A GCM weighting procedure based on explicit probabilistic models and accounting for observation uncertainty

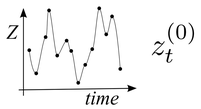
Benjamin Renard & Jean-PhilippeVidal

Irstea, UR HHLY Hydrology-Hydraulics, Centre de Lyon-Villeurbanne, France.

benjamin.renard@irstea.fr, jean-philippe.vidal@irstea.fr

Ingredients

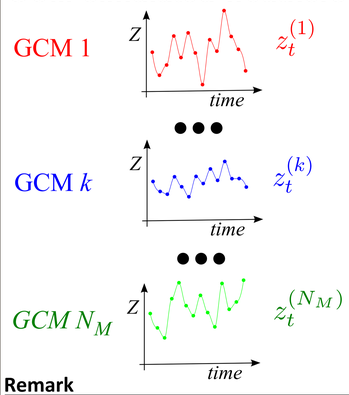
Observed weather



Remark

z_t may be multivariate (e.g. seasonal temperature-precipitation, or a spatial field)

GCM-simulated weathers



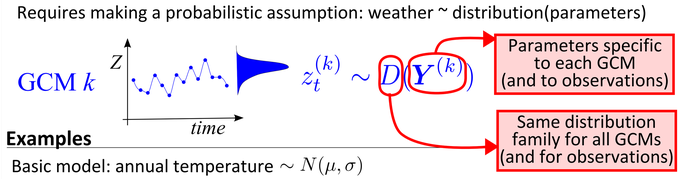
Remark

Replications welcome!

Definition of climates

Climate = *distribution* of weather

Requires making a probabilistic assumption: weather ~ distribution(parameters)



Examples

Basic model: annual temperature $\sim N(\mu, \sigma)$

Trend model: annual temperature $\sim N(\mu_0 + \mu_1 t, \sigma)$

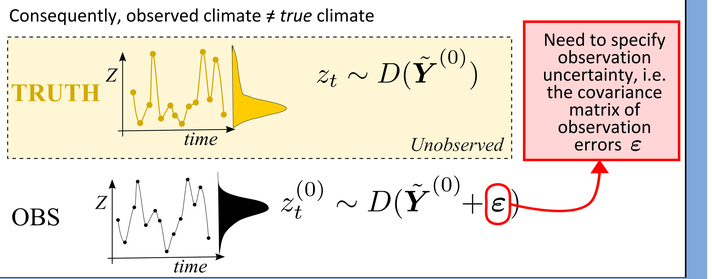
Bivariate model: annual (temperature,precipitation) $\sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

etc.

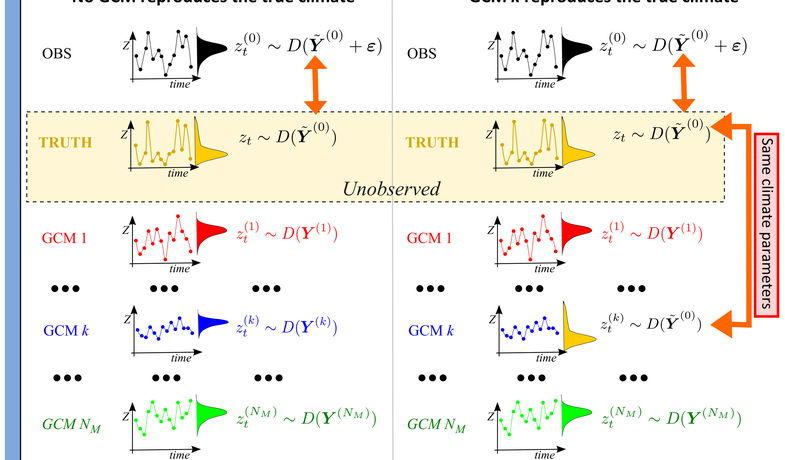
Observation uncertainty

Observations are uncertain $\Longrightarrow$ observed weather ≠ *true* weather

Consequently, observed climate ≠ *true* climate



Hypotheses



Hypotheses Weights

Basic steps

Compute likelihood of weather data (simulated + observed packed together) under each hypothesis

GCM weights are proportional to maximized log-likelihood

Weights can be normalized to add up to one

Technical details: see papers referenced below

Noteworthy properties

(based on synthetic experiments, not shown)

Ability to identify the correct GCM increases when:

- length of z_t increases
- level of error in other GCMs increases
- obs. uncertainty decreases

Similarly for ability to identify that no GCM is correct

When obs. uncertainty is very large, all GCMs get same weight

Renard, B., and J.-P. Vidal (2016a), Performance weighting of GCMs. Part 1: A method based on explicit probabilistic models and accounting for observation uncertainty., Climate Dynamics. To be submitted.

Renard, B., and J.-P. Vidal (2016b), Performance weighting of GCMs. Part 2: Daily atmospheric variability across Europe., Climate Dynamics. To be submitted.

