Quantifying the Human Influence on the Intensity of Extreme 1- and 5-Day Precipitation Amounts at Global, Continental, and Regional Scales

QIAOHONG SUN,^a FRANCIS ZWIERS,^{a,b} XUEBIN ZHANG,^c AND JUN YAN^d

^a *Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia, Canada*

^b *Nanjing University of Information Science and Technology, Nanjing, China*

^c *Climate Research Division, Environment and Climate Change Canada, Toronto, Ontario, Canada*

^d *Department of Statistics, University of Connecticut, Storrs, Connecticut*

(Manuscript received 12 January 2021, in final form 14 September 2021)

ABSTRACT: This study provides a comprehensive analysis of the human contribution to the observed intensification of precipitation extremes at different spatial scales. We consider the annual maxima of the logarithm of 1-day (Rx1day) and 5-day (Rx5day) precipitation amounts for 1950–2014 over the global land area, four continents, and several regions, and compare observed changes with expected responses to external forcings as simulated by CanESM2 in a large-ensemble experiment and by multiple models from phase 6 of the Coupled Model Intercomparison Project (CMIP6). We use a novel detection and attribution analysis method that is applied directly to station data in the areas considered without prior processing such as gridding, spatial or temporal dimension reduction, or transformation to unitless indices and uses climate models only to obtain estimates of the space–time pattern of extreme precipitation response to external forcing. The influence of anthropogenic forcings on extreme precipitation is detected over the global land area, three continental regions (the western Northern Hemisphere, western Eurasia, and eastern Eurasia), and many smaller IPCC regions, including central North America, East Asia, east-central Asia, eastern Europe, eastern North America, northern Europe, and western Siberia for Rx1day, and central North America, eastern Europe, eastern North America, northern Europe, the Russian Arctic region, and western Siberia for Rx5day. Consistent results are obtained using forcing response estimates from either CanESM2 or CMIP6. Anthropogenic influence is estimated to have substantially decreased the approximate waiting time between extreme annual maximum events in regions where anthropogenic influence has been detected, which has important implications for infrastructure design and climate change adaptation policy.

SIGNIFICANCE STATEMENT: All previous detection and attribution studies of observed changes in extreme precipitation (i) use station data that has been heavily processed via gridding, transformation, and spatial and temporal averaging or other dimension reduction approaches, as well as using climate models to estimate the responses to external forcing, and (ii) also use models to estimate the unforced natural variability of extreme precipitation. Both aspects reduce user confidence in detection and attribution results. This study uses station data directly and avoids difficult to verify model-based estimates of the unforced variability of precipitation extremes. Results confirm findings from previous studies, and extend them to a number of subcontinental regions, thus substantially increasing confidence in detection and attribution findings precipitation.

KEYWORDS: Extreme events; Pattern detection; Climate models

1. Introduction

Precipitation extremes represented by the annual maximum precipitation accumulations on one day (Rx1day) or five consecutive day (Rx5day) can cause severe impacts on society and are relevant to the design of infrastructure (Canadian Commission on Building and Fire Codes 2015; Sobie and Murdock 2014; Hershfield 1961). Intensification of such annual maxima has been observed over global land, different

continents (Min et al. 2011; Paik et al. 2020; Kirchmeier-Young and Zhang 2020), and in some regions, including notably dry and wet regions (Donat et al. 2016), the high latitudes of the Northern Hemisphere (Groisman et al. 2005; Westra et al. 2013), central North America, eastern North America, northern Central America, northern Europe, the Russian Far East, eastern central Asia, and East Asia (Sun et al. 2021). Previous detection and attribution analyses (Min et al. 2011; Zhang et al. 2013; Li et al. 2017; Dong et al. 2020; Kirchmeier-Young and Zhang 2020; Paik et al. 2020) suggest that human-induced increases in greenhouse gases have contributed to these observed changes.

Optimal fingerprint detection and attribution analysis is a widely used, standard method for detection and attribution (D&A) analysis in which observations are regressed via generalized total least squares (Allen and Stott 2003; Ribes et al. 2013) onto model simulated signals. However, this method assumes that regression residuals follow a symmetric Gaussian-like

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-21-0028.s1>.

Corresponding author: Qiaohong Sun, sunqh@uvic.ca

DOI: 10.1175/JCLI-D-21-0028.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

distribution, which may not be suitable for extreme values such as Rx1day and Rx5day that tend to have skewed distributions and heavy-tailed behavior. In previous studies, several approaches have been proposed, most of which avoid directly addressing the statistical properties of extremes. This is usually done by transforming extremes to a probability scale to obtain so-called probability indices (PIs) and applying optimal fingerprinting to temporal and spatial averages of the transformed data, as well as further dimension reductions in some cases (Min et al. 2011; Zhang et al. 2013; Paik et al. 2020; Kirchmeier-Young and Zhang 2020). We refer to these approaches as PI-based D&A approaches in the following section. Others have made inferences based on optimal fingerprinting applied to the parameters of extreme value distributions fitted to gridded observations within regions over decadal periods (e.g., Christidis et al. 2011) or averages of extremes over large regions in the case of surface air temperature (Wen et al. 2013; Yin et al. 2016; Seong et al. 2020). While such approaches are successful in detecting the effects of external forcing on extremes, none of the techniques developed so far can be applied directly to station data, and all require climate model-based estimates of the natural unforced variability of extremes.

Zwiers et al. (2011) proposed an approach for the detection and attribution analysis of changes in climate extremes that directly represents the statistical properties of the observed extremes to which it is applied via the use of extreme value distributions, and used their approach to study observed changes in temperature extremes, using forced signals estimated from CMIP3 simulations. In this approach, generalized extreme value (GEV) distributions are fitted to the observed annual temperature extremes using a time-evolving pattern of location parameters obtained from model-simulated extreme temperatures under anthropogenic (ANT) forcing or anthropogenic and natural (ALL) forcing combined. Simultaneous fitting of the GEV distributions at grid boxes in the domain of interest was performed using a profile likelihood method. Computational constraints meant that it was not possible to simultaneously consider multiple forced signals in the detection and attribution analysis. Wang et al. (2017) improved the method by using a coordinate descent algorithm for fitting, which is much faster than the profile likelihood approach and thus makes it feasible to perform a multiple-signal detection and attribution analysis for changes in temperature extremes. Wang et al. (2020) recently further improved the technique by using a combined score equation (CSE) method based on a weighted sum of marginal GEV scores to further improve computational efficiency and enable detection and attribution inferences that account for spatial dependence. They demonstrated the performance of this optimized technique by again considering observed changes in annual extreme temperatures, and showing that human influence on temperature extremes is detectable at the subcontinental scale.

Application of this method to precipitation requires consideration of the differences in scale between precipitation extremes simulated by climate models, which are the extremes of model-simulated gridbox-average precipitation amounts, and point scale observations from weather stations. Owing to the discontinuous nature of precipitation in both

time and space, gridding observed Rx1day and Rx5day values at stations into a coarser spatial resolution does not produce values that correspond to the extremes of precipitation averaged over grid boxes as simulated by climate models (Zhang et al. 2011). Also, such gridding may not produce values that can be interpreted as estimates of the extremes of gridbox-average precipitation in the observed climate. Here we will therefore adapt the updated Wang et al. (2020) technique to allow us to take spatial patterns of change in extreme precipitation into consideration and perform detection analysis for station data directly. The application of this approach, which is structurally different from that of previous studies, will help to enhance our confidence in detection and attribution results concerning extreme precipitation—an area where the evidence of human influence is still emerging in particular at the continental and smaller spatial scales.

We also expand the detection and attribution analysis to the regional scale to determine whether changes in extreme precipitation due to anthropogenic greenhouse gas changes are detectable and attributable at that scale. Detection and attribution of extreme precipitation changes at the regional and local scales is more difficult than at global or continental scales due to lower signal-to-noise ratios, poor observational coverage (Stott et al. 2010; Sarojini et al. 2016), and greater uncertainty concerning the representation of subgrid scale precipitation related processes in climate models (Fowler et al. 2021). Nevertheless, continued efforts to collect data have resulted in improved spatial and temporal coverage of observations, and analysis of this data shows (Sun et al. 2021) that the percentage of stations with significant increasing trends is significantly larger than that can be expected by chance in many regions, including central North America, eastern North America, northern Central America, northern Europe, the Russian Far East, eastern central Asia, and East Asia.

