

ISI-CFACT

Introduction

ISI-CFACT produces counterfactual climate data from past datasets for the ISIMIP project.

Counterfactual climate is a hypothetical climate in a world without climate change.

For impact models, such climate should stay as close as possible to the observed past, as we aim to compare impact events of the past (for which we have data) to the events in the counterfactual. The difference between past impacts and counterfactual impacts is a proxy for the impacts caused by climate change. We run the following steps:

- 1.** We approximate the change in past climate through a model with three parts. Long-term trend, an ever-repeating yearly cycle, and a trend in the yearly cycle. Trends are induced by global mean temperature change. We use a Bayesian approach to estimate all parameters of the model and their dependencies at once, here implemented through pymc3. Yearly cycle and trend in yearly cycles are approximated through a finite number of modes, which are periodic in the year. The parameter distributions tell us which part of changes in the variables can be explained through global mean temperature as a direct driver.

- 2.** We do quantile mapping to map each value from the observed dataset to a value that we expect it would have been without the climate-induced trend. Our hierarchical model approach provides us with a time evolution of our distribution through the time evolution of a gmt-dependent parameter. We first use this time-evolving distribution to map each value to its quantile in this time evolving distribution. We then use the distribution from a reference period in the beginning of our dataset where we assume that climate change did not play a role, to remap the quantile to a value of the variable. This value is our counterfactual value. Quantile mapping is different for each day of the year because our model is sensitive to the yearly cycle and the trend in the yearly cycle. This approach is illustrated Figure 1.

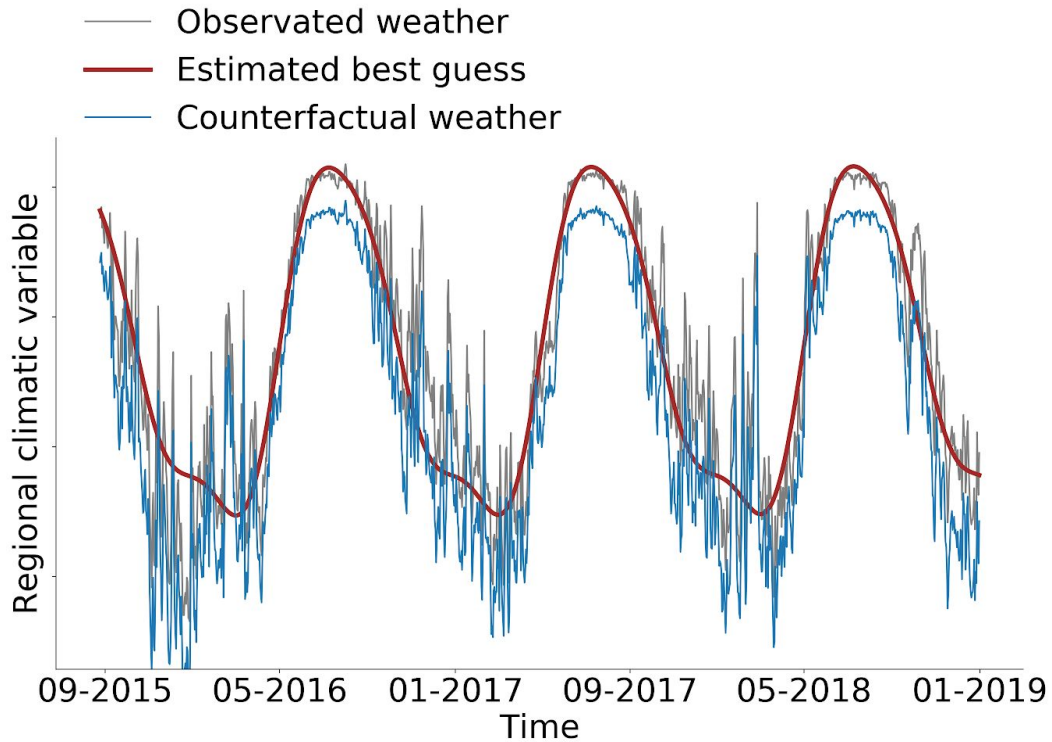


Figure 1: Model for a climate variable with yearly cycle. Grey is the original data, red is our estimation of change. Blue is the original data minus the parts that were estimated driven by global mean temperature change.

