Formulation of a statistical Emulator of the Climate Response to astronomical forcing variations

Nabila BOUNCEUR - Michel CRUCIFIX

nabila.bounceur@uclouvain.be - michel.crucifix@uclouvain.be

Georges Lemaître Centre for Earth and Climate Research, Chemin du Cyclotron 2, 1348 Louvain-la-Neuve , Belgium

UCL Université catholique de Louvain

Aims

Formulate and develop a GCM surrogate for the response of the climate's fast components (atmosphere, ocean, land surface) to variations of the astronomical forcing during the Pleistocene, requiring only a small amount of computing time and providing uncertainty estimates. For the construction and calibration of our climate simulator surrogate, we combine three methods. A "space filling design" to choose the set of input parameter used to run the experiments, a multivariate analysis technique to derive dominant modes of climate variability on the output data, and a multivariate Gaussian Process to emulate the simulator's response.

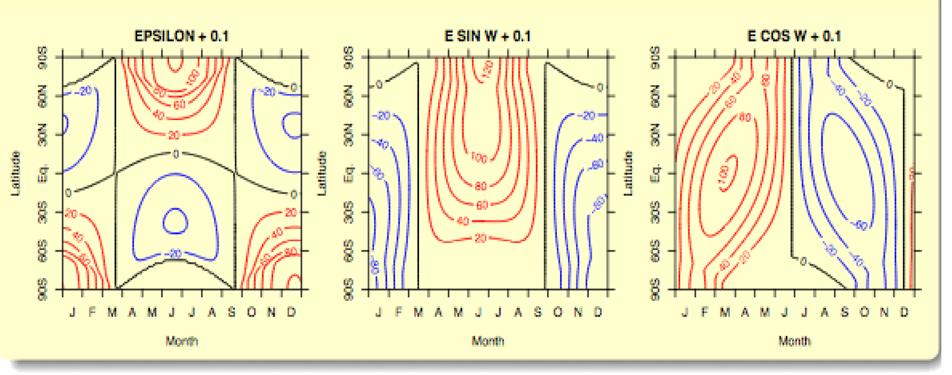
Hypothesis

Insolation during long time variations is influenced by the eccentricity e the longitude of the perihelion ϖ and the obliquity ε

To deal with astronomical theory of paleoclimate, input data are expressed in an adequate form and then consider the basis :

$$(\varepsilon,\ e\cdot\sin(\varpi),\ e\cdot\cos(\varpi))$$

Effect on spatio-temporal distribution of insolation [W/m^2]

