Climate Central Temperature Attribution System Methods

Version 1, June 2022

Andrew Pershing, Daniel M. Gilford, and Joseph Giguere

Background

The purpose of the Climate Central Temperature Attribution System (CC-TAS) is to provide high resolution estimates of how human-caused climate change has altered the likelihood of observed or forecasted air temperatures. The system works by taking the temperature *T(t, x)* observed or forecasted for day=*t* and location=*x* and comparing it to statistical distributions for *T(x)* representing the modern observed climate and a counterfactual climate without anthropogenic greenhouse gas emissions. The ratio of the likelihoods of *T* in these two climates gives an estimate of how human-caused climate change has influenced *T*.

Gilford et al. (2022, hereafter, G22) developed a multi-method framework for creating the modern[1] and counterfactual climates. They describe two approaches that are grounded in observational data and another approach that uses climate models. These three methods provide complementary assessments of how climate change influenced the likelihood of an event, and each includes a rigorous characterization of the uncertainties in its likelihood estimate.

The G22 empirical approaches begin by creating a time series of monthly average or median temperatures at a given location. For each month m, they use linear regression to compute

$$T_m(y, x) = \beta(m, x)\ \mathrm{GMT}(y) + c \tag{1}$$

where GMT(y) is the global mean temperature in year=y and $T_m(y, x)$ is the expected temperature for month m (Figure 1, step 1). The coefficient $\beta(m, x)$ is called the "scale factor", and it is the key output of this process. The scale factor describes how we expect the temperature at x to change in response to a change in GMT. It is the key characteristic calculated from empirical data; it enables the transformation of *T* likelihoods between a well-known observed climate and the forced+counterfactual climates defining the attribution estimates.

Median Scaling: The first method in G22 is a straightforward application of the scale factors. They characterize the distribution of daily temperatures over a reference period (they used 1985-2015) for each month m as a quantile distribution (i.e. $T(q)$ = the temperature such that q = the proportion of observed temperatures less than or equal to $T(q)$) (Figure 1, step 2). They then create the modern climate using the scale factor:

$$T_{\mathrm{modern}}(q) = T(q) + \beta(\mathrm{GMT}_{\mathrm{modern}} - \mathrm{GMT}_{\mathrm{reference}}) \tag{2}$$

[1] We use the term "modern" to describe the climate which we are comparing to the counterfactual. The exact time period depends on the context. For G22, modern = 2010-2019. For our operational system, modern = 2022

