normalized warming trends. *Geophysical Research Letters*, *48*, e2020GL091626. https://doi.org/10.1029/2020GL091626

Received 7 NOV 2020
Accepted 4 MAY 2021

Author Contributions:

Conceptualization: Xubin Zeng
Data curation: J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo
Formal analysis: J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo
Funding acquisition: Xubin Zeng
Investigation: Xubin Zeng, J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo
Methodology: Xubin Zeng, J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo
Software: J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo
Supervision: Xubin Zeng
Validation: J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo
Visualization: J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo

Abstract Surface air temperature trends and extreme events are of global concern and they are related. Here, we show that the occurrence of record hot years over different latitudes from 1960 to 2019 are more strongly correlated with the observational annual mean temperature trends normalized by internal variability. Compared with the raw trends showing Arctic amplification, the normalized trends show a tropical amplification over land. Two hot spots with more frequent occurrence of record hot years are identified: northern hemisphere ocean (vs. land) and southern hemisphere tropical land (vs. mid- and high-latitude lands). Ensemble mean results from 32 Earth system models agree with observations better than individual models, but they do not reproduce observed large differences in correlations across latitudes between normalized trends and record-breaking events over land versus ocean. Our results enable the quantification of record hot year occurrence through normalized warming trends and provide new metrics for model evaluation and improvement.

Plain Language Summary Surface air temperature trends and extreme hot events are of global concern and they are related. Based on generic time series, it was argued that the occurrence of extreme events depends on the normalized, rather than the raw, trends. However, it remains unknown whether this dependence is relevant to the comparison of extreme hot event occurrences over different regions, such as the Arctic versus Amazon. Here, we show that the occurrence of record hot years over different regions from 1960 to 2019 are more strongly correlated with the observational annual mean temperature trends normalized by internal variability. Compared with the raw trends showing Arctic amplification, the normalized trends show a tropical amplification over land. Two hot spots with more frequent occurrence of record hot years are identified: Northern hemisphere ocean (vs. land) and southern hemisphere tropical land (vs. mid- and high-latitude lands). Results averaged from 32 Earth system models agree with observations better than individual models, but they do not reproduce observed large differences in correlations across latitudes between normalized trends and record-breaking events over land versus ocean. Our results enable the quantification of record hot year occurrence through normalized warming trends and provide new metrics for model evaluation and improvement.

1. Introduction

Numerous global warming papers have studied surface air temperature (T) trends (e.g., Hansen et al., 2010; Zeng & Geil, 2016) and the occurrence of extreme hot events (e.g., Schar et al., 2004; Diffenbaugh, 2020). Both the trends and extreme events are of global concern and they are related, as summarized in the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (IPCC, 2013). It is generally understood that warming trends would enhance the occurrence of extreme events at a given region. For instance, the Arctic amplification (based on the raw T trends) has been emphasized in all five IPCC reports, but it can overshadow regions with smaller internal variability, where trends don't need to be as large to break records and affect the environment, ecosystem, and human well-being. These regions can be identified by trends normalized by internal variability (e.g., as represented by the standard deviations of interannual variations). Indeed, using a theoretical approach to quantify the effect of long-term trends on the expected number of extremes in generic time series, Rahmstorf and Coumou (2011) argued that the occurrence of extreme events depends on the normalized, rather than the raw, trends at a given location.

Writing – original draft: Xubin Zeng
Writing – review & editing: J. E. Jack Reeves Eyre, Ross D. Dixon, Jorge Arevalo

However, it remains unknown whether this dependence is relevant to the comparison of extreme hot event occurrences over different regions, such as the Arctic versus Amazon.

The goal of this study is to address three questions: (a) How different are the spatial distributions of raw and normalized annual mean T trends from observations? (b) What are their correlations with extreme hot events, such as record-breaking annual mean T? and (c) How well do Earth system models (ESMs) participating in the Coupled Model Intercomparison Project (CMIP5 Taylor et al., 2012 and CMIP6 Eyring et al., 2016) perform compared with observations?