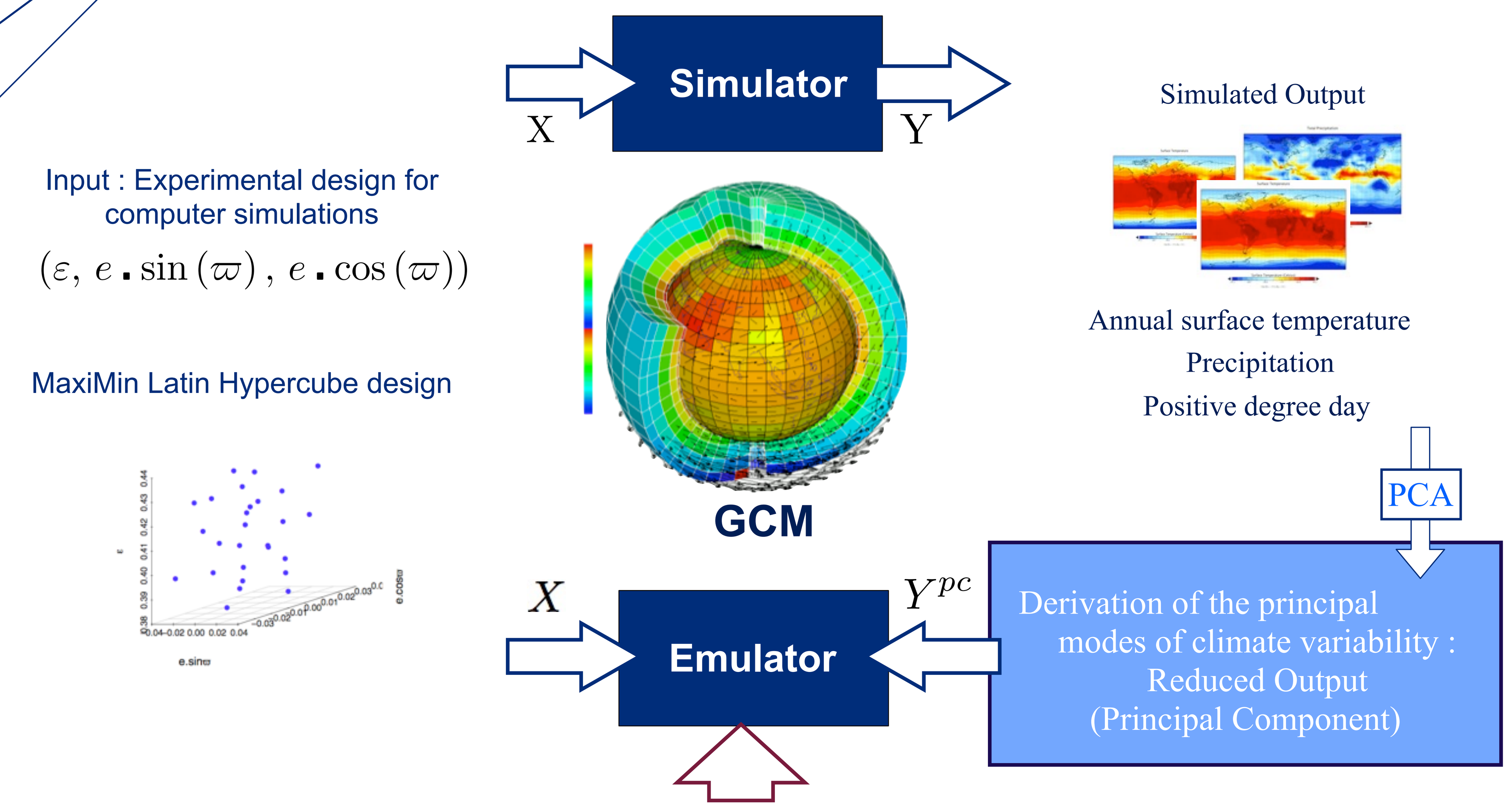


Methodology

The statistics of the fast components of the climate system could be in principle estimated with a general circulation model of the atmosphere and the ocean (AOGCM). However, the demand on computing resources would be far too excessive. One possible approach for dealing with a computationally expensive simulator is the construction of a statistical model based on the available runs, that can replace the simulator. This is a statistical regression problem. As a surrogate, an emulator is a term used to mean the full probabilistic specification for the statistic model. We use a Gaussian Process Emulator. It provides both an estimate of the model and quantifies uncertainty about evaluating the emulator at a limited number of input data.

Weighted principal component analysis (WPCA) is applied to project the output data onto a lower dimensional manifold to explore and extract the principal modes of climate variability. The principal components (PCs) are ranked according to the order of decreasing eigenvalues. In order to maximize the information about the model response, using an optimal number of experiments, the set of input parameters was designed following a space filling design.

The experimental design is chosen independent from the model. Here, we choose a MaxiMin Latin Hypercube Design which maximizes the minimum distance between design points but requires even spacing of the levels of each input data. This design permit to explore more parameters than a grid using the same ensemble size, as the later don't need to grow exponentially with the dimensionality of the parameter space.



The Emulator interpolate model reduced dimensional output space over parameter space reduced dimensional output space

$$\eta(.) = \beta h(.) + \Gamma(.), \quad \Gamma(.) \text{ is a GP}$$

A Gaussian Process (GP) is a stochastic process for which any output data has a multivariate Gaussian distribution as a prior. It is specified by a mean function and a covariance function. The set of simulations is used to update this prior. For the covariance function we use a separable squared exponential.

$$\eta(.) \sim GP\left(m(.), \sigma^2 c(.,.)\right)$$

$$m(.) = \beta h(.)$$

$h(.)$ is a vector of functions

β a regression parameters

$$c(x_1, x_2) = \Pi_i \exp\left(-\left(x_1^i - x_2^i\right)^2 / \lambda_i^2\right)$$

The output is than reconstructed from the principal component space to the original full space.

