Quick data description

August 1, 2020

This small directory is intended to provide a very simple example of our Kriging method, based on real data. It does not provide all data and material to generate all figures shown in the corresponding paper, but do allow the user to generate Figure 2b (ie constraint of historical + SSP2-4.5 scenario).

We provide the following files / directory:

- `Dataset_GSAT.Rdata`: the data file, in R format. This files contains 3 variables (arrays), which are detailed below.
- `main.R`: a minimal R-script to run our method (on the above mentioned, real-world data). This script can be run in R using `source("main.R")`.
- `routines`: a few R-routines, called by `main.R`.

The following data / variables are provided:

- `Z`, a `ny`×`Nmod` array: the GSAT simulated by individual models in concatenated historical + SSP2-4.5 scenarios (period: 1850–2100, i.e. `ny`=251 years). If several simulations (members) are available, the ensemble mean is given. We provide data for `Nmod`=22 CMIP6 models (consistent with our paper). Unit: Kelvin.

- `Z_fit`, a `ny`×3×`Nmod` array: an estimate of the forced response. The forced response is estimated to 3 subsets of forcings: ALL (i.e., all historical forcings), NAT (i.e., natural forcings only), GHG (i.e., GHG-only), following the methodology described in the paper.

- `Y`, a `nyo`×`ns` array: historical observations. Historical observations (let's say "best reconstruction") of GSAT (more precisely: anomalies wrt the 1961–1990 average), over the period 1850–2019. There are `nyo`=170 years of data. Then, in order to sample measurement uncertainty, an ensemble of `ns`=100 realizations is provided. The median of this ensemble is often taken as a best-estimate (here denoted "`median`"). Note that the median is provided in addition to the 100 realizations, leading to `ns`=101.

- `Z_pictl`, a `nyp`×`Nmod` array: pre-industrial control simulations.
  Times-series (`nyp`=1000 years at most) of pre-industrial control simulations. These simulations can be used to evaluate internal variability (i.e. how observations might differ from the forced response `Z_fit`). These time-series are expected to be stationary (although low-frequency variability can be found for some models); in practice, however, some models are not at equilibrium, and might exhibit a trend.

Any technical issue or question about this toy example, please contact me:
`aurelien.ribes@meteo.fr`.