2. Materials and Methods

We primarily use the observational annual mean T data-set (at 1° × 1° grids) for 1880–2019 from the Berkeley Earth product (Rohde et al., 2013). Additional data-sets for the same period are also used for uncertainty quantification, including the NASA GISS Surface Temperature Analysis (GISTEMP) product (at 2° × 2° grids) (Hansen et al., 2010; Lenssen et al., 2019), the NOAA Merged Land Ocean Global Surface Temperature Analysis (NOAAGlobalTemp) product (at 5° × 5° grids) (Smith et al., 2008; Zhang et al., 2019), and the HadCRUT4 product (at 5° × 5° grids) (Morice et al., 2012). To ensure the least missing data for the study period, we focus on the results in the last 6 decades from 1960 to 2019. To complement these observational products with some missing data, we also use the most recent reanalysis (at 0.25° × 0.25° grids) from the European Center for Medium-Range Weather Forecasts (ERA5) without any missing data (Hersbach et al., 2020).

Furthermore, we use the surface air temperature output of 17 CMIP5 models and 15 CMIP6 models, as summarized in Table S1. We use a slightly different six-decade period of 1955–2014, since 2014 is the final year of the historical simulations in CMIP6. CMIP5 historical simulations end in 2005 and are extended to 2014 using simulations with the middle of the Representative Concentration Pathways (RCP4.5) (Meinhausen et al., 2011).

The linear trend of annual mean T at each grid box is computed using the least-square approach. The internal variability is computed as the standard deviation of the detrended T data. The normalized trend is simply the linear trend divided by internal variability. The raw (or normalized) trends are expressed as a ratio to the global mean trend (or normalized trend).

3. Results

3.1. Raw and Normalized Temperature Trends

Figure 1a shows that the observed annual mean T trends from 1880 to 2019 over most of the Arctic more than double the global mean warming trend, demonstrating the existence of polar amplification. The warming trends over land are greater than the global mean that, in turn, is greater than the trends over most of oceanic regions. For the normalized T trends, however, the polar amplification largely disappears, and the amplification shifts to some tropical regions such as the Amazon (Figure 1c).

Regionally, the observed trends over North America are greater than the global mean (Figure 1a), but the normalized trends are less than the global mean (Figure 1c). Over portions of the North Indian Ocean and South Atlantic, the situation is opposite, with the trends less than the global mean (Figure 1a) but the normalized trends greater than the global mean (Figure 1c).

Most grid boxes over Southern Hemisphere high latitudes have at least 20 years of missing data in Figure 1a. Global grids with valid data reached ~100% around 1960 for the Berkeley and GISTEMP datasets (Figure S1), while their percentages are lower for NOAAGlobalTemp and HadCRUT4 datasets. To ensure the least missing data for the study period, we focus on the results in the 6 decades from 1960 to 2019, as shown in Figure S2 for the four observational datasets and ERA5 reanalysis. The patterns in Figures 1a and 1c from 1880 to 2019 are also seen for the 1960–2019 period in Figures S2a and S2b.

The zonal mean trends from 1960 to 2019 show the Arctic amplification, which is stronger over ocean than over land (Figure 1b). Over midlatitudes and tropics, land warming is stronger than ocean warming. For