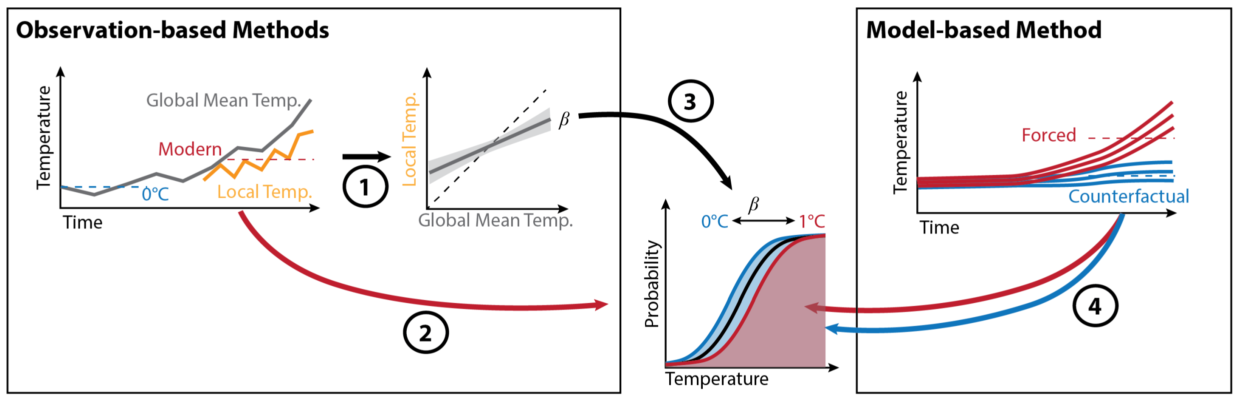

Figure 1. Attribution method described in G22. The two observation-based methods begin by (1) relating the local temperature (orange) to GMST (gray) to get β, the change in local mean temperature with a change in GMST. This includes an error estimate from the linear regression. (2) Observational data are used to characterize a climatological (1985– 2015) distribution of temperatures (black curve). (3) Then, the median-scaling method uses β to shift the climatological distribution backward to a pre-industrial counterfactual climate (blue curve and shading) and either backward or (typically) forward to a forced distribution of temperatures contemporary with the events being attributed (red curve and shading); shifts based on GMST are assumed to be completely driven by historical human-emitted greenhouse gases. The quantile-scaling method uses the same procedure, but models separate β values across thirty specified distribution quantiles. The model-based method uses climate model projections to characterize the local temperature under natural forcing (blue lines) or in a climate forced by human-emitted greenhouse gases (red lines). (4) The forced and counterfactual distributions of temperatures used to quantify attribution estimates are then inferred directly from the three methods.

where $\mathrm{GMT}_{\mathrm{modern}}$ is the global mean temperature averaged over the modern period (they used the ten-year period around 2015, i.e., 2010-2019), and $\mathrm{GMT}_{\mathrm{reference}}$ is the average over the reference period (Figure 1, step 3). The counterfactual climate is created in an analogous way:

$$T_{\mathrm{counter}}(q) = T(q) + \beta(\mathrm{GMT}_{\mathrm{counter}} - \mathrm{GMT}_{\mathrm{reference}}) \quad (3)$$

where $\mathrm{GMT}_{\mathrm{counter}}$ is the global mean temperature averaged over the counterfactual period (they used 1885-1915). If $\beta>0$ as is typically found in a human-warmed climate, then $T_{\mathrm{modern}}(q) > T(q) > T_{\mathrm{counter}}(q)$.

For an observed T, they can estimate q in both the modern and counterfactual climates (q_c(I), where c=modern or counter). They then report a “probability ratio” expressed by the cumulative probabilities of T in each climate:

$$PR(T) = q_{\mathrm{modern}}(T)/q_{\mathrm{counter}}(T) \quad (4)$$

Quantile Scaling: The G22 median scaling method changes the center of the temperature distribution. Their quantile scaling method allows for the shape of the distribution to change as well. They do this by computing $\beta_q(m, x)$:

$$T_{m,q}(y, x) = \beta_q(m, x)\ \mathrm{GMT}(y) + c \quad (5)$$

where q is a particular quantile. They used 30 evenly spaced quantiles between 0.01 and 0.99. They then apply equations (2) and (3) using each β_q to create the modern and counterfactual distributions.

Daily temperatures from Berkeley Earth gridded at 1.25° formed the basis of their empirical analysis; G22 GMT estimates were drawn from the Met Office Hadley Centre/Climatic Research Unit Temperature data set, version 5 (HadCRUT5).

Model-based Method: G22's model based method used an ensemble of CMIP5 simulations. Daily Berkeley Earth data were used to bias-adjust 24 historical+RCP8.5 model simulations using the method of Lange (2019). Eleven of these models also had pre-industrial control runs, and the same bias adjustment was applied to these simulations. After bias adjustment, they identified the year when the model's running global mean temperature reached GMT_{modern} (they used GMT_{modern} =1.07° as the attributable warming since the preindustrial period). They then used the 31 years centered on this year to represent the modern climate and estimated $q(T)$ for each of the 24 historical+RCP8.5 simulations. The counterfactual distribution was pooled from of all eleven available pre-industrial control simulations. They then computed probability ratios by dividing the 24 forced simulations by the pooled counterfactual simulation.