In this paper we consider whether human influence on extreme precipitation is robustly detectable at different scales using different datasets with a novel detection and attribution analysis method in which changes at stations are evaluated by fitting nonstationary GEV distributions at individual stations to the logarithms of observed Rx1day and Rx5day values. We will examine changes in extreme precipitation as recorded in the Sun et al. (2021) dataset using estimates of externally forced changes from two ensembles of historical climate change simulations: a large ensemble simulation performed with CanESM2 (Arora et al. 2011), which is a coupled Earth system model developed and run by the Canadian Centre for Climate Modeling and Analysis (CCCma) that participated in phase 5 of the Coupled Model Intercomparison Project (CMIP5), and a multimodel ensemble of simulations from CMIP6. The remainder of this paper is structured as follows. Section 2 describes the observational data and model simulations used in this study and the detection and attribution technique, including a detailed description of how uncertainty from natural unforced variability is considered. Detection and attribution analysis results are provided in section 3. Conclusions and discussion are given in section 4.

FIG. 1. Locations of long-term stations for extreme indices (dots), divisions of 4 continents (blue boxes), and 16 regions (red boxes) used in the IPCC WGI Sixth Assessment Report.

2. Data and methods

a. Data and data processing

We use Rx1day and Rx5day from a stations-based observational dataset, which includes 14 796 land-based observing stations with more than 30 years of record during 1900–2018 across the globe (Sun et al. 2021). These station observations are collected from various sources that contributed to HadEX2 (Donat et al. 2013) with additional stations that were not available to HadEX2 such as a large collection of Chinese and Russian data. To consider long-term changes, we first selected 5081 stations that have records more than 45 years during 1950–2014, including at least 3 years during 2010–14.

We perform the detection and attribution analysis at the global (GLB), continental, and regional spatial scales. At the continental scale, we divided the global land area into four continents as in previous studies (Zhang et al. 2013; Paik et al. 2020), including the western Northern Hemisphere land area (NA; 45°W–180°), western Eurasia (EU; 25°W–60°E) and eastern Eurasia (AS; 60°E–180°) from the Northern Hemisphere, and the region of Oceania (OC; 110°E–180°) from the Southern Hemisphere. At the regional scale, we selected 16 regions from Iturbide et al. (2020) for IPCC AR6 reference land regions with more than 30 stations that satisfy the data completeness criteria described above (Fig. 1). Iturbide et al. (2020) updated the previous set of IPCC AR5 Working Group I (WGI) reference regions for the regional synthesis of model-projected climate change information (in particular for the new CMIP6 simulations).

We used two large ensemble simulations produced with CanESM2 (Arora et al. 2011; <https://open.canada.ca/data/en/dataset/aa7b6823-fd1e-49ff-a6fb-68076a4a477c>), each containing 50 perturbed initial condition members—one with anthropogenic and natural forcings combined (ALL forcings) from

1950 to 2020, and the other with natural (NAT) forcings for the same period. The model, forcings, variable names, and file formats all follow those used in CMIP5. The NAT simulations were run to 2005 using observed estimates of historical changes in solar irradiance and volcanic aerosol loadings, and were extended into 2020 by repetition of the 1997–2008 solar cycle (Arora et al. 2011). The ALL simulations were run to 2005 using CMIP5 historical forcings that combine natural and anthropogenic forcings (such as greenhouse gases, aerosol emissions, and land use), and then to 2020 using representative concentration pathway 8.5 forcings.

Multimodel climate simulations from the CMIP6 experiments are also utilized. Simulations from 10 CMIP6 models (Table 1) that provide data with at least three ensemble members for both ALL and NAT forcing are used in this study. To ensure that both ALL and NAT signal estimates are equally affected by sampling uncertainty, we required the ALL and

TABLE 1. List of CMIP6 model simulations used in this study.

Model	All	NAT	Ensemble size
ACCESS-CM2	✓	✓	3
BCC-CSM2-MR	✓	✓	3
CanESM5	✓	✓	10
CNRM-CM6-1	✓	✓	3
CESM2	✓	✓	3
HadGEM3-GC31	✓	✓	4
IPSL-CM6A-LR	✓	✓	9
MIROC6	✓	✓	3
MRI-ESM2-0	✓	✓	5
NorESM2-LM	✓	✓	3
Total	10 models, 46 runs for each forcing		

NAT ensembles from each model to be of equal size. In total, 10 models met this requirement and together provided a total of 46 runs for each forcing. As the historical simulations for CMIP6 are available through 2014, we perform analysis over 1950–2014, allowing analysis of observations of a period that is 10 years longer than most previous results using simulations from CMIP5. Annual Rx1day and Rx5day values are extracted from CanESM2 and CMIP6 simulations.

b. Detection and attribution method

Extreme value theory suggests that the behavior of observed annual extremes can be approximated with a generalized extreme value (GEV) distribution with location, scale, and shape parameters μ , σ , and ξ , respectively. Nonstationary behavior can be represented by treating these parameters as functions of covariates that alter the characteristics of extremes. In detection and attribution analyses, the parameters of the GEV distribution can be treated as functions of climate change signals estimated from climate model simulations to determine whether the observations provide evidence for the presence of those signals. In Wang et al. (2020), observed extremes $\mathbf{Y}_{ts}$ at time t and geographical location s were approximated by a GEV distribution with a location parameter μ_{ts} that varies with t and s and scale and shape parameters σ_s and ξ_s that vary only with s . The forced signals were assumed to affect the location parameter via a simple linear additive model

$$\mu_{ts} = \alpha_s + \mathbf{X}_{ts}\boldsymbol{\beta}, \quad (1)$$

where $t = 1, \dots, n$ is time (yr), $s = 1, \dots, m$ is geographical location, and the signal matrix $\mathbf{X}_{ts}$ contains location parameters from GEV distributions fitted to ensembles of forced simulations, with one column for each type of forcing considered. The GEV parameters that are to be estimated from the observations are therefore α_s , σ_s , ξ_s and the scaling factor vector $\boldsymbol{\beta}$, which does not depend on location.

This approach is appropriate for temperature, since we would expect forcing to shift the distribution to the right on the temperature scale, but it is less appropriate for precipitation, because scaling arguments suggest an elongation of the precipitation distribution to the right rather than a simple shift to the right. Also, assuming $\boldsymbol{\beta} = 1$, this supposes that any scale mismatch between models and observations can be accounted for simply by shifting the location to the left or right. This seems contrary to results from recent temperature scaling research (e.g., Li et al. 2019a,b, 2021; Sun et al. 2020) and PI-based D&A approaches (Min et al. 2011; Zhang et al. 2013; Paik et al. 2020), which suggests that temperature scaling of precipitation extremes in climate models is consistent with that in observations despite the differences between the type of precipitation extremes represented by models (annual extremes of gridbox averages) and observations from weather stations (annual extremes of point values).

The Clausius–Clapeyron relation suggests that, broadly speaking, it would be reasonable to expect extreme precipitation to intensify exponentially with warming, as is seen in both observations and models (Sun et al. 2021; Li et al. 2021). Thus, the extremes of the logarithm of precipitation should

increase more or less linearly with temperature. To take advantage of the possibility that both observed and modeled precipitation extremes may scale similarly with temperature, we fit GEV distributions to log-transformed observations, $\ln(\text{Rx1day})$ and $\ln(\text{Rx5day})$, with GEV location parameters of the form

$$\mu_{ts} = \mu_{bs} + \sum_{i=1, \dots, l} \beta_i \ln(\gamma_{tsi}), \quad (2)$$

where b is a fixed base time and $\boldsymbol{\gamma}_{ts} = (\gamma_{ts1}, \dots, \gamma_{tsl})^T$ is an $l \times 1$ vector containing the signals from l forcings of interest, expressed as the ratio between the median of Rx1day or Rx5day at time t and that at time b . As above, we assume that the scale and shape parameters σ_s and ξ_s vary only with geographical location s . The quantile function for this model of the log of $\mathbf{Y}_{ts}$ has the form

$$Q_{\ln(\mathbf{Y}_{ts})}(p) = \mu_{bs} + \sum_{i=1, \dots, l} \beta_i \ln(\gamma_{tsi}) + \sigma_s Q(p; \xi_s), \quad (3)$$

where $Q(p; \xi)$ is the p -quantile of a standard GEV distribution with location zero, scale 1, and shape ξ ; that is, $Q(p; \xi) = (1/\xi_s) \{[-\ln(p)]^{-\xi_s} - 1\}$. Thus, the corresponding quantiles of Rx1day and Rx5day have the form

$$Q_{\mathbf{Y}_{ts}}(p) = e^{\mu_{bs} + \sigma_s Q(p; \xi_s)} \prod_{i=1, \dots, l} \gamma_{tsi}^{\beta_i}. \quad (4)$$

The quantiles are constant in time (as they should be) if the β_i values are zero or if $\prod_{i=1, \dots, l} \gamma_{tsi}^{\beta_i} = 1$. On the other hand, the quantiles will exhibit scaling that is inherited from $\prod_{i=1, \dots, l} \gamma_{tsi}^{\beta_i}$ if they are not constant in time. In the one-signal case, for example, if Clausius–Clapeyron scaling with warming is seen in that signal, and if the scaling factor for that signal is near unity, then the quantiles will scale similarly with warming.