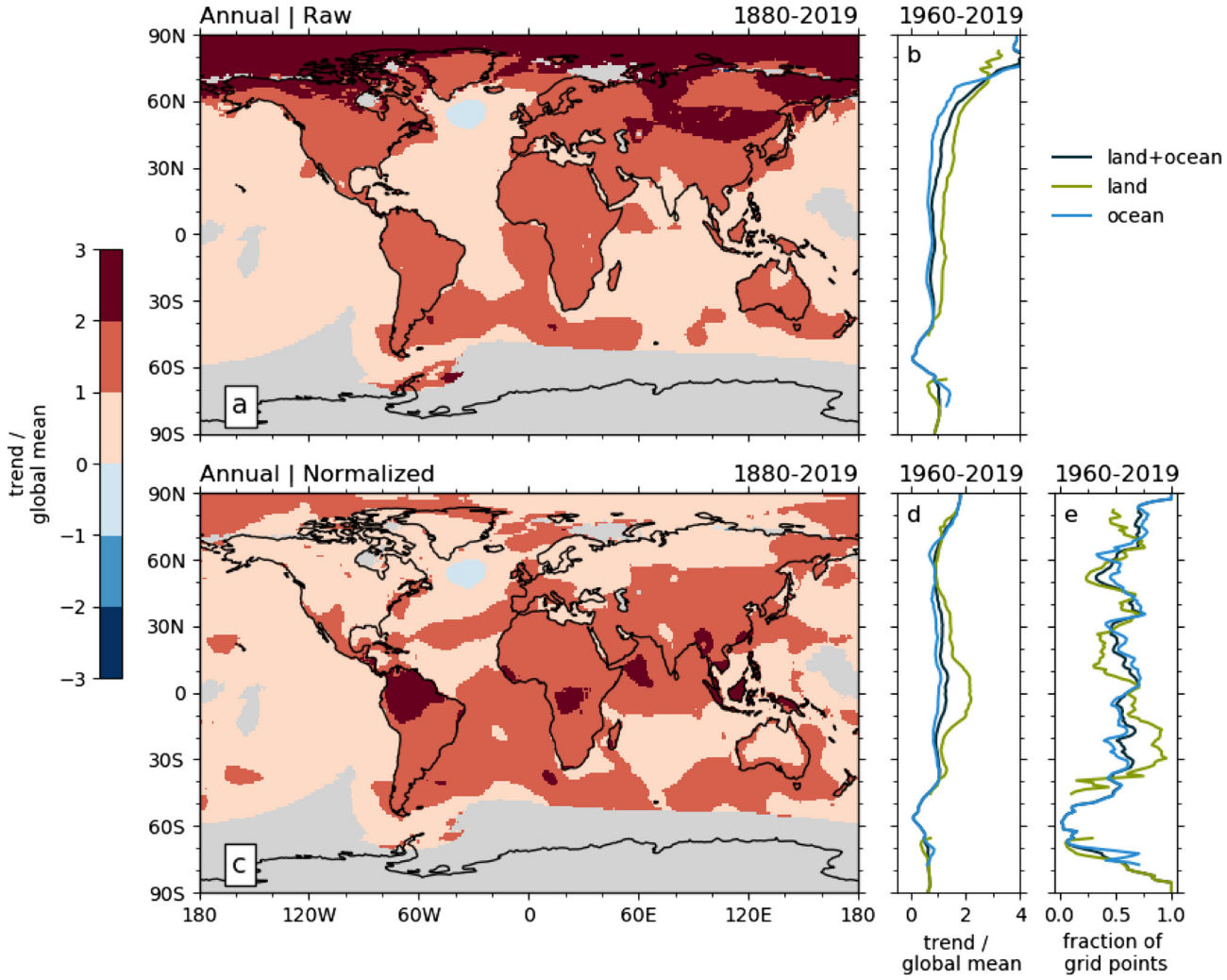

Figure 1. (a) Berkeley annual mean temperature trends from 1880 to 2019, expressed as a ratio to the global mean trend and (b) the zonal mean of this quantity for the period of 1960–2019 over land, ocean, and land + ocean. (c) and (d) Same as (a) and (b) but for normalized trends. (e) The fraction of grid boxes over land, ocean, and land + ocean across latitudes with the warmest years occurring in the last decade (2010–2019) during the 6 decades of 1960–2019. Gray shading in (a) and (c) represents grid boxes with more than 20 years of missing data. In (b), (d) and (e), grid boxes are excluded from the zonal mean if they have more than 5 years of missing data, and zonal means are not shown for latitudes with less than 10° longitude of valid data of the appropriate type (land or ocean).

the zonal mean normalized trends, there is a tropical amplification over land, along with a weaker Arctic amplification over ocean (Figure 1d).

The different spatial distributions of raw and normalized trends are confirmed using three other global observational datasets and ERA5 reanalysis (Figure S3). The question is: How are they correlated with extreme hot events?

3.2. Correlations of Trends With Extreme Hot Events

There are various metrics to characterize extreme hot events, and one simple metric is to compute when the warmest year occurred at each grid box. For a stationary annual time series (i.e., without any trend) for 6 decades, the probability for the warmest year to occur in the last decade is 1/6. Figure 1e shows that the fraction of grid boxes across latitudes with warmest years occurring in the last decade (2010–2019) during the 6 decades of 1960–2019 is much greater than 1/6 (i.e., 0.17) at most latitudes, because of global warming. Over land (excluding Antarctic), the highest fraction occurred over the tropics. Over ocean, high fractions