Variables

To avoid large relative errors in the daily temperature range as pointed out by Piani et al. (2010), we de-trend the daily temperature range $tasrange = tasmax - tasmin$ and the skewness of the daily temperature $tasskew = (tas - tasmin) / tasrange$ and derive $tasmin$ and $tasmax$ from tas , $tasrange$ and $tasskew$.

A counterfactual huss is derived from the counterfactual tas , ps and $hurs$ using the equations of Buck (1981) as described in Weedon et al. (2010).

Variable	Short name	Unit	Statistical model
Near-Surface Air Temperature	tas	K	Gaussian
Range of Daily Temperature	tasrange	K	Gaussian
Skewness of Daily Temperature	tasskew	1	Gaussian
Daily Minimum Near-Surface Air Temperature	tasmin	K	Derived from tas, tasrange and tasskew
Daily Maximum Near-Surface Air Temperature	tasmax	K	Derived from tas, tasrange and tasskew
Surface Downwelling Longwave Radiation	rlds	W / m ²	Gaussian
Surface Downwelling Shortwave Radiation	rsds	W / m ²	Gaussian
Surface Air Pressure	ps	Pa	Gaussian
Near-Surface Wind Speed	sfcWind	m / s	Weibull
Precipitation	pr	kg / m ² s	Bernoulli-Gamma
Near-Surface Relative Humidity	hurs	%	Gaussian
Near-Surface Specific Humidity	huss	kg / kg	Derived from hurs, ps and tas

Table 1: Specs of climate variables for the ISIMIP3b counterfactual climate datasets. The variables tasrange and tasskew are auxiliary variables to calculate tasmin and tasmax.

Model

A global mean temperature timeseries without yearly variations is used as predictor for the model. To generate this predictor, global mean temperature is preprocessed using singular spectrum analysis.

The variables *tas*, *rlds*, and *ps* are modeled with a Gaussian distribution with a time varying mean value. The mean value is a linear function of the global mean temperature change plus a yearly cycle. This yearly cycle is modeled with one mode for all variables except *tasskew*, where two modes are used. The parameters of the yearly cycle are also a linear function of the global mean temperature.

Tasrange and *tasskew* are modeled with a Gaussian distribution as described above. But those variables are bounded. *Tasrange* is positive and *tasskew* between 0 and 1. In the quantile mapping step, values that are close to the boundary can get mapped to values outside the defined range. To avoid this, such values are not quantile mapped with the effect that the counterfactual value is the same as the historic value in those cases. This happens only rarely, as the value has to be already close to the boundary which is unlikely for both variables.

The variables *hurs* and *rsds* are also bounded, *hurs* is between 0 and 1 and *rsds* is always non-negative. Those variables are also modeled with a Gaussian distribution. Values that are outside the defined range after quantile mapping are reset to the closest boundary value.

The *sfcWind* variable is modeled with a Weibull distribution using two parameters.

The shape parameter `_alpha_` is assumed to be free of trend.

Both parameters need to be positive.

Therefore, the model output is transformed with the logistic function to produce positive outputs.

Precipitation is modeled with a mixed Bernoulli-gamma distribution.

This approach enables to model the probability of rain-days and precipitation amounts on rain days with one distribution.

The Bernoulli-gamma distribution has three parameters. All three parameters are modeled as a linear function of the global mean temperature.

The model does not contain a yearly cycle for any parameter.

Results

We here present summaries of each variable for both the original data based on GSWP3 and the GSWP3-W5E5, and the corresponding counterfactual. ISI-CFACT removes annual trends as well as trends in the yearly cycle. Hereby the trend is regarded with the global mean temperature as independent variable. As a visual check, we show a map of the slope of a linear trend with time as independent variable, relative to the standard deviation of the slope in all grid-cells (see Figure 2 and Figure 3). By construction, our method should reduce this relative slope to a value close to zero. The trend calculated for this visual check is based on a simpler method that disregards the yearly-cycle of the variables.