These methods form the basis of our system. When operationalizing the G22 methods, we identified several challenges including discrepancies between model-derived forecast data and the interpolated Berkeley data. This prompted is to identify new and improved ways to optimize the computations and to improve their accuracy. We identified six specific issues that prompted improvements to our system:

1. The coarse Berkeley data could not resolve complex coastlines and treated many important coastal cities as if they were in the ocean.
    - ➡ **Solution**: higher spatial resolution
2. The temperature distributions in many areas, especially in the tropics, are very narrow. This means that any bias between the forecast and the climatology is magnified at these locations. Because GFS and Berkeley are fundamentally different estimates of temperature, we did not trust our calculations in these regions.
    - ➡ **Solution**: use the CFSR, which uses the same data assimilation procedures as GFS, to establish observed temperature climatologies.
3. CMIP5 has been superseded by CMIP6.
    - ➡ **Solution**: switch to CMIP6
4. Even after debiasing, the climate model output can still contain differences in the distribution of temperatures among the different models. Comparing the forced output to the single pooled control can amplify these differences.
    - ➡ **Solution**: used paired models and only compare each forced simulation to the control run made with the same model.
5. Using quantiles to represent temperatures and *PR* as the main attribution metric is designed to consider very warm events. We want our system to be able to speak to less extreme conditions and also cold events.
    - ➡ **Solution**: use parameterized distributions (skew-normal) and base the attribution metric on the ratio of the probability distribution functions (PDFs), instead of cumulative probability distribution functions (CDFs).

CC-TAS Version 1.0, Spring 2022

The system builds on G22. Most importantly, it uses their approach of computing scale factors and then using them to transform an observed climatology into the modern and counterfactual climatologies used to compute an attribution metric. It also uses a similar strategy of directly estimating changes in the distributions based on climate models. The major differences are

- Berkeley Earth has been replaced with high resolution CFS reanalysis temperatures to define the climatology
- ERA5 reanalysis is used to compute the scale factors
- Paired forced and pre-industrial CMIP6 models
- The quantile-based method (CDF values) for assigning the likelihood of the observed temperatures has been replaced with a method that fits a skew-normal distribution to the observations or models and then uses that distribution to estimate likelihoods.
- The occurrence ratio (*OR*) is used as the attribution metric instead of the probability ratio (*PR*)

Details and rationale are described below.

Climatologies

We use the National Center for Environmental Prediction's climate forecast system reanalysis (CFSR) as the basis for our historical climatologies. CFSR is available starting in 1979 and running to present at 0.5°-by-0.5° resolution. The data are produced using the same data, data assimilation, and model core as the GFS forecasts.

CFSR consists of two sequential products:

- CFSR (v1) January 1979 to December 2010
- CFSv2 January 2011 to present

The data are publicly available at 6hr resolution. To get daily maximum (Tmax) and minimum (Tmin), we used the xarray groupby('time.day').max() (or min) function. We downloaded the entire record from 1979-2021 and processed to Tmin and Tmax.

We define the reference climatology for day=d as the CFSR data for the 31 day period centered on day d over the years 1991-2020 (i.e. the NOAA climate normal period).

Observations

For our nowcast and forecasts operations, we get temperature data from NOAA's Global Forecast System (GFS). GFS is an operational model that is run several times a day to support weather forecasts. We download the forecast for the next 5 days and then compute each days min and max temperatures. GFS is run at a 0.25°-by-0.25° resolution. We use xesmf.Regridder (method = "conservative") to transfer the GFS data up to the coarser CFSR grid.

We can also apply the attribution methods to the CFSR record. We use these data for hindcasting past conditions.

Scale Factors

Scale factors (the β s in the equations above) are central to the two empirical attribution methods. These are computed by regressing a temperature time series for a particular quantile against global mean temperature. We used the Hadley Centre's HadCRUT temperature product as our global mean temperature.

We originally tried computing scale factors using the 1979-2020 CFSR database. However, we found that this was not sufficient to get reliable scale factors. In particular, if a region was very hot or very cool for the last several years (for example, because of the PDO phase), then the scale factors would be stronger or weaker than we would expect based on climate projections and our experience with the Berkeley data.