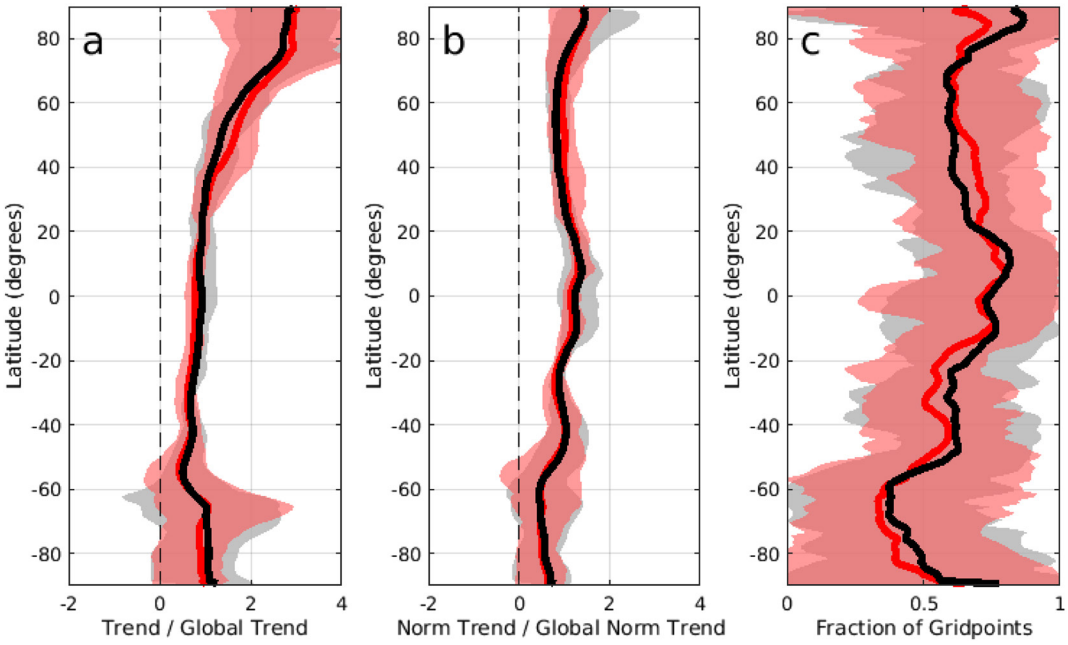

Figure 2. (a) The zonal mean of the annual temperature trends from 1955 to 2014, expressed as a ratio to the global mean trend, over land + ocean, from 17 Coupled Model Intercomparison Project (CMIP5) models (black) and 15 CMIP6 models (red) with filled envelopes showing the model spread and thick lines for the ensemble means. (b) Same as (a) but for normalized trends. (c) The fraction of grid boxes over land + ocean across latitudes with the warmest year occurring in the last decade (2005–2014) during the 6-decade period of 1955–2014.

occurred at low, mid, and high latitudes. The conclusions are similar using three other global observational datasets and ERA5 reanalysis (Figure S3).

Correlations of the fractions in Figure 1e with the raw trends in Figure 1b are much higher over land + ocean (0.46) and ocean (0.54) than over land (0.02). They are consistently lower by about 0.2 than those with the normalized trends in Figure 1d (0.66, 0.81, and 0.18, respectively). The coefficient of determination, or the correlation squared, represents the fraction of variance in Figure 1e that can be explained by the raw or normalized trends through a linear regression. Its ratios using the normalized versus raw trends over land + ocean is much greater than 1, varying from 1.8 to 2.8 for the four observational datasets, with the value from ERA5 in between (2.2) (Table S2).

While the raw and normalized trends over ocean are lower than those over land between 40°S and 70°N (Figures 1b and 1d), the fractions in Figure 1e are mostly higher over ocean between 5°N and 60°N, consistent with the stronger correlation over ocean than over land. This suggests that more attention should be paid to extreme hot events over ocean in these latitudes, consistent with the recognition of marine heatwaves and their impacts on ecosystem and environment in the recent IPCC Special Report (IPCC, 2019).

The higher correlation of the fractions with the normalized trends are confirmed using three other global observational datasets and ERA5 reanalysis (Table S2). The question is: Can CMIP models reproduce these observed results?

3.3. CMIP Model Evaluations

CMIP5 and CMIP6 ensemble mean results are similar (Figure 2 and S4). These results are also similar to observations in Figure 1, such as the strong Arctic amplification in the raw trends, much weaker Arctic amplification in the normalized trends, and high fractions occurring over different latitudes. However, while observations show a large contrast between land and ocean in normalized trends (Figure 1d) and record-breaking events (Figure 1e), these contrasts are quite small in the ensemble mean results (Figure S4).