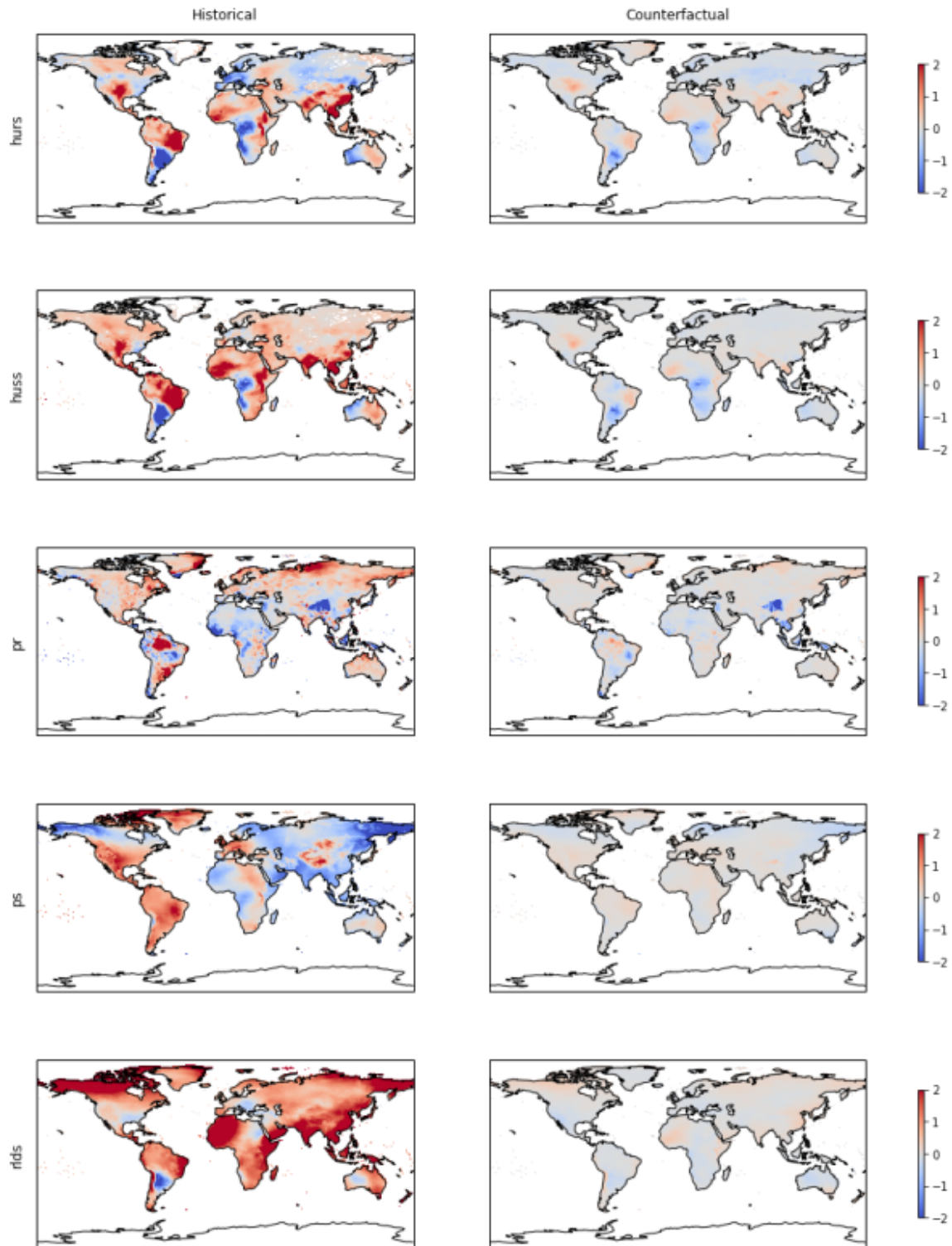


Figure 2: Maps of linear trends in historical (left) and counterfactual (right) climate for the GSWP3 dataset. The linear trends for hurs, huss, pr, ps, rlds are calculated with time as independent variable and without a yearly cycle. The grid-cell show the slope of that trend relative to the standard deviation of the slope in all grid-cells