A potential limitation is that the scaling is identical for all quantiles. There is some evidence from climate models that scaling may vary with rarity (Kharin et al. 2018; Li et al. 2019a, 2021), but this would be hard to discern reliably from observations. On balance, however, we think this statistical model reflects a reasonable set of compromises while offering the ability to incorporate expected scaling behavior in the center of the Rx1day or Rx5day distribution, allowing signal representation in a relative (rather than absolute) form, and enabling robust noise reduction in the signal (discussed below) through smoothing in time and space. Thus we apply the linear additive model to the location parameter of GEV distributions fitted to samples of observed $\ln(\text{Rx1day})$ and $\ln(\text{Rx5day})$ values.

1) SIGNAL ESTIMATION FROM ENSEMBLE SIMULATIONS

The signal vector $\boldsymbol{\gamma}_{ts}$ is obtained from climate model output. A large body of scaling research and model assessment papers (e.g., Kharin et al. 2013; Li et al. 2021) suggests that $\boldsymbol{\gamma}_{ts}$ varies smoothly in space, at least for the ALL, anthropogenic (ANT), and natural (NAT) signals. There are regions, such as regions affected by changes in the downwelling branches of the Hadley circulation, where we expect these factors to

decrease over time, eventually attaining values less than one, but these are not areas where we have very much station data. There are also some areas where we might expect these factors to increase relatively quickly over time well past one (corresponding to super Clausius–Clapeyron scaling), particularly in tropical convergence zones—but we again have relatively few data in such areas over land. Generally, we could however expect homogeneous scaling over very broad regions. This suggests that the noise in the γ_{ts} estimates could be reduced by averaging over large regions. Here, we use a $k \times k$ gridbox moving window to smooth the signal across space and use a quadratic B -spline to smooth the signal across time. An advantage of using the B -splines is that we do not need to make the assumption that the location parameter is constant within 10-yr periods, as in previous studies (Zwiers et al. 2011; Wang et al. 2017).

Raw signal estimation at each location from an individual climate model using a specified forcing (described for Rx1day) is as follows:

- 1) For the entire base period 1950–2014, pool Rx1day values from within a $k \times k$ gridbox region from all ensemble members and estimate the median; for example, if $k = 3$ and there are four ensemble members, then the median is obtained from $9 \times 65 \times 4 = 2340$ Rx1day values.
- 2) For each year, pool Rx1day values from the $k \times k$ grid boxes and from all ensemble members; estimate the median and divide by the median for the base period. Continuing the example, for each year, the median value in the numerator is obtained from $9 \times 4 = 36$ Rx1day values. Assign this ratio to the location at the center of the region at this year.
- 3) Apply a quadratic B -spline with knots placed every N years to the raw signal estimates to obtain the smoothed signals $\hat{\gamma}_{ts}$ across time. The B -spline smoothing procedure is described in detail in the online supplemental material and is illustrated in Fig. S1.

The signal estimation process described above is applied directly to the CanESM2 simulations, and is further modified as follows for the multimodel CMIP6 ensemble so that each model is given equal weight in the final smoothed signal estimate:

- 1) For each year, interpolate the raw signals from individual models to a common grid.
- 2) Average the model signals to obtain a multimodel signal field on the interpolation grid.

At each location on the interpolation grid, fit a B -spline with knots at every N years for that location to obtain smoothed signal estimates $\hat{\gamma}_{ts}$.

The result for both ensembles is a smooth function in time and space describing the model response to the forcing that was applied in those ensembles (either ALL or NAT). The choice of moving window size k reflects a trade-off between the amount of spatial smoothing, and thus reduction of noise from internal climate variability, and avoidance of overly smoothing spatial features that reflect regions with heterogeneous precipitation climatologies. Here, we use $k = 3$. We expect the signal to be smooth in space because modeling evidence suggests scaling is broadly homogeneous, and roughly consistent with the Clausius–Clapeyron relation, over broad

land areas in the extratropics (where most observations are located). On the other hand, the number of knots determines the degree of temporal smoothing. We placed spline knots at 5-yr intervals, which again reflects a trade-off between noise reduction and the loss of the signal detail associated with episodic volcanic forcing events, such as associated with the eruption of Mt. Pinatubo in 1991.

We estimate the smoothed signal for ALL and NAT forcing at each grid box. The signals are expressed as the ratio between the median of annual maxima at time t and that at a fixed base time and then log-transformed [Eq. (2)]. Thus estimates of the ANT forcing signal are obtained as the ratio between the ALL and NAT signal estimates (ALL/NAT). Assuming that forcing i causes intensification of Rx1day over time in a region centered on location s , we should see $\hat{\gamma}_{tsi} < 1$ when t is near the beginning of the 1950–2014 analysis period, $\hat{\gamma}_{tsi} \approx 1$ in the middle of the period, and $\hat{\gamma}_{tsi} > 1$ toward the end of the period. We would expect ANT forcing to produce such a response. An example of a smoothed ALL signal in Rx1day from CanESM2 at a midlatitude Northern Hemisphere location (latitude = 54.4162°N, longitude = 19.6875°E) is displayed in Fig. S1 in the online supplemental material. Signals obtained in this way well represent observed changes in extreme precipitation where they are detected in observations. Anticipating one such result, which will be discussed in section 3, Fig. S2 shows the comparison between changes in the ALL signal estimates from CanESM2 and CMIP6 over time and changes in the observations over northern Europe, indicating increasing trends in both ALL signals and the observations.

2) DETECTION ANALYSIS

Once the signals are obtained, GEV distributions with location parameters described by (2) can be fitted to $\ln(\text{Rx1day})$ and $\ln(\text{Rx5day})$. Since we expect the scaling to be more or less homogeneous in space and independent of measurement scale, we do not grid the observational data to a grid scale consistent with that of the models. Instead, using Eq. (2), we fit nonstationary GEV distributions to all station values in a domain of interest directly. Wang et al. (2020) proposed a combined score equation (CSE) fitting method based on a weighted sum of marginal GEV scores (i.e., gradients of the log-likelihood function) to account for spatial dependence and improve efficiency, where the weights are a function of the spatial dependence of the extremes. Specifically, at time t and location s , let $f(\cdot; \boldsymbol{\mu}_{ts}, \boldsymbol{\sigma}_s, \boldsymbol{\xi}_s)$ be the density function of the GEV distribution with location, scale, and shape parameters $\boldsymbol{\mu}_{ts}$, $\boldsymbol{\sigma}_s$, and $\boldsymbol{\xi}_s$ respectively, and let $\mathbf{S}_{ts} = d\log f[\ln(\mathbf{Y}_{ts}); \boldsymbol{\mu}_{ts}, \boldsymbol{\sigma}_s, \boldsymbol{\xi}_s]/d\boldsymbol{\mu}_{ts}$. Then the CSE for β over a domain with m stations is

$$\sum_{t=1}^n \ln(\gamma_t) \mathbf{W}_t^{-1} \mathbf{S}_t = 0, \quad (5)$$

where $\boldsymbol{\gamma}_t = (\gamma_{t1}, \dots, \gamma_{tm})$, $\mathbf{W}_t^{-1}$ is $m \times m$ weight matrix, and $\mathbf{S}_t = (S_{t1}, \dots, S_{tm})^T$. Estimating equations of site-specific parameters $(\boldsymbol{\mu}_{bs}, \boldsymbol{\sigma}_s, \boldsymbol{\xi}_s)$ are based on the derivatives of the log likelihood with respect to these parameters, which only

depend on data at each site s . Given a set of initial values of $(\mu_{bs}, \sigma_s, \xi_s)$, β is estimated by solving the CSE; given β , $(\mu_{bs}, \sigma_s, \xi_s)$ is estimated by their estimating equations at each s ; the iteration continues until convergence.

In our application, we opt not to explicitly model spatial dependence and thus use the special case with the identity weight matrix, referred to as the CSE-I method, for estimating the scaling factors in model (2). While not explicitly accounted for in the GEV fitting procedure, spatial dependence will be accounted for in the confidence interval estimation procedure that will be described next. The CSE-I method has been shown to provide more robust results in the presence of uncertainty in the signal estimates and has more reliable confidence intervals (Wang et al. 2020). The CSE-I method is equivalent to the profile likelihood method of Zwiers et al. (2011), but computationally much faster. Additional details about the CSE-I method can be found in Wang et al. (2020).