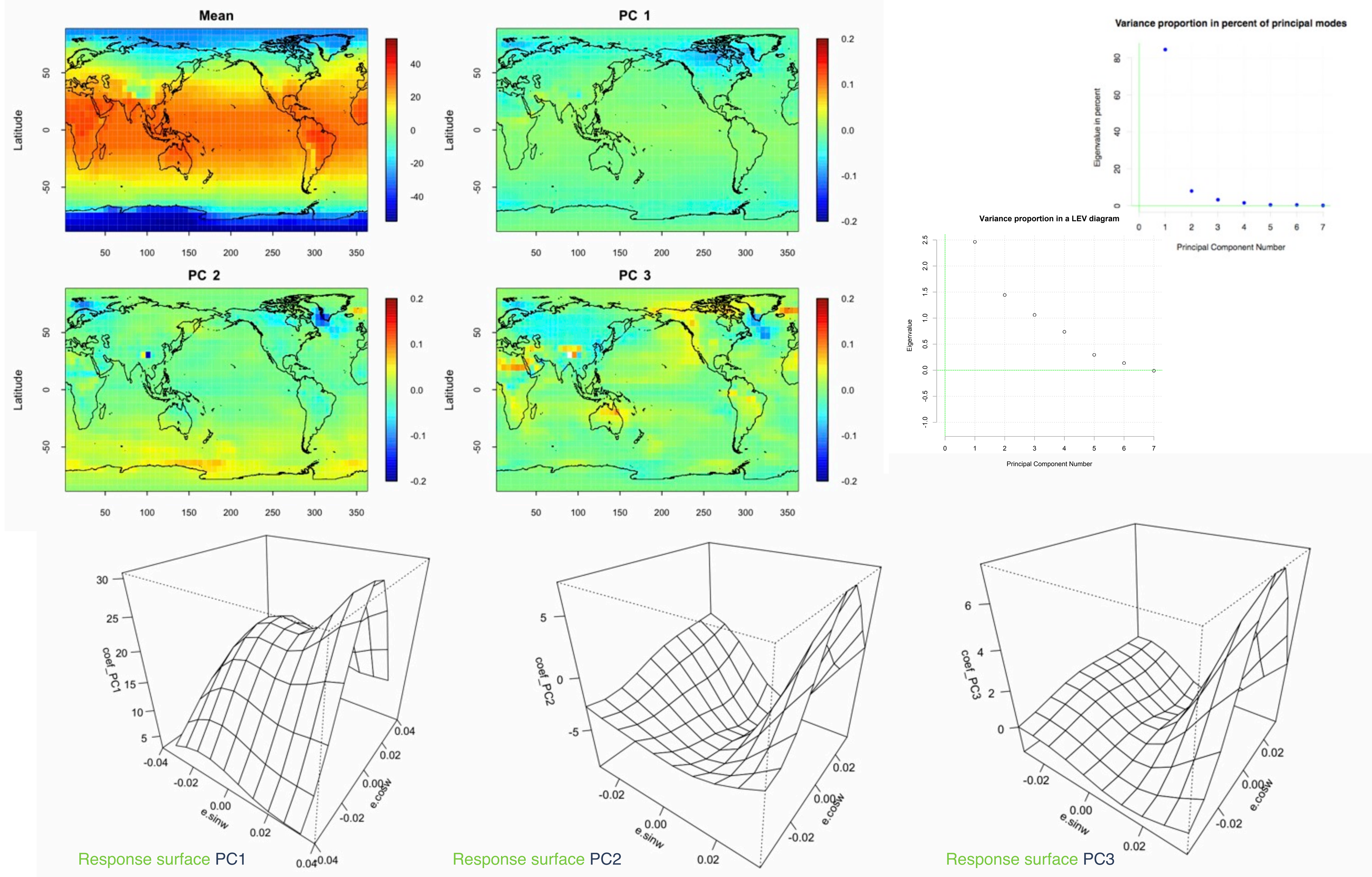
Application

For a first application, we have developed and designed an emulator of a three-dimensional Earth system model of intermediate complexity (LOVECLIM, Goosse et al., 2010), considering the principal components of its response (mean surface temperature). The first three principal component account for 99% of the data variance (figure 1).