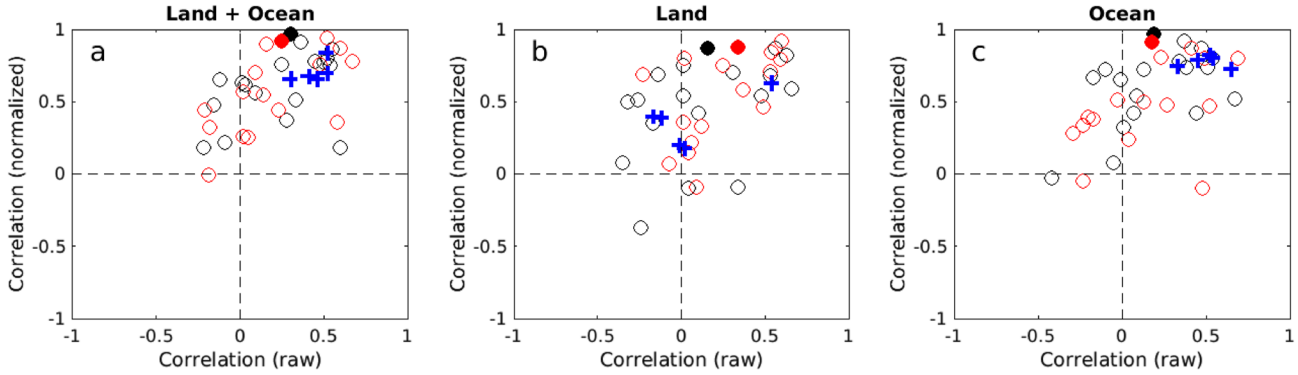

Figure 3. Scatterplots of correlation coefficients computed using raw trends versus those using normalized trends for (a) Land + ocean, (b) Land, and (c) Ocean. Each correlation is computed across latitudes between the zonal mean trend (raw or normalized) and the fraction of grid boxes (at each latitude) with the warmest year occurring in the final decade of the six-decade period of 1955–2014. Black (or red) open circles represent results for individual Coupled Model Intercomparison Project (CMIP5) (or CMIP6) models, with the filled circles for correlations using the multi-model ensemble means. Results for specific models are provided in Table S1. Blue "+" denotes results from the four observational products and ERA5 reanalysis.

Figure 3 and Table S1 provide the correlations across latitudes of the fractions in Figures 2c and S4c with the raw and normalized trends for each model. For ensemble means, the stronger correlations with the normalized trends (than with the raw trends) are qualitatively consistent with observations. Quantitatively, however, the ratios of the coefficient of determination using the normalized versus raw trends over land + ocean are much greater (>10, Table S1) than observations in Table S2 (1.8–2.8, as mentioned earlier). The correlations with the normalized trends from CMIP5 to CMIP6 ensemble means are similar over land versus over ocean (Figure 3 and Table S1), in contrast to the large differences from observations and ERA5 reanalysis (with an average of 0.36 over land and 0.78 over ocean; Figure 3 and Table S2). The correlations for individual models are widely scattered, and partly for this reason, the observational and reanalysis results are within the spread of the multi-model ensemble over land + ocean, land, and ocean (Figure 3).

Only one ensemble member from each model is used in Figure 3. To assess the impact of internal climate variability on the results of individual models (e.g., Dai & Bloecker, 2019), Figure 4 shows the corresponding results from four CMIP6 models with 11–32 ensemble members. The ensemble mean results from two models (CanESM5 and CESM2) are largely consistent with observations, while those from one model (CNRM-CM6-1) have very different results from observations, with a correlation with normalized (or raw) trends of about 1 (−0.5). Similar to those in Figure 3, the correlations with the normalized trends from each model's ensemble means are similar over land versus over ocean, in contrast to the large differences from observations and ERA5 reanalysis (Figure 4). For each model, the observational and reanalysis results are within the spread of the widely-scattered ensemble over land-ocean, land, and ocean, also similar to those in Figure 3.

In general, ensemble mean results do not reproduce the relationships seen in observations, because ensemble means largely remove internal climate variability while observations include such variability. At the same time, at 60 year time scale, a large portion of the known internal climate variability is removed (e.g., Dai & Bloecker, 2019; Zeng & Geil, 2016), so the above-mentioned large differences between ensemble means and observations are unexpected. To rigorously compare model results against observations, a large ensemble (e.g., with >100 members) is needed to generate a reliable probability distribution which can be used to ascertain model-observation agreement (or disagreement) if observations fall within (or outside of) the given (e.g., 5th–95th) percentile range.