3) BOOTSTRAP CONFIDENCE INTERVAL ESTIMATION

In traditional detection and attribution analyses, estimates of uncertainty due to natural internal variability are obtained from long control simulations. Further, it is assumed that the variability affecting the signal estimates has the same characteristics as the internal variability affecting the observations (e.g., Ribes et al. 2013). Internal variability estimates from climate models are very hard to verify, and there is concern that their use leads to considerable underestimation of scaling factor uncertainty (DelSole et al. 2019). We have therefore used an appropriately constructed block bootstrap procedure to infer confidence intervals for the scaling factor.

Scaling factor confidence intervals are constructed by means of a two-level block bootstrap procedure that accounts for both the effects of the signal uncertainty (level 1), which arises from variability simulated by the climate models, and the natural internal variability that affects the observations (level 2). The procedure, which is described in detail in the online supplemental material, is designed to account for sampling uncertainty, the effects of spatial dependence and the effects of temporal dependence on time scales up to 5 years. At the first level, resampling is performed to generate additional bootstrap samples of ensemble of simulations in a way that retains temporal signal structure, spatial dependence, and as much temporal dependence as possible. These samples are used to generate additional realizations of the estimates of that model's smoothed signal so as to account for signal uncertainty due to model-simulated internal variability. At the second level, a block bootstrap sampling procedure is applied to residuals of the observations after removal of the scaled signals and refitting each set of the bootstrapped signal estimates from the first level, which accounts for internal variability within the observations, as generated by the climate system itself. In total, we generate 32 bootstrap replicates of the signal estimates and 32 bootstrap replicates of the scaling factor estimates for each replicate of the signal estimates, resulting in $32 \times 32 = 1024$ estimates of β that are used to construct scaling factor confidence intervals. We use the 5th and 95th percentiles of these scaling factor estimates as the

bounds for approximate 90% confidence intervals and claim detection if this interval lies above zero.

We conduct one-signal (ALL) and two-signal analyses (ANT, NAT) for the changes of precipitation extremes in global, continental, and regional domains. Similar to the standard fingerprinting setting, a response to external forcing is said to be “detected” if the corresponding scaling factor is found to be significantly greater than zero. Further, if the corresponding confidence interval also covers unity, attribution is considered.

4) EXTREME PRECIPITATION CHANGES ATTRIBUTABLE TO EXTERNAL FORCING

After conducting the detection analysis, we use the fitted nonstationary GEV models to estimate possible changes in the likelihood of extreme precipitation attributable to external forcing estimating the waiting time (the inverse of the probability of occurrence) in the 2010–14 climate for Rx1day or Rx5day events that are estimated to have been 20-yr events in the 1950–54 climate. Waiting times are constant in a stationary climate—a value more extreme than the 20-yr return value is expected to recur once every 20 years on average. In contrast, waiting times have an instantaneous interpretation in a nonstationary climate that is specific to the date for which the waiting time is determined; they indicate the instantaneous rate at which the specified threshold is likely to be exceeded (Zwiers et al. 2011). The calculation of waiting times proceeds as follows: 1) we assume that the climate remained roughly stationary over the period 1950–54 and thus calculate a representative location parameter μ_{1950s} for the period that is equal to the average of location parameters for 1950–54 from Eq. (2); 2) we calculate 20-yr return values (i.e., 95th percentiles) of the distribution with location parameter μ_{1950s} and the corresponding scale and shape parameters for individual stations over the domain of interest; 3) we then set the location parameter to μ_{2010s} , taken as the average location parameter over 2010–14, and calculate the probabilities of exceeding the estimated 1950–54 20-yr return levels with the changed location parameters; and 4) the average exceedance probability across all stations in the domain is calculated and its inverse is considered as the waiting time (expected return period) in the 2010–14 climate for the domain for the recurrence of circa 1950–54 20-yr extreme events. If the intensity or frequency of extreme precipitation has increased over time, we would expect waiting times for these events to be reduced in the recent 2010–14 climate.

Also, we convert these attributable location parameter changes to attributable annual extreme precipitation changes using the following steps: 1) for each domain of interest, we compute the median value of Rx1day or Rx5day (X50) over 1950 to 2014 for individual stations in the domain using the quantile estimation function [Eq. (4) with location parameters estimated using Eq. (2) above]; 2) we then calculate the trend in annual X50 values for 1950–2014 at individual stations and express the trend as a percentage relative the median X50 value; 3) we calculate the median percentage change across the spatial domain to obtain a rough estimate of attributable

changes in extreme precipitation over the domain; and finally, 4) these attributed changes in intensity are expressed as sensitivities per degree of warming in global mean surface temperature, using the global mean surface temperature of the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) (Hansen et al. 2010). The observed warming in global mean surface air temperature in the NASA/GISS dataset is about 0.86°C over the 1950–2014 period. We estimate attributable changes in Rx1day and Rx5day over 1950–2014 for domains where anthropogenic influence is detected using both CanESM2 and CMIP6 signals.

3. Results

a. Observed and modeled trends

Figure 2 shows spatial patterns of trends for the log-transformed Rx1day and Rx5day from observations, the CanESM2 large ensemble, and the CMIP6 multimodel ensemble mean during 1950–2014. The observational results show that precipitation extremes have intensified over the majority of land areas, including large regions of Europe and the eastern part of North America, but with a mixed pattern of wetting and drying over eastern China, the west coast of South America, and Australia. The ALL forcings CMIP6 MME and CanESM2 large ensemble show positive trends for both Rx1day and Rx5day almost everywhere with broad agreement on the sign of changes, while the NAT simulations do not exhibit such trends. Observations show much greater than simulated trends over the NH mid- to high-latitude regions, especially in large regions over Europe, while the ensemble of model simulations shows weaker than observed trends. Results for Rx5day are very similar to those for Rx1day. In interpreting these figures, it should be noted that the observations reflect a combination of internal unforced variability and forced variability, while the ensemble means would be much less influenced by unforced internal variability since much of this variability is filtered away through the calculation of the ensemble means.

b. Detection and attribution analysis at global and continental scales

1) SINGLE-SIGNAL ANALYSIS

Figure 3 shows scaling factor best estimates for ALL based on single-signal analyses. The ALL scaling factor estimates are significantly greater than zero for both Rx1day and Rx5day in analyses over global land, indicating that the combined effect of external anthropogenic and natural forcing can be detected robustly in extreme precipitation regardless of whether the CanESM2 or CMIP6 MME signal estimate is used. The corresponding 90% confidence intervals include unity when using the CanESM2 signal, indicating that the model-simulated change signals in extremes under ALL forcing are consistent with changes in the observations at the global scale. Scaling factors for the CMIP6 signals are slightly smaller, with the CMIP6 Rx1day signal being marginally inconsistent with unity. Note that our bootstrap confidence intervals do not account for model uncertainty, which is

difficult to quantify with ensembles of opportunity, like the CMIP6 MME ensemble. Thus, the actual scaling factor uncertainty for CMIP6-based signals is likely larger than shown.

It should be noted that scaling factors can be affected by several factors, including differences in the magnitudes of trends between observations and model simulations, mismatch of spatial patterns of changes, and internal variability that obscures the response to external forcing. Observations show an intensification of precipitation extremes over Europe and eastern North America, but a mixed pattern of wetting and drying over eastern China, the west coast of South America, and Australia. In contrast, the CMIP6 MME and CanESM2 large ensemble ALL simulations show positive Rx1day and Rx5day trends almost everywhere (Fig. 2; see also Fig. S9). Overall, the medians of the linear trends for the GLB from CanESM2 and CMIP6 MME seem to be slightly larger than that from observations for Rx1day, which is consistent with scaling factors over global land that tend to be smaller than unity.

Large-scale circulation change that occurs in response to anthropogenic forcing could also be a factor. Changes in precipitation extremes projected by climate models under future forcing scenarios (e.g., Li et al. 2021) provide hints of what may also be happening under historical forcing. In particular, the signature of Hadley and Walker circulation changes seems evident, with projected reductions in the intensity of precipitation extremes in regions where the downwelling branches of these circulations are found [e.g., see Li et al. (2021), their Fig. 5]. Projected changes over Oceania appear to be quite uncertain at lower levels of forcing, with the possibility that tropical overturning circulation change may be damping changes in precipitation extremes in some parts of the region. Circulation change may also be a factor over the interior of large midlatitude continental regions such as North America (Li et al. 2021, their Fig. 5; Li et al. 2019a).

