Methods

We document here the general approach used by WWA for the analysis of temperature extremes, which also applies for the study of the Siberian 2020 heat event. After having agreed on a definition of the event to be analysed, the analysis steps include: (i) trend calculation from observations; (ii) model validation; (iii) multi-method multi-model attribution and (iv) synthesis of observational and model results to assist the formulation of an attribution statement. This approach has also been described in peer-reviewed literature, for example, in Van Oldenborgh et al. 2018 (heat extremes), Van der Wiel et al. 2017 (precipitation extremes). We refer the reader to these papers for more background information. We aim to include as many models and model-setups as available, provided they pass the model validation. Contributing groups are permitted to use statistical methods that differ from the main WWA method. These are also briefly described below.

Statistical methods for the analysis of temperature extremes

Trend calculation from observations (Observed probability and probability ratio)

We first analyze observed temperatures and estimate how rare the current event is, and whether or not there is a trend in temperature. To do so, we analyse time series from stations or gridded observational data sets where long records of observed data are available.

The method differs depending on the type of extreme considered, which ranges from annual maximum/minimum values --- often three-day maximum temperatures, to seasonal means. For an event definition of a short-duration event where we choose the highest or lowest value in a season or year, we fit a Generalized Extreme Value Distribution (GEV) that theoretically describes these so called block maxima. It is described by three parameters: the position parameter μ, the scale parameter σ and the shape parameter ξ. The shape parameter is constrained with a Gaussian of width σ=X/2 added to the likelihood function, where X takes the default value of 0.4 but can be changed if required. For an event definition of, for instance, summer mean temperature, we fit either a Gaussian distribution, which sometimes describes aggregate quantities well or a Generalized Pareto Distribution (GPD), which only describes the tail of the distribution. We specify the threshold parameter μ and fit the scale parameter σ and the shape parameter ξ. In this statistical approach, global warming is factored in by allowing the fit to the distribution to be a function of the (low-pass filtered) GMST. We usually assume that the scale parameter σ and shape parameter ξ are constant, thus the PDF is shifted up or down with low-pass filtered GMST but does not change shape. In this way, it results in a distribution that varies continuously with GMST. This distribution can be evaluated for a GMST in the past (e.g., 1900) and for the current GMST. A 1000-member non-parametric bootstrap procedure is used to estimate confidence intervals for the fit.

We can then assess the probability of occurrence of the observed event in the present climate, p_1, and past climate, p_0. These probabilities are communicated as return periods of the event in the present and past: $1/p_1$ and $1/p_0$ respectively. The probability ratio PR is evaluated as the ratio of p_1 to p_0 and indicates how much more likely (PR>1) or unlikely (PR<1) the event has become. An alternative measure of the change is the change in intensity, ΔT in degrees (per degree increase in GMST).

Model validation

Model validation is the process of determining which of the available model ensembles perform well enough against observations to be used specifically for the attribution of the event as specified by the event definition. If an insufficient number of model ensembles perform well enough (less than two), the extreme event will not be analysed further and a formal attribution statement cannot be made, with any conclusions drawn being based on observations only.

Models are validated on general properties and statistical properties. General property assessments include checking if the seasonal cycle (timing of peaks and troughs) and the climatological spatial pattern resemble those in observations, and if the model resolution is sufficient to represent the spatial scale of the event. The statistical property assessment concerns the similarity of the probability distribution from climate model simulations to observations. The same statistical model (GEV, GPD or Gauss) is fitted to the simulations as to observations. It is then checked whether the best estimate for each of the fit parameters from simulations lies within the 95% confidence interval of the fit parameters from observations. It is also checked (at least by eye) that the fits applied agree with the model values, and that there is no notable temporal trend in variability. If these conditions are met, the validation result is “good”. If these conditions are only partly met, for example, if the 95% confidence interval of the fit parameters from observations and simulations only just overlap, or if the peak in the seasonal cycle is shifted by one month, the validation result is considered “reasonable”. If, per framing/model setup we have five or more models that are “good” according to the validation, we do not use the “reasonable” models for that framing/model setup but only the “good” models.

Attribution

To assess the role of climate change, we compare observations with results from climate models that are available and suitable for the variable under study in these locations. This answers the question whether and to what extent external drivers, in particular human-caused climate change, can explain the temperature trends in the observational data. Including models allows us to give the causation of a trend. We use the approximate return time of the event as found in the observations to determine the event threshold for analysis, or, in case of too extreme a return period, a lower return period, e.g., the lower bound of the distribution’s 95% confidence interval. For transient simulations of the changing climate, we again calculate how the probability of the event is changing over time in the model data, by fitting the temperature values to a distribution that shifts proportional to the smoothed global mean (model) temperature. For other models separate runs for the current climate and a counterfactual climate without anthropogenic emissions are available, the probability ratio and change in intensity are then simply computed from these runs, either parametrically by fitting an extreme value function or, if there are enough data, non-parametrically.