To correct this, we used the ECMWF ERA5 reanalysis, which extends back to 1950. We downloaded the data at 3 hourly resolution and then computed daily min and max temperature. For a day of the year d, we extracted the 31 days surrounding that data over the entire record and regridded from the ERA5 0.25° grid to the CFS 0.5° grid (esmf "conservative" method). We then computed the quantiles in each year to get a time series for each quantile level. For each quantile level, we computed $\beta_q(d, x)$. We used 24 periods in the year: the 1st and 15th calendar day of each month. During testing we found that we could gain computational efficiency by reducing the number of quantiles without noticeable changes in the output. We now use 21 evenly-spaced quantiles between 0.02 and 0.98, including the median (q=0.5). The R^2 statistic and the p-value of scale factor regressions were retained.

Skew Normal Distributions

The most significant change between V1 and the methods described in G22 is the use of parameterized distribution functions. We found that both Tmax and Tmin are well-approximated by the skew-normal distribution. This function is described by the location parameter (analogous to a median) = L, the scale parameter (analogous to the variance) =S, and the shape parameter=A that defines the level of skewness.

The SciPy stats package includes a function to fit each built-in distribution to data. We used this function to fit skew-normal distributions to the reference climatologies. The counterfactual distributions for median scaling are constructed from the reference distributions by adding $\beta(GMT_{cf} - GMT_{ref})$ to the location parameter. The modern distributions for median scaling are created in the same way.

The quantile scaling method poses a challenge to using parameterized distributions like the skew normal. Because this method transforms the approximation of the CDF represented by the quantile distributions, it impossible to use the built-in routines. We use SciPy's optimize.leastsq to find the [A, L, S] at each location that minimizes the residuals between the observed quantile distribution and the skew-normal CDF using [A, L, S]. We use the least-squares fit to the CDF for the modern and counterfactual distributions.

We also fit skew-normal distributions to the data from the climate models (see below).

Estimating the Modern Climate

The Gilford et al. methodology uses annual global mean temperature (GMT) to define the global climate state. Ideally, we would like to define GMT(y) as the mean over some interval p of years (for example, 11) surrounding y. However, this definition creates a problem for the last p/2 years in the record. This is especially troubling for our main application of estimating the attributable state of the climate for the current year.

One approach we considered was to define GMT as the mean of the prior p years. This is highly conservative in that it will always be cooler than the value we would get using the centered mean. For p = 11, it would be about 0.2°C below the “true” GMT. We decided to define GMT(y) as

- the centered 11 year mean, if $y < 2022-6 = 2016$
- γ_{30} y + c_{30}, if $y >= 2016$

where γ_{30} and c_{30} are the coefficients from the linear regression of GMT against year over the 30 preceding years (e.g. 1992-2021 for y=2022). Note that the values estimated using the 30 year regression are strongly related to those using the 11 year mean ($R^2=0.97$, $p<0.01$).

Model-based Method

We accessed CMIP6 output (daily tasmax and tasmin) from the Google Cloud archive. We found 24 models that had historical runs, forced projections (SSP3-7.0 if available, SSP5-8.5 in

Table 1. List of climate models from CMIP6

Organization	**Model Names**
Australian Community Climate and Earth System Simulator	ACCESS-CM2, ACCESS-ESM1-5
Alfred Wegener Institute	AWI-CM-1-1-MR
Euro-Mediterranean Center on Climate Change	CMCC-ESM2
Centre National de Recherches Météorologiques	CNRM-CM6-1-HR, CNRM-CM6-1, CNRM-ESM2-1
Canadian Centre for Climate Modelling and Analysis	CanESM5
EC Earth Consortium	EC-Earth3-AerChem, EC-Earth3-Veg-LR, EC-Earth3-Veg, EC-Earth3
Geophysical Fluid Dynamics Laboratory	GFDL-CM4, GFDL-ESM4
Institute for Numerical Mathematics, Russian Academy of Sciences	NM-CM4-8, INM-CM5-0
Institut Pierre Simon Laplace	IPSL-CM6A-LR
JAMSTEC, AORI, NIES, R-CCS	MIROC6
Max Planck Institute	MPI-ESM-1-2-HAM,MPI-ESM1-2-HR, MPI, ESM1-2-LR, MRI-ESM2-0
NorESM Climate Modelling Consortium	NorESM2-MM
Research Center for Environmental Changes, Academia Sinica	TaiESM1

some cases) and pre-industrial control runs (Table 1). We concatenated the historical and projections for each model to create a single “forced” simulation for each model.