At the continental scale, the ALL signal scaling factors are also significantly greater than zero for Rx1day over AS, NA, and EU, indicating that influence of external forcing on extreme precipitation is also detectable on continental scales. The observed spatiotemporal evolution in extreme precipitation over EU is well represented by the ALL forcing climate change signals, especially when using the CanESM2 signal, for which the scaling factor is close to unity. In contrast, the ALL signal scaling factors are smaller than one over AS and NA land areas. A large uncertainty in these areas is whether models can adequately represent precipitation changes associated with changes in aerosols that affect precipitation producing process via changes in cloud microphysics (Rosenfeld et al. 2014). These areas are also strongly affected by low-frequency variability associated with ENSO and the interdecadal Pacific oscillation (e.g., King et al. 2013), which will have affected the single observational realization of change that is available for analysis. Not surprisingly, the ALL forcing signals are not detectable over OC. The observational study of Sun et al. (2021) shows that trends in extreme precipitation are strongly mixed in the region, with more stations showing negative trends than positive trends. This is an area where the projected responses to forcing are complex, with considerable intensification near the equator that is associated

FIG. 2. Spatial patterns of long-term trends [unit: $\% (65 \text{ yr})^{-1}$] in log-transformed Rx1day and Rx5day during 1950–2014 from (a),(b) observations, the CanESM2 large ensemble with (c),(d) ALL forcing and (e),(f) NAT forcing, and CMIP6 multimodel simulations with (g),(h) ALL forcing and (i),(j) NAT forcing. To make the spatial pattern of observation more comparable to that from model simulations, the log-transformed annual maxima from the stations are averaged within each $5.0^\circ \times 5.0^\circ$ grid box, and then linear trends are calculated for each grid cell. Annual extremes from CanESM2 and CMIP6 are regridded to the same $5.0^\circ \times 5.0^\circ$ resolution grid and then the linear trends were calculated. The numbers indicate the median of linear trends over grid cells over global land.

with strengthening of the ITCZ, but comparatively weak responses over Australia (Kharin et al. 2013; Li et al. 2021). This area is also strongly affected by the interdecadal Pacific oscillation, and changes in other modes of variability that affect Australia, such as the Antarctic Oscillation (Dey et al. 2019; Thompson et al. 2011). Results for Rx5day are similar to

those for Rx1day, except the scaling factors based on the CanESM2 ensemble are not significantly greater than zero over AS.

Results may be somewhat sensitive to the variant of the CSE method that we used to fit our extreme value models. Specifically, we used the CSE method with the identity weight

FIG. 3. Point estimates of the scaling factors for Rx1day (black) and Rx5day (orange) and the corresponding 90% confidence intervals for single-signal (ALL) detection analysis at the global and continental scales using CanESM2- and CMIP6-based signal estimates. Best estimates (data points) and 5%–95% uncertainty ranges (error bars) of scaling factors are displayed for ALL signals.

matrix, referred to as the CSE-I method, which gives equal weight to all stations. Regions that are more densely sampled, such as Europe, can therefore more strongly influence the estimation of the global scaling factors than other regions. We therefore repeated the analysis using a weight matrix that reflects the spatial variation in the density of stations as follows. The weight given to each station in a given $5.0^\circ \times 5.0^\circ$ grid box was set to $1/N$, where N is the number of stations in the grid box. Thus each $5.0^\circ \times 5.0^\circ$ grid box with data receives the same weight, with individual stations within densely sampled grid boxes having small weights relative to stations in less densely sampled grid boxes. Results (see Fig. S10) show that the scaling factors for GLB decrease slightly compared the case with the CSE-I method (Fig. 3), which is likely due to the larger weight now given to regions in which the number of stations is small and the change trends are not very obvious, such as South America, Africa, and Australia. Results for continental and smaller scales, however, do not change very much. We therefore present CSE-I method results below, noting also that it has been shown to provide more robust results in the presence of uncertainty in the signal estimates and has more reliable confidence intervals (Wang et al. 2020).

2) TWO-SIGNAL ANALYSIS

Figure 4 shows the best estimate scaling factors for ANT and NAT in two-signal analyses at the global and continent scales. Even when separately accounting for the naturally forced signal, anthropogenic influence can be detected at the 10% significance level in Rx1day over the global land area and most continents, including AS, EU, and NA, with ANT scaling factors that are significantly greater than zero when using either CanESM2 or CMIP6 signals. For EU, ANT scaling factors include unity, indicating that the observed changes

are largely consistent with anthropogenic influence. As in the one-signal analyses, anthropogenic influence is not detectable in OC. The CanESM2 NAT signal scaling factors are not significantly larger than zero in any domain, but lie above zero in the global land, NA, and OC domains when using CMIP6 signals. The finding that NAT signals may be detected in some regions is not robust since results vary with model simulations and detection methods. Also, it is likely that some NAT signal details have been lost with the smoothing process, such as signals from episodic volcanic forcing. Previous studies used the optimal fingerprinting method also do not find detectable NAT signals over these regions (Paik et al. 2020). The Rx5day results are very similar to those for Rx1day, with the anthropogenic influence being detected over the global land area and three continents (AS, EU, and NA); NAT signals are, however not detected when using signals from either CanESM2 or CMIP6 simulations (Fig. 4).

Our results for RX1day over Asia are consistent with those of Dong et al. (2020), who used the optimal fingerprinting method and CMIP5 models. In contrast with Dong et al., however, we were able to detect anthropogenic influence on Rx5day with our technique in both one- and two-signal analyses. Our global and continental results are also consistent with Paik et al. (2020), who conducted a two-signal detection analysis using CMIP6 models in an optimal detection and attribution method and detected ANT signals for the global, AS, EU, and NA domains (Paik et al. 2020; see also Fig. S1). Consistency between these studies and the study reported in this paper, using different datasets and data processing methods, different detection and attribution methods, and different models, substantially increases confidence in the finding that human influence on extreme precipitation is now robustly detectable in land-based precipitation observations at global and continental scales.

FIG. 4. Point estimates of the scaling factors for Rx1day and Rx5day and the corresponding 90% confidence intervals for two-signal (ANT and NAT) detection at the global and continental scales using CanESM2 and CMIP6 signal estimates. Best estimates (data points) and 5%–95% uncertainty ranges (error bars) of scaling factors are displayed for ANT (red) and NAT (blue) signals.

c. Detection and attribution analysis at regional scale

1) SINGLE-SIGNAL ANALYSIS

Figure 5 summarizes single-signal detection results over 16 of the 43 IPCC AR6 reference land regions that have spatial and temporal coverage satisfying our coverage requirements. The CanESM2 ALL forcings Rx1day signal is detected in 8 out of 16 regions, including eastern Europe (EEU), northern Europe (NEU), western and central Europe (WCE), central North America (CAN), eastern North America (ENA), eastern Siberia (ESB), western central Asia (WCA), and eastern central Asia (ECA). Scaling factor confidence intervals include unity over six of those regions (NEU, EEU, WCE, CNA, ENA, WCA, and ECA), indicating that the model response supports the attribution of the observed spatiotemporal evolution to the combination of forcings. The CMIP6 ALL forcings signal is similarly detected in six regions (EEU, NEU, WSB, CNA, ENA, and ECA) with scaling factor confidence intervals that include unity for all six regions. Observations show intensification of precipitation extremes over NEU, while model simulations tend to underestimate these trends over large parts of the region (Fig. 2), which is consistent with scaling factors that are slightly larger than unity over NEU. In contrast, scaling factor confidence intervals include negative values over regions such as SAU, WNA, RFE, and MED, indicating less consistency between observations and model simulations and larger internal variability (Fig. 2). Note that confidence intervals for larger domains are narrower than those for the smaller regional scale due to larger scaling factor sampling uncertainty in smaller domains, as has also been noticed by Kirchmeier-Young and Zhang

(2020). Consistent with Kirchmeier-Young and Zhang (2020; see their Fig. 4), we robustly detect anthropogenic influence on extreme precipitation in eastern and central North America, but not western North America. They obtained their results using an optimal fingering technique and three large ensemble simulations, which were performed with CanESM2, the Canadian Regional Climate Model CanRCM4, and the Community Earth System Model. This consistency, using different data processing, models, and analysis methods, strengthens confidence in our understanding of human influence on extreme precipitation in North America, and the regional variations in the strength of that signal relative to other causes of change.