4. Conclusions

The observational data analyses and model evaluations, here, suggest that the normalized T trends deserve more attention to understand the occurrence of extreme hot events over different regions, and both raw and normalized trends and their relations with extreme hot events should all be included as metrics in ESM development, tuning, and evaluation.

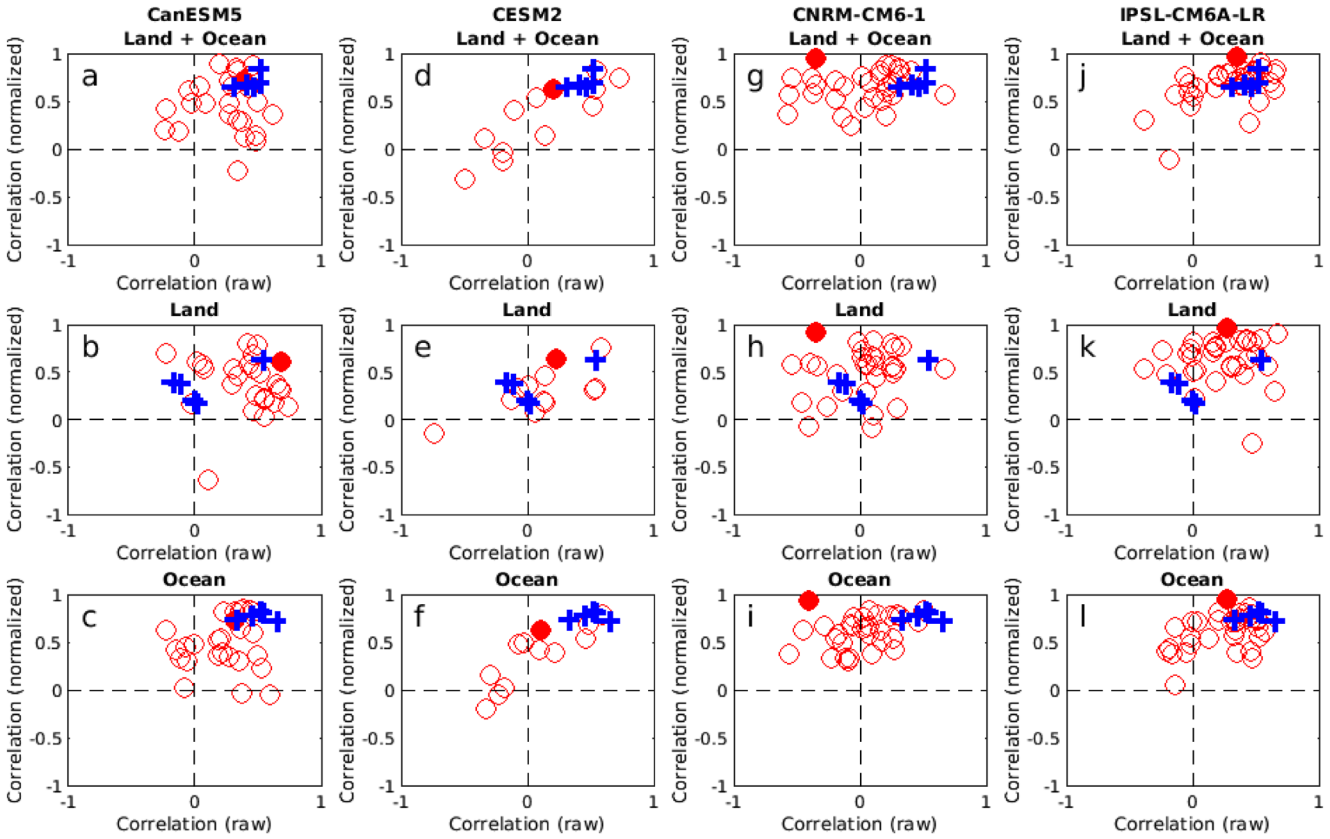

Figure 4. Scatterplots of correlation coefficients computed using raw trends versus those using normalized trends for land + ocean (top row), land (middle row), and ocean (bottom row). Each correlation is computed across latitudes between the zonal mean trend (raw or normalized) and the fraction of grid boxes (at each latitude) with the warmest year occurring in the final decade of the six-decade period of 1955–2014. Results for four Coupled Model Intercomparison Project (CMIP6) models are shown in four columns. Red open circles in each column represent results for individual ensemble members of a CMIP6 model, with the filled circles for correlations using the ensemble means. Blue "+" denotes results from the four observational products and ERA5 reanalysis.