Case studies

Example 1: Icelandic storms

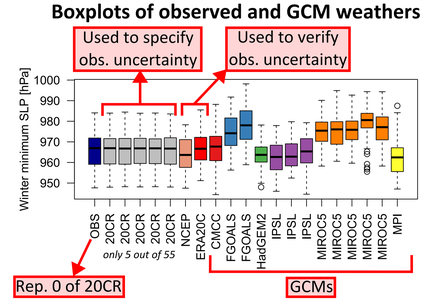
Basic data: daily Sea Level Pressure (SLP) fields, period 1950-2004

"Observed" data: reanalyzes 20CR (56 replications), NCEP/NCAR and ERA20C

Simulated data from 6 CMIP5 GCMs, some of them with replications

Weather variable = winter minimum SLP in Reykjavic $z_t^{(k)} \sim GEV(\mu, \sigma, \xi)$

Boxplots of observed and GCM weathers



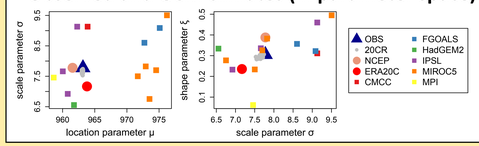
Variability in 20CR replications is very small $\Longrightarrow$ observation uncertainty very small

But comparison with ERA20C and NCEP/NCAR suggests it may be *too* small

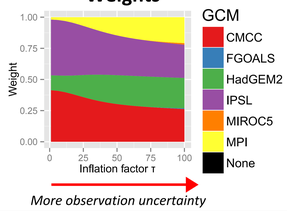
Introduction of an inflation factor to artificially increase observation uncertainty

Weights markedly depend on this inflation factor, which is at best a placeholder...

Observed and GCM climates (in parameter space)



Weights



Example 2: NAO poles (same data sources)

Bivariate weather variable = winter average SLP in (Reykjavic;Lisbon)

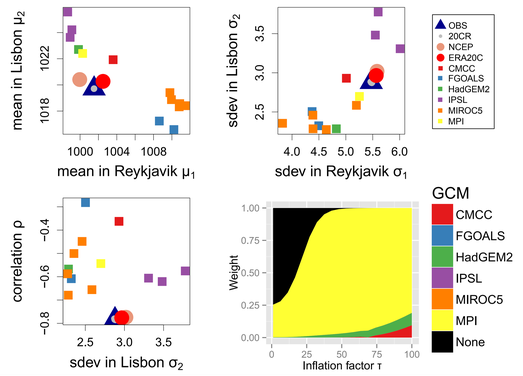
$$\boldsymbol{z}_t^{(k)} \sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right)$$

Similar result regarding 20CR variability => similar inflation factor

Weight for $H^{(0)}$ "no GCM reproduces the true climate" is non-zero...

Due to all GCMs underestimating correlation between NAO poles?

Observed/GCM climates (in parameter space) and weights



Example 3: Daily atmospheric variability accross Europe (Now 23 GCMs)

Weather variable = daily SLP field over the North Atlantic => 1760-dimensional variable!

Derivation of a tractable probabilistic model

Principal Component Analysis (PCA) of observed daily SLP field

8 components explain ~75% of total variability

$\Longrightarrow$ 9 basis functions (BF): the 8 PCs + the interannual mean field

Raw SLP data projected into the 9 BFs $\Longrightarrow$ 9-dimensional variable $z_t^{(k)}$

For each component, a 1-dimensional Gaussian model that includes:

- seasonality in the mean and the standard deviation
- autocorrelation as an order-1 autoregressive model

Results

HadGEM2-CC has a weight close to 1

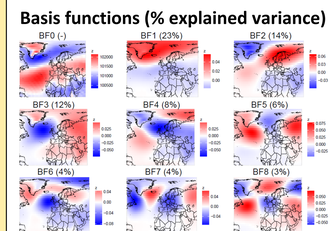
Same analysis repeated for 12 European regions relevant for renewable energy production (François et al. ,2016)

Longitudinal gradient in GCM performance

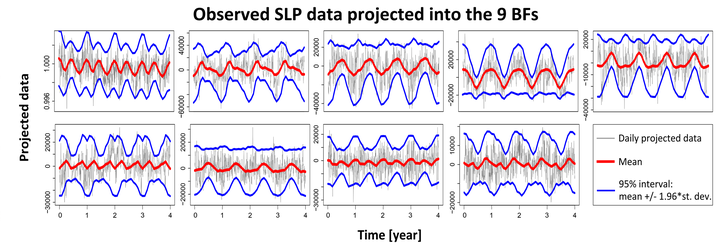
Strong impact of observation uncertainty

François, B., B. Hingray, D. Raynaud, M. Borga, and J. D. Creutin (2016), Increasing climate-related-energy penetration by integrating run-of-the river hydropower to wind/solar mix, Renewable Energy, 87, Part 1, 686-696.

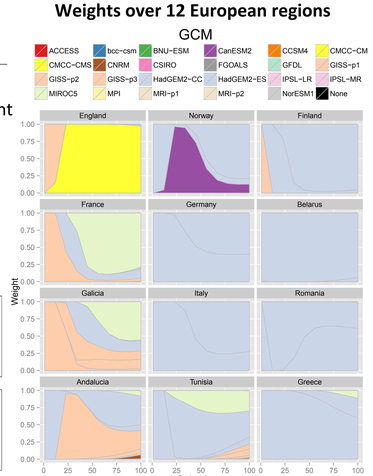
Basis functions (% explained variance)



Observed SLP data projected into the 9 BFs



Weights over 12 European regions



Conclusion & perspectives

Main properties of the proposed weighting procedure

Based on an explicit probabilistic model describing GCM and historical weathers, which may include seasonality, trend, correlation, skew, etc.

Can use several GCM replications whenever available

Accounts for observation uncertainty - indeed, strongly depends on it!

Considers the $H^{(0)}$ hypothesis that no GCM reproduces the true climate

Open questions and perpectives

Key role of obs. uncertainty in measuring GCM performance ... but obs. uncertainty poorly known!
$\Longrightarrow$ **a reliable quantification of uncertainty in reanalyzes is of prime importance**

What should be done when the $H^{(0)}$ weight (no GCM reproduces the true climate) is close to one?

Relationship with attribution and internal variability: distinguish climate properties that should be reproduced by GCMs (and should therefore be included in the probabilistic model) from those that should not (and should therefore NOT be included in the probabilistic model)

Usage of the weights to propagate GCM uncertainty in a modeling chain for impact studies