Figure 5 shows that similar results were obtained for Rx5day, but with the influence of external forcings additionally detected over WSB, the Russian Arctic region (RAR), the Russian Far East (RFE), and the Mediterranean (MED) based on CanESM2 signal, while the CMIP6 ALL forcing Rx5day signal is additionally detected over RAR but not over WCE. The 90% confidence intervals for scaling factors for ALL forcing Rx5day signals from both climate models include unity over EEU, NEU, CNA, ENA, RAR, and WSB, indicating that the magnitudes of model simulated changes are comparable to those observations.

2) TWO-SIGNAL ANALYSIS

Figure 6 shows the best estimates of the scaling factors and their 90% bootstrap confidence intervals for two-signal detection of CanESM2 and CMIP6 signals in 16 regions. The ANT signal is robustly detected in Rx1day over the same six regions (EEU, NEU, WSB, CNA, ENA, and ECA) as in the one-signal analysis, regardless of whether CanESM2 or CMIP6

FIG. 5. Point estimates of the scaling factors for Rx1day (black) and Rx5day (orange) and the corresponding 90% confidence intervals for single-signal (ALL) detection analyses using CanESM2 and CMIP6 signals over 16 regions used in the IPCC WGI Sixth Assessment Report. Best estimates (data points) and 5%–95% uncertainty ranges (error bars) of scaling factors are displayed for ALL signals.

signals are used. The ANT signal scaling factors include unity in all six cases, supporting the attribution of the observed spatiotemporal evolution of Rx1day to anthropogenic influence. The CanESM2 and CMIP6 ANT signals are also detected in East Asia (EAS) but do not include unity. Evidently, both ensembles of climate model simulations overestimate observed changes in EAS. In addition, the CanESM2 ANT signal is also detected in ESB and WCE. Consistent with the continental-scale analyses, the NAT forcing signals are not robustly detected in any region.

Turning to Rx5day, we detect the ANT signal over more regions than was the case for Rx1day, with the CanESM2 ANT signal detected in five additional regions (ESB, MED, WCA, RFE, and RAR) and the CMIP6 ANT signal detected in one additional region (RAR) (Fig. 6). NAT scaling factors are again not significantly different from zero in almost all regions, except over ENA and WCA when using CanESM2 signals and WSB when using CMIP6 signals.

Overall, the regions in which we detect the ANT signal in the two-signal analyses are similar to those in which we detect ALL with the one-signal analysis. All findings point to

anthropogenic forcing as the likely cause of the widespread observed intensification of Rx1day and Rx5day and indicate that natural external forcing is unlikely to have contributed significantly. Detection becomes more difficult at the regional scale, and in regions where circulation change and the influence of aerosols on precipitation producing processes may have also affected changes in extreme precipitation.

d. Estimated extreme precipitation changes attributable to anthropogenic influence

Figure 7 shows the estimated waiting time (i.e., the inverse of the annual probability of occurrence) in the 2010–14 climate for events that are estimated to have been 20-yr events in the 1950–54 climate. These estimates are reported for domains where ANT signals are detected using both CanESM2 and CMIP6 signals. Globally, waiting times are estimated to have changed substantially, with the waiting time for 1950–54 1-day extreme 20-yr events reduced to 16 years (15–17 years) and 17 years (16–18 years) respectively in the 2010–14 climate when using CanESM2 and CMIP6 signals (Fig. 7a). The percent changes in Rx1day correspond well

FIG. 6. Point estimates of the scaling factors for CanESM2 and CMIP6 simulated ANT and NAT signals in Rx1day and Rx5day and their corresponding 90% confidence intervals for two-signal detection analyses over 16 IPCC AR6 reference land regions. Best estimates (data points) and 5%–95% uncertainty ranges (error bars) of scaling factors are displayed for ANT (red) and NAT (blue) signals.

with changes in waiting time over the global land. The median attributable increases in Rx1day are estimated to be 6.0% (4.0%–7.6%) per 65 years over the 1950–2014 period when using CanESM2 signals, which is consistent with the observed changes reported by Sun et al. (2021). This is equivalent to an attributable intensification per 1°C of global warming of about 7.0% (4.7%–8.8%) when expressed relative to the observed warming in global mean surface air temperature of ~0.86°C over this same period based on the GISS dataset (Hansen et al. 2010), which is also consistent with the sensitivity estimated by Sun et al. (2021). Estimates of the attributable increase in Rx1day intensity and the corresponding sensitivity to warming are slightly smaller when using CMIP6 signals.

Waiting times for circa 1950–54 20-yr Rx1day events have also decreased substantially at the continental scale, including AS, NA, and EU, and regional scale, including EAS, ECA, EEU, ENA, and NEU. The largest changes in the likelihood of extreme precipitation are observed in EU at the continental scale and NEU at the regional scale, which may be related to the substantial warming that has occurred in these regions as suggested in Sun et al. (2021). The estimated attributable

waiting time changes and intensity changes tend to be somewhat smaller in most domains when using CMIP6 signals rather than CanESM2 signals.

Waiting times for the recurrence of 1950–54 20-yr Rx5day events in the 2010–14 climate are also estimated to be substantially reduced for the globe, two continents (NA and EU), and several regions including EEU, ENA, NEU, RAR, and WSB, when using signals from either CanESM2 or CMIP6. Waiting times for such events are estimated to have decreased to fewer than 17 years for the globe (Fig. 7b), and the median values of attributable increases in Rx5day are estimated to be 5.8% (3.7%–7.6%) and 4.4% (2.2%–5.9%) per 65 years when using CanESM2 and CMIP6 signals respective (Fig. 7d). This corresponds to an attributable intensification per 1°C of global warming of about 6.7% (4.3%–8.8%) and 5.1% (2.6%–6.9%) when using CanESM2 and CMIP6 signals respectively. This somewhat lower sensitivity to warming than for Rx1day is consistent with observations findings (e.g., Sun et al. 2021) and the expectation that the sensitivity to warming would transition smoothly from the sensitivity of the annual mean precipitation (approximately 2%–3% per 1°C of

FIG. 7. (a),(b) Estimated waiting times for 1950–54 20-yr extreme events in the 2010–14 climate for two-signal (ANT and NAT) detection, (c),(d) the median values of attributable relative changes, and (e),(f) the median values of the sensitivity of precipitation extremes to temperature change [percentage change per 1-K increase in global mean surface temperature] in Rx1day and Rx5day for spatial domains where ANT signals are detected robustly. Black plots are for CanESM2; orange plots are for CMIP6.

warming; Boer 1993; Allen and Ingram 2002) to the Clausius–Clapeyron rate or perhaps higher for the annual extremes of daily or shorter accumulations (Min et al. 2011).

Regionally, the largest change in the likelihood of a 1950–54 20-yr Rx5day is observed in NEU, where waiting times for 20-yr events are estimated to decrease to about 13 years when using CanESM2 signals and the median attributable intensification in Rx5day intensity is about 9.5% per 65 years. Decreases in the waiting time and attributable changes are somewhat smaller when using CMIP6 signals, except over WSB. Overall, it can be concluded that external forcing, dominated by anthropogenic influence, has contributed to the observed increases in precipitation extremes at different spatial scales. Further intensification of these extremes is projected as the climate continues to warm (Kharin et al. 2013, 2018; Kirchmeier-Young and Zhang 2020; Li et al. 2021), with clear model-based evidence that intensification is essentially a monotonic function of warming on both global and regional scales (Li et al. 2021).

4. Conclusions and discussion

We have performed a detection and attribution analysis of the observed changes in Rx1day and Rx5day at the global, continental, and regional scales. Analysis was performed directly on meteorological station data without additional

data processing such as gridding, transformation to unitless probability indices, or spatiotemporal dimension reduction. We used records from 5081 land-based stations spanning the period 1950–2014 with at least 45 years of coverage, including at least 3 years in the period 2010–14. Changes at these stations were evaluated by fitting nonstationary generalized extreme value (GEV) distributions at individual stations to the logarithm of observed Rx1day and Rx5day values. Nonstationarity in the GEV distribution was permitted by using location parameters that were allowed to be linearly dependent on climate model-simulated responses in $\ln(\text{Rx1day})$ and $\ln(\text{Rx5day})$ to different forcings, with the restriction that for each signal, the climate model signal scaling coefficient (scaling factor) should be the same for all locations in the analysis domain. We analyzed changes at stations in domains ranging from global to continental to subcontinental regions. Climate model simulated responses to external forcing were estimated from large ensembles of ALL and NAT forcings simulations performed with CanESM2, and from a multimodel ensemble of CMIP6 ALL and NAT forcings simulations. As in conventional optimal fingerprinting, detection and attribution analysis was performed by making statistical inferences about the scaling factors to evaluate whether the climate model simulated signals were detectable in a given domain and, if so, whether a scaling factor different from unity was needed to

alter their amplitudes. Inferences were performed with a two-stage bootstrapping procedure that simultaneously accounts for signal uncertainty due to climate model simulated unforced internal variability and scaling factor uncertainty due to unforced natural variability in the observations. The bootstrapping procedure fully retains spatial dependence and retains temporal dependence up to approximately ENSO time scales. Unlike optimal fingerprinting analysis, difficult to verify climate model estimates of internal variability are not used.