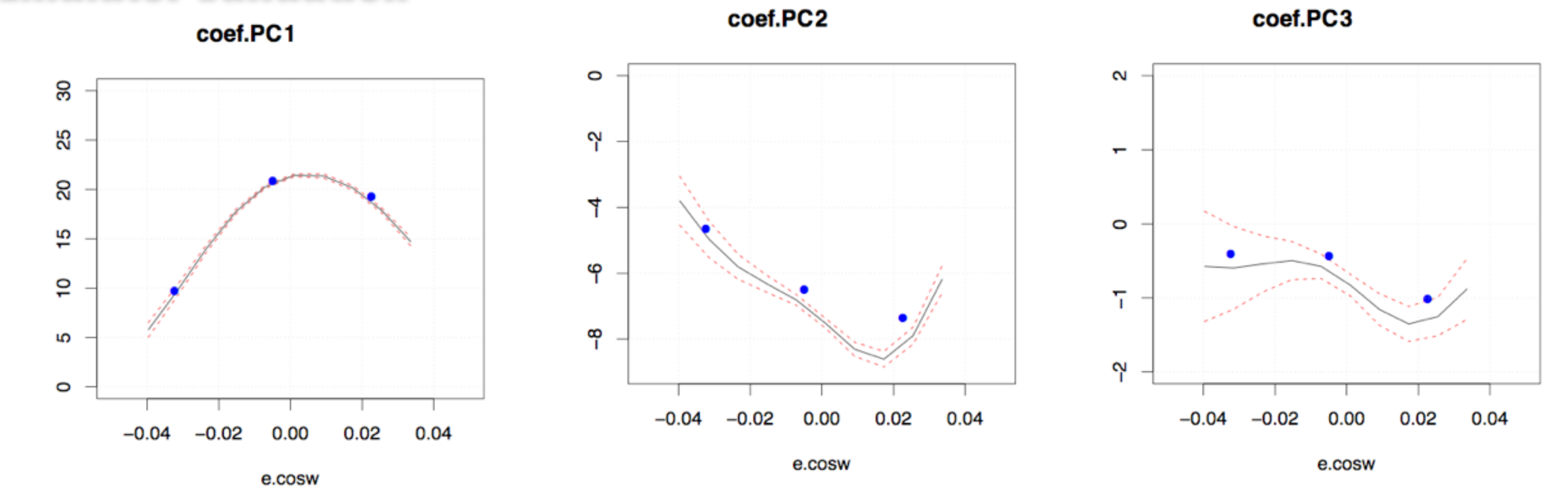
They correspond to the polar mode, planetary waves mode and the monsoon one (figure 2). GP emulation allow to better capture the nonlinearities of the output without necessitating an excessive number of experiments. We have considered 3 levels for the 3 input parameter and made 27 computer experiments.

Figure 2: Principal component analysis : annual mean surface temperature [○C]

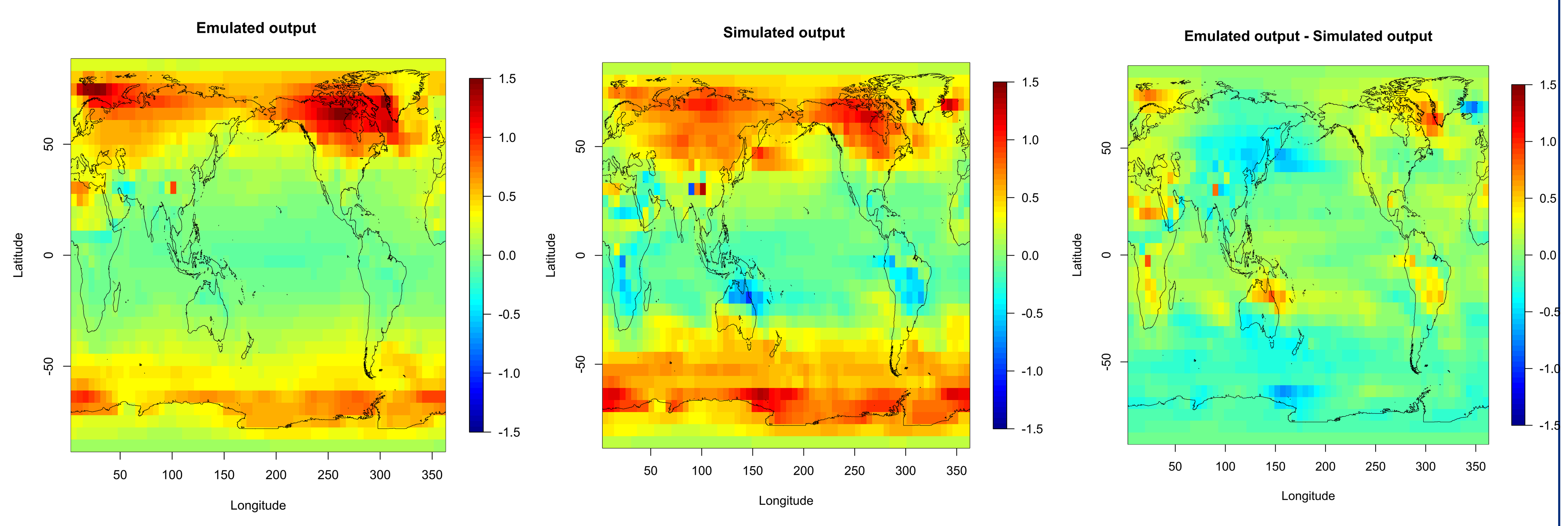
Figure 1



Emulator Validation



Annual mean surface temperature ε = 0.383972400 , e.sinw = -.001461294 , ε = 22○



Experiment design and emulators : efficient computing and understanding