We used xesmf’s regridder to regrid each model to a common 1.5°-by-1.5° grid. For the 16 models that had coarser resolution, we used bi-linear interpolation (xesmf’s “bi-linear” method). For the 8 models that had finer resolution, we used the xesmf “conservative” method.

As in G22 we used the Lange (2019) methodology to bias-adjust the output of each individual climate model. We used the same 1991-2020 CFSR climatology (regridded to 1.5°) as in our empirical method as the reference data for the debiasing. The debiasing trained by the relationship between the reference climate and the forced simulations was likewise applied to debias each paired pre-industrial control simulation.

For each model, we identified the first year when its representation of GMT (smoothed using an 11year running mean) was greater than or equal to the modern GMT as defined above. We then select the 31 year period centered around that year to assess the climate corresponding to that GMT. For each of the 24 periods in the year, as with the empirical method, we extract the 31 day period centered around the target day. We then fit skew normal distributions to the Tmin and Tmax data. We also extract the same range of days from the last 31 years of the pre-industrial control run and fit a counterfactual skew normal distribution.

Attribution Estimates

We now have several skew normal distributions representing the modern and counterfactual climates (Table 2). The pair of distributions created by each method allow us to independently estimate the change in the likelihood of the temperature in question.

Let $T(d, y)$ be the observed or forecasted temperature for day d in year y. We find d_j = the day from the 24 periods that is closest to d. For each method, we acquire $SN_{counter}(d_j)$ and $SN_{modern}(d_j)$, the files with the parameters for the skew normal distributions for day d_j.

Table 2. Summary of the attribution methods. The first row shows the global mean temperature relative to 1850-1899 for the three periods: reference, counterfactual, and modern. The remaining rows summarize the process for creating the skewed normal distributions that are used to estimate the likelihoods.

	Reference (1991-2020)	**Counterfactual (1885-1915)**	**Modern (y=2022)**
GMT (rel. 1850-1899)	0.90	-0.06	1.27
Median-scaling	Direct fit to 1991-2020 CFSR data	Reference distribution with location parameter shifted by β *(GMT_{cf}- GMT_{ref})	Reference distribution with location parameter shifted by β *(GMT_{mod}- GMT_{ref})
Quantile-scaling	Fit to 1991-2020 quantile distribution	Fit to reference quantile distribution shifted by β_q *(GMT_{cf}- GMT_{ref})	Fit to reference quantile distribution shifted by β_q *(GMT_{mod}- GMT_{ref})
Models	N/A	Fit to model M’s control run	Fit to model M’s forced run

Once we have the skew normal distributions, we use them to get PDF_{modern} (T) and $PDF_{counter}$ (T), the PDF values from each climate and with each method. We then compute:

$$OR(T) = \mathrm{PDF}_{modern}(T) / \mathrm{PDF}_{counter}(T)$$

which we refer to as the "occurrence ratio."

The occurrence ratio is our main attribution metric. It approximates the ratio of the probability of encountering a temperature close to T in the two climates. Values of OR>1 indicate that anthropogenic climate change has made those conditions more likely, and values <1 indicate that climate change made those conditions less likely.

Rationale for the Occurrence Ratio

Using the occurrence ratio as the primary attribution metric is a departure from established attribution methodologies. The probability ratio (sometimes called the hazard ratio) is used in most studies. It is defined as (cf. eqn. 4):

$$PR(T) = (1-\mathrm{CDF}_{modern}(T)) / (1-\mathrm{CDF}_{counter}(T))$$

The *PR* makes a statement about the likelihood of temperatures greater than or equal to *T*. It is well-suited to the needs of most attribution studies which focus on extreme conditions.

A key difference between our application and most published studies is that we want to be able to make a statement for any *T*, not just extreme values. We also want to be able to discuss conditions like cold days that are becoming less likely. The *OR* provides a more intuitive way to do this as it refers to specific conditions and doesn't require the *a priori* choice of direction of change.