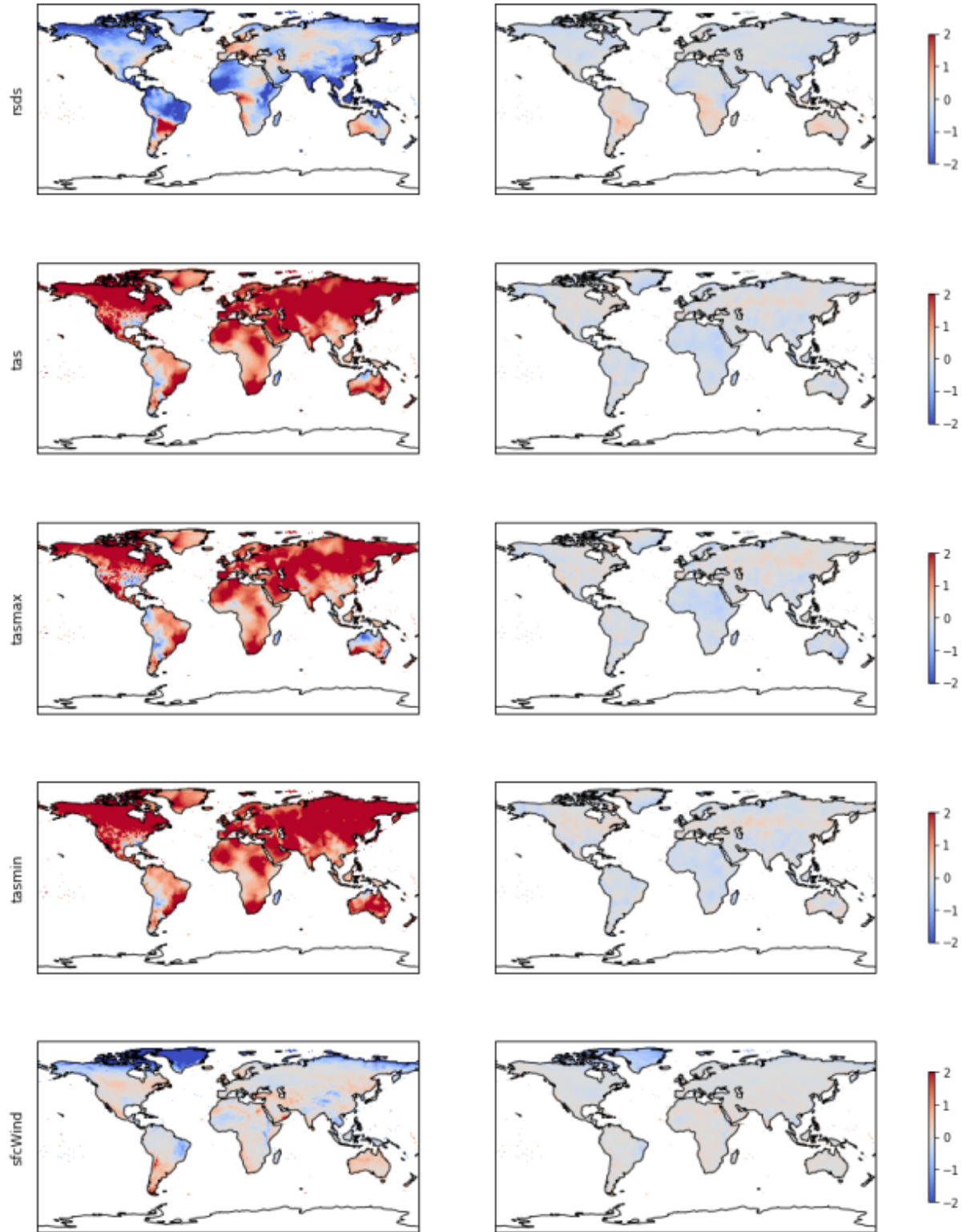


Figure 3: Same as Figure 2 for *rsds*, *tas*, *tasmin*, *tasmax* and *sfcWind*.

References

- Buck, A.L.: New Equations for Computing Vapor Pressure and Enhancement Factor, *J. Appl. Meteorol.*, 20, 1527–1532, 1981.
- Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, *J. Hydrol.*, 395, 199–215, <https://doi.org/10.1016/j.jhydrol.2010.10.024>, 2010.
- Weedon, G. P., Gomes, S., Viterbo, P., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: The WATCH forcing data 1958–2001: A meteorological forcing dataset for land surface and hydrological models, in: Technical Report no 22., available at: <http://www.eu-watch.org/publications/technical-reports> (last access: July 2016), 2010.