The combined effect of anthropogenic and natural external forcing (ALL) on the intensification of extreme 1-day (Rx1day) and 5-day (Rx5day) precipitation accumulations is detected robustly at the global scale and at continental scales over western Eurasia, eastern Eurasia, and the western Northern Hemisphere. At the regional scale, we find that external influence can be detected robustly over central North America, East Asia, eastern central Asia, eastern Europe, eastern North America, northern Europe, and western Siberia for Rx1day, and over central North America, eastern Europe, eastern North America, northern Europe, the Russian Arctic region, and western Siberia for Rx5day. We obtained consistent results using simulated responses from different models (CanESM2 and CMIP6), although the scaling factors from CMIP6 mostly fall below those from CanESM2. We also used a two-signal variant of the analysis procedure to determine whether anthropogenic (ANT) and natural (NAT) external forcing signals could be detected when they included as separate signals. The ANT signal was detected at the global scale and over similar continents and regions for both Rx1day and Rx5day as in the analysis using just the ALL signal. In contrast, we found that the NAT signal was generally not detectable, and that it does not contribute significantly to observational trends over most regions.

Our results agree with earlier studies that applied the optimal fingerprinting method to heavily processed data using model simulations similar to those used in our study (Paik et al. 2020; Kirchmeier-Young and Zhang 2020). This consistency across different datasets, different data processing approaches (none in our case), different climate models, and different methods substantially enhances our confidence in the finding that human influence has contributed significantly to the observed changes in the annual extreme precipitation over the global land, the continents, and some regions. Consistent with Paik et al. (2020), scaling factors estimated from our study fall below unity, particularly at the global scale, while the 90% uncertainty ranges of scaling factors include unity in Zhang et al. (2013) and are larger than unity in Min et al. (2011). It is hard to judge whether scaling factors in our study are generally underestimated given that the available observations provide rather an incomplete view of global extreme precipitation changes, with large data voids over land areas and virtually no rain gauge data over oceans. Nevertheless, to the extent that results give this impression, there are some possibilities as to why that may be the case. One possibility is that the spatial and temporal smoothing of signal estimates has not been able to remove sufficient noise, which would lead to underestimation of scaling factors, particularly at the regional scale. Also, the somewhat higher equilibrium

and transient climate sensitivities of CMIP6 models might be a small contributing factor leading to scaling factors less than unity, given the observed warming of 0.86°C over the period 1950–2014 and the corresponding DAMIP historical ALL simulations warming (based on the models used in our study) of 1.21°C over the same period. Assuming that the mismatch between observed and simulated changes in extreme precipitation can be entirely explained by the difference in observed and simulated warming rates on the basis of the Clausius–Clapeyron relation, we would expect the scaling factor on the ANT signal to be approximately $\beta = 1.07^{1.21}/1.07^{0.86} = 1.07^{0.35} = 1.024$. This is unlikely to be a discernable departure from unity given scaling factor estimation uncertainty. It seems more likely to us that the plausible explanations for scaling factor estimates that lie below unity would include things like observational uncertainty due to uneven observational coverage, signal uncertainty, possible anthropogenically forced circulation changes, and regional variations in aerosol forcing and response.

Anthropogenic influence is estimated to have substantially decreased the waiting time for the recurrence of extreme events in regions where anthropogenic influence has been detected, and over the global land area, it appears to have intensified extreme 1-day accumulations at a rate of about 7% per 1°C of warming in global mean surface air temperature. Previous studies (Kharin et al. 2013, 2018; Kirchmeier-Young and Zhang 2020; Li et al. 2021) have suggested that the intensification of extreme precipitation and resulting decreases in the waiting time for extreme events will continue with additional warming in future climate. The application of these scaling factors from our study can be used to impose an observational constraint to extreme precipitation projections and thus contribute to the establishment of emergent constraints (e.g., Stott and Kettleborough 2002; Tokarska et al. 2020a,b). Since our analysis is performed directly on meteorological station data, projected changes in the logarithm of future extreme precipitation can be multiplied by these scaling factors to obtain constrained projections that are directly interpretable at stations. Evidence of anthropogenic influence in extreme precipitation at the regional scale has important implications for infrastructure design and climate change adaptation planning and implementation.

Finally, our results substantially increase confidence in detection and attribution findings for extreme precipitation, while some caveats do remain. A potential limitation of the assumption that only the GEV location parameter is affected by external forcing is that the attributed changes in extreme precipitation are identical for all quantiles. In contrast, some previous model-based studies (e.g., Kharin et al. 2018; Li et al. 2019a, 2021) suggest that the relative changes in probability could be larger for rarer, more extreme events. A second potential limitation is that because we have smoothed the model simulated externally forced signals at a 5-yr time scale, we may have lost some detail from the NAT signals, such as signals from episodic volcanic forcing, that may potentially aid in their detection. Third, our method is not “optimal” in the sense that an “optimal fingerprinting” method is considered to be optimal. Specifically, uncertainty in the scaling factor estimates can likely be further reduced by explicitly

accounting for spatial and temporal dependence through the use of an appropriately constructed weight matrix in the CSE fitting procedure that we use. In addition, we have used signals from two types of ensembles: large initial conditions ALL and NAT forcing ensembles from a single model (CanESM2) and multimodel ALL and NAT forcing ensembles of opportunity from CMIP6. Thus, our results depend on signal estimates extracted from structurally different ensembles. In the case of CMIP6, the choice of models to include was constrained by ensemble sizes for individual models (our bootstrapping scheme requires at least three simulations per model), and by a desire to have ALL and NAT ensembles of the same size from each model. This latter requirement further limited the number of models that we could use, but ensures that differences between our ALL and NAT signal estimates were not due to the use of different samples of models for the extraction of those two signals. Relaxing these conditions would have allowed larger ALL and NAT forcing multimodel ensembles, thus possibly reducing signal estimation uncertainty.

Acknowledgments. The study was supported by the Pan-Canadian Global Water Futures (GWF) research program. We acknowledge the modeling group for producing the large ensemble of CanESM2. We also 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 (<https://esgf-node.lnl.gov/projects/cmip6/>), and the multiple funding agencies who support CMIP6 and ESGF. We also thank three anonymous reviewers whose constructive and insightful comments helped us to improve this paper.