Method of Météo-France (MF)

The Météo-France method uses the same statistical models, but differs in two ways. MF assumes first that the covariate (here the GMST) is also random, and its confidence interval is taken into account. Secondly, the distribution of observations is fitted in a Bayesian way with a synthesis of a set of CMIP5 models as prior. The uncertainty of the covariate is inferred by the uncertainty of the decomposition of the GMST as the sum of an EBM model (natural forcings), a smoothing spline of time (anthropic forcings) and a residual Gaussian random term (natural forcings). This decomposition is applied to each model of a set of CMIP5 models, and used to fit the distribution of the variable of interest (GEV for maxima, etc). The confidence intervals are inferred by resampling 1000 times each model. Then, either the return time has already been specified and is used as a definition to calculate the PR and the ΔI for each CMIP5 model separately, or, if Météo France provides an observational analysis, the CMIP5 multi-model distribution is used as a prior to calculate the PR and the ΔI from the observations of the GMST and the variable of interest with Bayesian methods (a combination of the Gaussian conditioning theorem and the Markov Chain Monte Carlo approach, see Robin et al, 2020, in review). See also Ribes et al. (2020) and Ribes et al. (2017) for further information.

Method for CMIP6 data (ETH Zurich)

The CMIP6 data is analysed using the same statistical models as the main method. However, the parameter uncertainty is estimated in a Bayesian setting using a Markov Chain Monte Carlo (MCMC) sampler and not via bootstrapping. Using an affine-transformation invariant MCMC sampler (Goodman and Weare, 2010; Foreman-Mackey et al., 2013) we generate at least 25’000 non-independent samples. From these we calculate the best estimate of the parameters (median) and their uncertainty (using the 2.5 % and 97.5 % percentile). We use noninformative priors, except for the shape parameter of the GEV. ξ is constrained with a Gaussian prior with a standard deviation of 0.15.

Method for SMILES

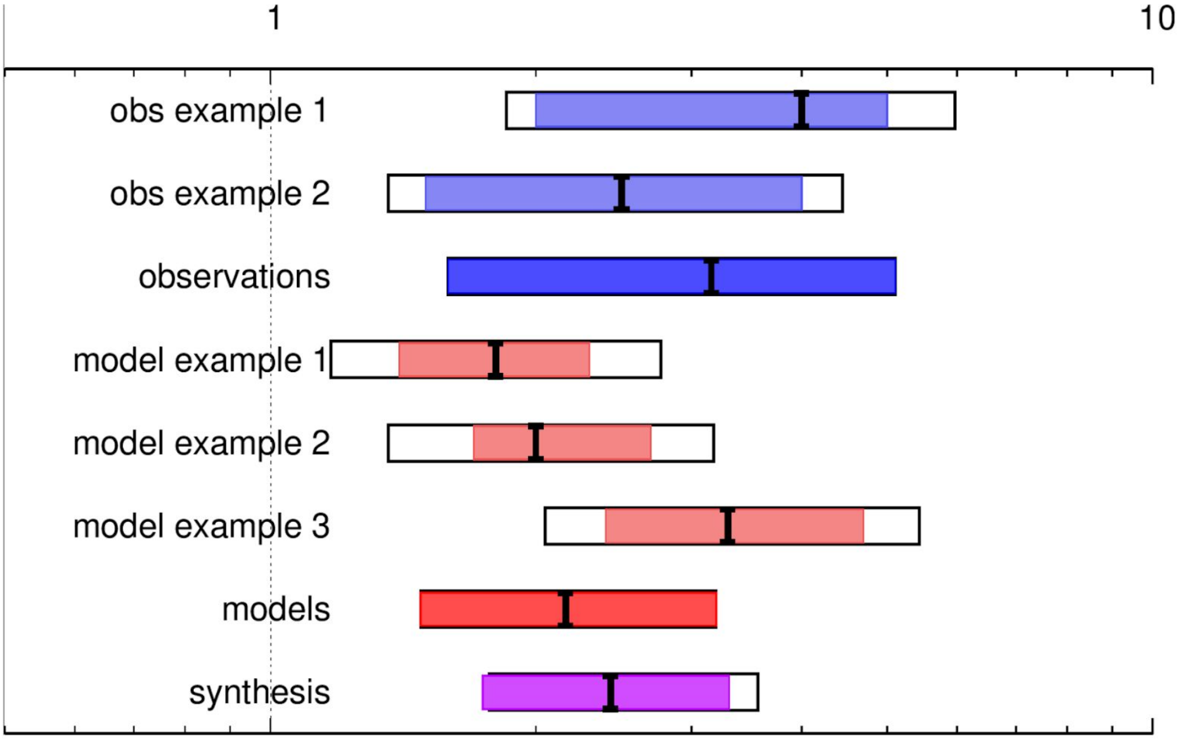
As a minor variant to the main method, distributions can be fitted to ensemble members of a single model initial-condition large ensemble (SMILE) at a given point in time (rather than fitting distributions across time). For a subset of CMIP5 and CMIP6 models, such SMILEs are available with ensemble members ranging from 16 to 100, depending on the climate model. This approach has the benefit that the changing background climate is taken into account automatically, such that GMST does not need to be invoked as a covariate. It also does not assume constant distributional parameters, as those change automatically as a result of the changing background climate (e.g., a forced future decrease in temperature variability associated with the loss of cryosphere at high latitudes).