Combining Estimates and Accounting for Uncertainty

Our procedure provides two empirical estimates and 24 estimates from the paired climate models. G22 took the mean of the climate model estimates (24 models) and then took the mean of the model average and the two empirical methods. We use this same approach; however, additional testing may indicate a different weighting scheme.

G22 used a Monte Carlo procedure to estimate confidence intervals around the attribution metrics from the two empirical methods. A full implementation of their Monte Carlo approach would require us to fit multiple skew normal distributions for each time period. This is computationally challenging. We therefore focused on developing a conservative methodology that could be applied within our operational time limits. This includes the definition of our new attribution metric, the Climate Shift Index.

Our procedure is:

1. Convert each of the 26 estimates of OR into a climate factor (CF):

$$CF(T) = 2*\log_2(OR(T))$$

This creates a linear scale centered on 0. Note that *CF* values above 8 or below -8 were replaced with 8 or -8, respectively.

2. Average the 22 model-based climate factors together to create a single model average. As in G22, we found that the model-based estimates are generally the most conservative estimates (CF values closest to 0).

3. On rare occasions when the median or quantile method could not make an estimate (something observed occasionally for very extreme temperatures), replace that method with the model-based method

4. If there is any disagreement in the sign of the CF among the median, quantile, and model-average, then the average CF is set to 0

5. If all three agree on the sign, then average the empirical methods together. Finally, take the average of the two averages.

This gives the value of the climate factor (which can be converted back to OR if desired) at the 0.5° CFSR grid points. For the initial Climate Central mapping tool, we averaged the climate factor into NOAA climate divisions.

We then converted the average climate factor into the categorical Climate Shift Index. For warm temperatures, the Climate Shift Index is

CSI	Descriptor	OR range	CF range
0	No effect	1/5-1.5	-1.17 - 1.17
1	Moderate	1.5-2	1.17-2
2	Strong	2-3	2-3.17
3	Very Strong	3-4	3.17-4
4	Extreme	4-5	4-4.64
5	Exceptional	>5	4.64

The CSI values for cold conditions are defined analogously:

CSI	OR range	CF range
0	1/5-1.5	-1.17 - 1.17
-1	1/2 - 1/1.5	-2 - -1.17
-2	1/3-1/2	-3.17 - -2
-3	1/4-1/3	-4 - -3.17
-4	1/5-1/4	-4.64 - -4
-5	<1/5	-4.64