REFERENCES

- Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 228–232, <https://doi.org/10.1038/nature01092>.
- , and P. A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting, part I: Theory. *Climate Dyn.*, **21**, 477–491, <https://doi.org/10.1007/s00382-003-0313-9>.
- Arora, V. K., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, <https://doi.org/10.1029/2010GL046270>.
- Boer, G. J., 1993: Climate change and the regulation of the surface moisture and energy budgets. *Climate Dyn.*, **8**, 225–239, <https://doi.org/10.1007/BF00198617>.
- Canadian Commission on Building and Fire Codes, 2015: National Building Code of Canada: 2015. National Research Council of Canada, <https://doi.org/10.4224/40002005>.
- Christidis, N., P. A. Stott, and S. J. Brown, 2011: The role of human activity in the recent warming of extremely warm daytime temperatures. *J. Climate*, **24**, 1922–1930, <https://doi.org/10.1175/2011JCLI4150.1>.
- DelSole, T., L. Trenary, X. Yan, and M. K. Tippett, 2019: Confidence intervals in optimal fingerprinting. *Climate Dyn.*, **52**, 4111–4126, <https://doi.org/10.1007/s00382-018-4356-3>.
- Dey, R., S. C. Lewis, J. M. Arblaster, and N. J. Abram, 2019: A review of past and projected changes in Australia's rainfall. *Wiley Interdiscip. Rev.: Climate Change*, **10**, e577, <https://doi.org/10.1002/wcc.577>.
- Donat, M. G., and Coauthors, 2013: 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, <https://doi.org/10.1002/jgrd.50150>.
- , A. L. Lowry, L. V. Alexander, P. A. O'Gorman, and N. Maher, 2016: More extreme precipitation in the world's dry and wet regions. *Nat. Climate Change*, **6**, 508–513, <https://doi.org/10.1038/nclimate2941>.
- Dong, S., Y. Sun, and C. Li, 2020: Detection of human influence on precipitation extremes in Asia. *J. Climate*, **33**, 5293–5304, <https://doi.org/10.1175/JCLI-D-19-0371.1>.
- Fowler, H. J., and Coauthors, 2021: Anthropogenic intensification of short-duration rainfall extremes. *Nat. Rev. Earth Environ.*, **2**, 107–122, <https://doi.org/10.1038/s43017-020-00128-6>.
- Groisman, P. Ya., R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. N. Razuvayev, 2005: Trends in intense precipitation in the climate record. *J. Climate*, **18**, 1326–1350, <https://doi.org/10.1175/JCLI3339.1>.
- Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, **48**, RG4004, <https://doi.org/10.1029/2010RG000345>.
- Hershfield D. M., 1961: Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years. U.S. Weather Bureau Technical Paper 40, 65 pp., https://reduceflooding.com/wp-content/uploads/2018/09/TechnicalPaper_No40.pdf.
- Iturbide M., and Coauthors, 2020: An update of IPCC climate reference regions for subcontinental analysis of climate model data: Definition and aggregated datasets. *Earth Syst. Sci. Data*, **12**, 2959–2970, <https://doi.org/10.5194/essd-12-2959-2020>.
- 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, <https://doi.org/10.1007/s10584-013-0705-8>.
- , G. M. Flato, X. Zhang, N. P. Gillett, F. Zwiers, and K. J. Anderson, 2018: Risks from climate extremes change differently from 1.5°C to 2.0°C depending on rarity. *Earth's Future*, **6**, 704–715, <https://doi.org/10.1002/2018EF000813>.
- King, A. D., L. V. Alexander, and M. G. Donat, 2013: Asymmetry in the response of eastern Australia extreme rainfall to low-frequency Pacific variability. *Geophys. Res. Lett.*, **40**, 2271–2277, <https://doi.org/10.1002/grl.50427>.
- Kirchmeier-Young, M. C., and X. Zhang, 2020: Human influence has intensified extreme precipitation in North America. *Proc. Natl. Acad. Sci. USA*, **117**, 13 308–13 313, <https://doi.org/10.1073/pnas.1921628117>.
- Li, C., and Coauthors, 2019a: Larger increases in more extreme local precipitation events as climate warms. *Geophys. Res. Lett.*, **46**, 6885–6891, <https://doi.org/10.1029/2019GL082908>.
- , F. Zwiers, X. Zhang, and G. Li, 2019b: How much information is required to well constrain local estimates of future precipitation extremes? *Earth's Future*, **7**, 11–24, <https://doi.org/10.1029/2018EF001001>.
- , —, —, —, Y. Sun, and M. Wehner, 2021: Changes in annual extremes of daily temperature and precipitation in CMIP6 models. *J. Climate*, **34**, 3441–3460, <https://doi.org/10.1175/JCLI-D-19-1013.1>.
- Li, H., H. Chen, and H. Wang, 2017: Effects of anthropogenic activity emerging as intensified extreme precipitation over China. *J. Geophys. Res. Atmos.*, **122**, 6899–6914, <https://doi.org/10.1002/2016JD026251>.

- Min, S. K., X. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, **470**, 378–381, <https://doi.org/10.1038/nature09763>.
- Paik, S., S. K. Min, X. Zhang, M. G. Donat, A. D. King, and Q. Sun, 2020: Determining the anthropogenic greenhouse gas contribution to the observed intensification of extreme precipitation. *Geophys. Res. Lett.*, **46**, e2019GL086875, <https://doi.org/10.1029/2019GL086875>.
- Ribes, A., S. Planton, and L. Terray, 2013: Application of regularised optimal fingerprinting to attribution. Part I: Method, properties and idealised analysis. *Climate Dyn.*, **41**, 2817–2836, <https://doi.org/10.1007/s00382-013-1735-7>.
- Rosenfeld, D., and Coauthors, 2014: Global observations of aerosol–cloud–precipitation–climate interactions. *Rev. Geophys.*, **52**, 750–808, <https://doi.org/10.1002/2013RG000441>.
- Sarojini, B. B., P. A. Stott, and E. Black, 2016: Detection and attribution of human influence on regional precipitation. *Nat. Climate Change*, **6**, 669–675, <https://doi.org/10.1038/nclimate2976>.
- Seong, M. G., S. K. Min, Y. H. Kim, X. Zhang, and Y. Sun, 2020: Anthropogenic greenhouse gas and aerosol contributions to extreme temperature changes during 1951–2015. *J. Climate*, **34**, 857–870, <https://doi.org/10.1175/JCLI-D-19-1023.1>.
- Sobie, S. R., and T. Q. Murdock, 2014: Analysis of climate change projections for the Ministry of Transportation and Infrastructure highway risk assessment. *Pacific Climate Impacts Consortium*, University of Victoria, https://www2.gov.bc.ca/assets/gov/driving-and-transportation/environment/climate-action/moti_climate_report_v6.pdf?forcedownload=true.
- Stott, P. A., and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature*, **416**, 723–726, <https://doi.org/10.1038/416723a>.
- , N. P. Gillett, G. C. Hegerl, D. J. Karoly, D. A. Stone, X. Zhang, and F. Zwiers, 2010: Detection and attribution of climate change: A regional perspective. *Wiley Interdiscip. Rev.: Climate Change*, **1**, 192–211, <https://doi.org/10.1002/wcc.34>.
- Sun, Q., F. W. Zwiers, X. Zhang, and L. Li, 2020: A comparison of intra-annual and long-term trend scaling of extreme precipitation with temperature in a large-ensemble regional climate simulation. *J. Climate*, **33**, 9233–9245, <https://doi.org/10.1175/JCLI-D-19-0920.1>.
- , X. Zhang, F. W. Zwiers, S. Westra, and L. V. Alexander, 2021: A global, continental and regional analysis of changes in extreme precipitation. *J. Climate*, **34**, 243–258, <https://doi.org/10.1175/JCLI-D-19-0892.1>.
- Thompson, D. W., S. Solomon, P. J. Kushner, M. H. England, K. M. Grise, and D. J. Karoly, 2011: Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nat. Geosci.*, **4**, 741–749, <https://doi.org/10.1038/ngeo1296>.
- Tokarska, K. B., G. C. Hegerl, A. P. Schurer, P. M. Forster, and K. Marvel, 2020a: Observational constraints on the effective climate sensitivity from the historical period. *Environ. Res. Lett.*, **15**, 034043, <https://doi.org/10.1088/1748-9326/ab738f>.
- , M. B. Stolpe, S. Sippel, E. M. Fischer, C. J. Smith, F. Lehner, and R. Knutti, 2020b: Past warming trend constrains future warming in CMIP6 models. *Sci. Adv.*, **6**, eaaz9549, <https://doi.org/10.1126/sciadv.aaz9549>.
- Wang, Z., Y. Jiang, H. Wan, J. Yan, and X. Zhang, 2017: Detection and attribution of changes in extreme temperatures at regional level. *J. Climate*, **30**, 7035–7047, <https://doi.org/10.1175/JCLI-D-15-0835.1>.
- , —, —, —, and —, 2020: Toward optimal fingerprinting in detection and attribution of changes in climate extremes. *J. Amer. Stat. Assoc.*, **116** (533), 1–13, <https://doi.org/10.1080/01621459.2020.1730852>.
- Wen, Q. H., X. Zhang, Y. Xu, and B. Wang, 2013: Detecting human influence on extreme temperatures in China. *Geophys. Res. Lett.*, **40**, 1171–1176, <https://doi.org/10.1002/grl.50285>.
- Westra, S., L. V. Alexander, and F. W. Zwiers, 2013: Global increasing trends in annual maximum daily precipitation. *J. Climate*, **26**, 3904–3918, <https://doi.org/10.1175/JCLI-D-12-00502.1>.
- Yin, H., Y. Sun, H. Wan, X. Zhang, and C. Liu, 2016: Detection of anthropogenic influence on the intensity of extreme temperatures in China. *Int. J. Climatol.*, **37**, 1229–1237, <https://doi.org/10.1002/joc.4771>.
- Zhang, X., L. Alexander, G. C. Hegerl, P. Jones, A. K. 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, <https://doi.org/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, <https://doi.org/10.1002/grl.51010>.
- Zwiers, F. W., X. Zhang, and Y. Feng, 2011: Anthropogenic influence on long return period daily temperature extremes at regional scales. *J. Climate*, **24**, 881–892, <https://doi.org/10.1175/2010JCLI3908.1>.