All SMILE data is interpolated conservatively to a 2.5°x2.5° grid in the beginning. The SMILEs have a common start year of 1950, therefore probabilities and probability ratios (PR) are calculated for year 1950 (using pooled data from 1950-1954), 2020 (2018-2022) and 2050 (2048-2052). They are thus not directly comparable to the other results in this study, which focus on probability ratio relative to 1900 instead of 1950. Consequently, one might expect the PR of the SMILEs to be slightly smaller, as they do not include the climate change (i.e., GMST increase) between 1900 and 1950. In practice, though, the PR values are very similar given the confidence interval of the PR values.

Synthesis

We synthesise the results from observations and climate models that have passed our evaluation in order to assess the overall magnitude of change in likelihood of the event occurring, and whether the observed change is attributable to anthropogenic climate change. For each region or definition of the event(s) studied, the synthesis is presented in the form of a bar diagram (see figure below). In the synthesis figure the blue bars show the confidence intervals of the Risk Ratio results from observations and the red bars that from models. If there are multiple observational estimates, we make the simple assumption that the uncertainty due to natural variability is totally correlated between them, as they record the same reality. Differences between observational estimates are then assumed due to `representation uncertainty', and this independent uncertainty is added to each observational estimate and shown in the figure as a white extension to each blue bar. The dark blue bar below represents the combined observational result - a direct average of the extended boxes. For the models, the natural variability is assumed to be uncorrelated, that is that each model creates independent atmospheric realisations. The model uncertainty is comprised of the uncertainty due to natural variability and the model spread. The model spread can be estimated by comparing the spread of the results to the natural variability from the fits. It is depicted by the white boxes and is added to the model spread in quadrature to the natural variability, but only if the χ^2/dof statistics is greater than one. The bright red bar in these figures indicates the total uncertainty of the models. The total model uncertainty consists of a weighted mean using the (uncorrelated) uncertainties due to natural variability plus an independent common model spread added to the uncertainty in the weighted mean if appropriate. Finally, observations and models are combined into a single result in two ways if they seem to be compatible. Firstly, we neglect model uncertainties beyond the model spread and compute the weighted average of models and observations: this is indicated by the magenta bar. As model uncertainty can be larger than the model spread, secondly, we also show the more conservative estimate of an unweighted average of observations and models, indicated by the white box around the magenta bar in the synthesis figures.

Synthesis of example synthesis, using common interval 1900-2018, with reference values -0.1617 0.9417, representation uncertainty (2σ) factor 1.39, model chi2/dof = 2.407, model uncertainty (2σ) factor 1.42, Coloured bars denote natural variability, boxes also take representativity / model errors into account. In the synthesis, the purple bar denotes the weighted average of observations and models, the box denotes the unweighted average. (eps, pdf, plot script, raw data)



References

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). emcee: The MCMC hammer. Publications of the Astronomical Society of the Pacific, 125, 306–312, https://doi.org/10.1086/670067.

Goodman, J., & Weare, J. (2010). Ensemble samplers with affine invariance. Communications in Applied Mathematics and Computational Science, 5(1), 65–80, https://doi.org/10.2140/camcos.2010.5.65.

Ribes, A., Zwiers, F. W., Azaïs, J.-M., and Naveau, P. (2017): A new statistical approach to climate change detection and attribution, Clim. Dyn., 48, 367–386, https://doi.org/10.1007/s00382-016-3079-6.

Ribes, A., Thao, S., and Cattiaux, J. (2020): Describing the relationship between a weather event and climate change: a new statistical approach, Journal of Climate, 33, 6297–6314, https://hal.archives-ouvertes.fr/hal-02122780.

Robin, Y. and Ribes A. (2020): Non-stationary GEV analysis for event attribution combining climate models and observations, Advances in Statistical Climatology, Meteorology and Oceanography (in review).

Van der Wiel, K., Kapnick, S. B., van Oldenborgh, G. J., Whan, K., Philip, S. Y., Vecchi, G. A., Singh, R. K., Arrighi, J. and Cullen, H. (2017): Rapid attribution of the August 2016 flood-inducing extreme precipitation in south Louisiana to climate change, Hydrol. Earth Syst. Sci., 21, 897–921, https://doi.org/10.5194/hess-21-897-2017.

Van Oldenborgh, G. J., Philip, S. Y., Kew, S. F., van Weele, M., Uhe, P., Otto, F. E. L., Singh, R., Pai, I., Cullen, H. and AchutaRao, K. (2018): Extreme heat in India and anthropogenic climate change, Natural Hazards and Earth System Sciences, 18, 365–381, https://dx.doi.org/10.5194/nhess-18-365-2018.