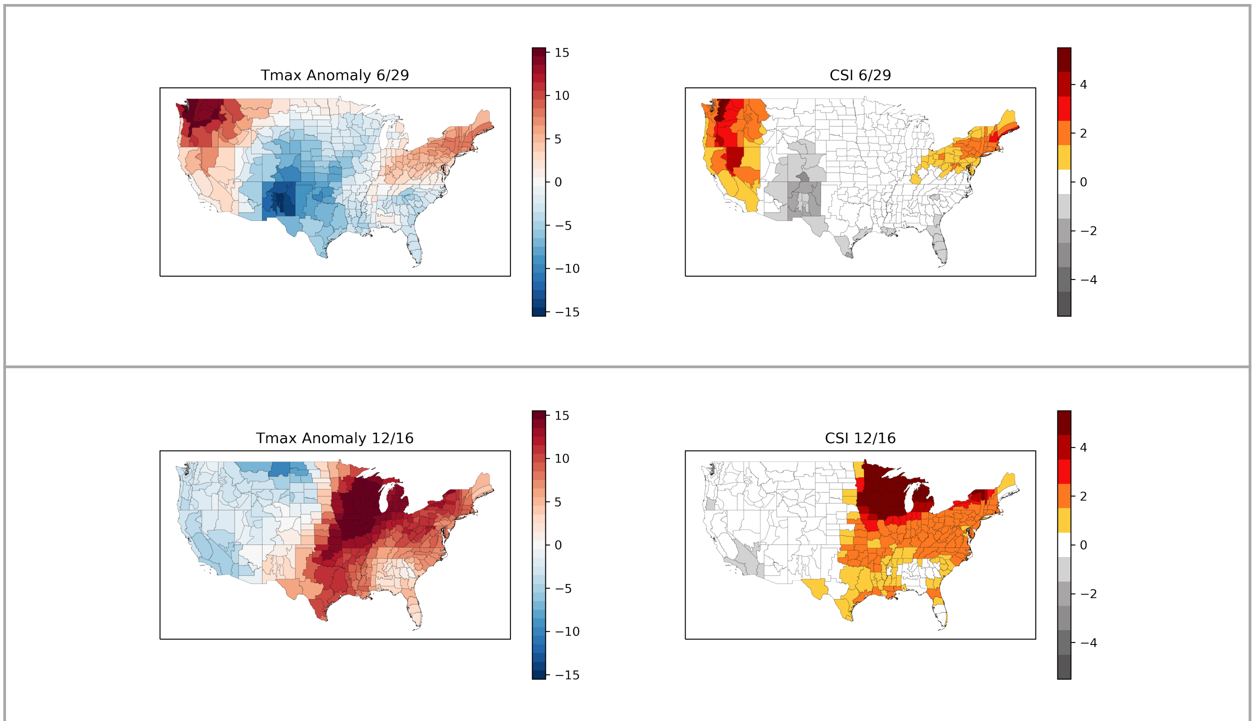

Figure 2. CC-TAS output for June 29, 2021 (top row) and December 16, 2021 (bottom row). Left. CFS Tmax expressed as an anomaly off of the 1991-2020 normal (anomaly is in °C). Right. The Climate Shift Index value for that day.

The CSI scale embeds several strategies meant to make the estimates conservative. First, it creates a very wide range of values around 0. This, plus the criterion that the three methods must agree on the sign, reduces the number of false positives. Second, the categorical nature of the scale means that we are effectively rounding toward zero. An event with an OR of 2.8 (CF = 2.97) becomes a 2 on the CSI scale. Finally, the CSI caps the maximum value at 5. This limits extrapolation into the tails of the distributions. The conservative approach and focus on lower bounds is consistent with other attribution studies that found a higher degree of certainty around lower bounds (Risser and Wehrner, 2017)

Example Results

Our initial public work will focus on the contiguous U. S. Our main product will be daily maps of the climate factor averaged over the NOAA climate divisions (Figure 2). We show the results for June 29, 2021, during the Pacific Northwest heat dome, and December 16, during an unusual winter heat event.

The Pacific Northwest heat dome is clearly identifiable and is quantified in our system with Climate Shift Index levels above 2 throughout Oregon and Washington and a band of level 4 and 5 in central Washington. Philip et al. (2021) estimated that the PR for this event was 175, a value much larger than what is indicated by the CSI. While our system is more conservative (by design), it clearly identified this as an important event and would have been able to do so days in advance. Our system also provides attribution estimates across the U.S.. For example, the

CSI identifies the cool temperatures in New Mexico and Colorado as conditions that are less likely due to climate change. It also identifies an event in the Northeast with a significant climate fingerprint (CSI>1).

In many ways, the December 16 event was even more striking than the Pacific Northwest event. A large region that encompassed all of Minnesota and Wisconsin and parts of Nebraska, South Dakota, North Dakota, Iowa, and Illinois reached level 5 on the CSI.

Acknowledgments

We are wish to express our gratitude to K. Haustein, F. Otto, C. Tebaldi, and M. Wehner who generously provided feedback on an earlier draft of this document. We also want to acknowledge the technical team at Climate Central, especially Z. Bleki, D. Dodson, S. Katsnelson, and M. Vieuille who worked closely with us to translate our ideas into an elegant operational system.

References

Gilford, D. M, Pershing, A. J., Strauss, B. H., Haustein, K., and Otto, F. E. L. (2022) A multi-method framework for global real-time climate attribution. Advances in Statistical Climatology, Meteorology and Oceanography, in review (April 17, 2022)

Lange, S. (2019) Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geoscientific Model Development, 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019.

Mark D. Risser and Michael F. Wehner (2017) Attributable human-induced changes in the likelihood and magnitude of the observed extreme precipitation in the Houston, Texas region during Hurricane Harvey. Geophysical Review Letters. 44, 12,457–12,464. https://doi.org/10.1002/2017GL075888