Our data analyses here have also identified two hot spots where record-breaking hot years occur more frequently: Northern hemisphere ocean (vs. land) and southern hemisphere tropical land (vs. mid- and high-latitude lands), and they should receive more attention due to their significant impacts on ecosystem and environment.

Further studies are needed on the relationship of raw and normalized trends with other metrics of extreme hot events in observational data analysis and CMIP model evaluations. Also needed are the studies using seasonal or monthly mean temperature data. Furthermore, research is needed to understand how an individual model's performance for the 1960–2019 period is related to its projection of extreme hot events in the future. Finally, a large ensemble (e.g., with >100 members) should be used to more rigorously compare ESM results with observations.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability Statement

The HadCRUT4 temperature product were obtained from www.metoffice.gov.uk/hadobs/hadcrut4/, Berkeley Earth product from berkeleyearth.org/data, GISTEMP product from data.giss.nasa.gov/gistemp/, and NOAAGlobalTemp product from www.ncei.noaa.gov/data/noaa-global-surface-temperature/v5/. The ERA5 reanalysis data were obtained from the Copernicus Climate Data Store.

Acknowledgments
This research was supported by the DOE ESM Program (DE-SC0016533 and through the LLNL subcontract B639244) and the NASA MAP Program (NNX14AM02G). We thank the anonymous reviewer for constructive comments and suggestions for our significant revision of the original manuscript. We acknowledge the World Climate Research Program, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP5 and CMIP6. We thank the climate modeling groups for producing and making available their model output, and the Earth System Grid Federation (ESGF) for archiving the data and providing access.

References

Dai, A., & Bloecker, C. E. (2019). Impacts of internal variability on temperature and precipitation trends in large ensemble simulations by two climate models. *Climate Dynamics*, *52*, 289–306. https://doi.org/10.1007/s00382-018-4132-4

Diffenbaugh, N. S. (2020). Verification of extreme event attribution: Using out-of-sample observations to assess changes in probabilities of unprecedented events. *Science Advances*, *6*, eaay2368. https://doi.org/10.1126/sciadv.aay2368

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. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Reviews of Geophysics*, *48*, RG4004. https://doi.org/10.1029/2010RG000345

Hersbach, H., Berrisford, B.P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*, 1999–2049. https://doi.org/10.1002/qj.3803

IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of working group I to the Fifth assessment report of the intergovernmental Panel on climate Change. In T. F. Stocker (Ed.). Cambridge University Press.

IPCC (2019). Special report on the Ocean and cryosphere in a changing climate. H.-O. Pörtner, et al. (Ed.).

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A, Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP uncertainty model. *Journal of Geophysical Research: Atmospheres*, *124*(12), 6307–6326. https://doi.org/10.1029/2018JD029522

Meinhausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L., Lamarque, J. F., et al. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, *109*. https://doi.org/10.1007/s10584-011-0156-z

Morice, C. P., Kennedy, J. J., Rayner, N. A, & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research*, *117*, D08101. https://doi.org/10.1029/2011JD017187

Rahmstorf, S. & Coumou, D. (2011). Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 17905–17909. https://doi.org/10.1073/pnas.1101766108

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., Wurtele, J., et al. (2013). Berkeley Earth temperature averaging process. *Geoinformatics & Geostatistics*, *1*, 2. https://doi.org/10.4172/gigs.1000103

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A, & Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature*, *427*, 332–336. https://doi.org/10.1038/nature02300

Smith, T. M., Reynolds, R. W., Peterson, T. C., & Lawrimore, J. (2008). Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). *Journal of Climate*, *21*, 2283–2296. https://doi.org/10.1175/2007JCLI2100.1

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Zeng, X. & Geil, K. (2016). Global warming projection in the 21st century based on an observational data-driven model. *Geophysical Research Letters*, *43*, 10947–10954. https://doi.org/10.1002/2016GL071035

Zhang, H.-M., Huang, J. B., Menne, M., Yin, X., Sánchez-Lugo, A., Gleason, B., et al. (2019). Updated temperature data give a sharper view of climate trends. *Eos*, *100*. https://doi.org/10.1029/2019EO128229