



EC-Earth meeting 5-6 May 2015

Do different machines simulate different climates?

François Massonnet

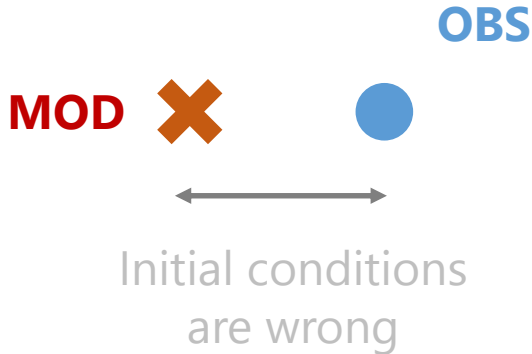
M. Asif, O. Bellprat, E. Exarchou, M. Ménégoz,
C. Prodhomme, F. J. Doblas-Reyes

Identifying sources of model error

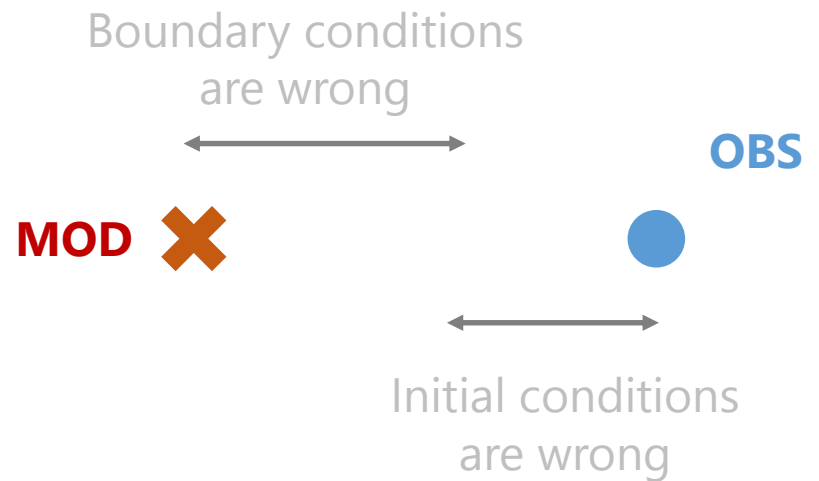
OBS

MOD

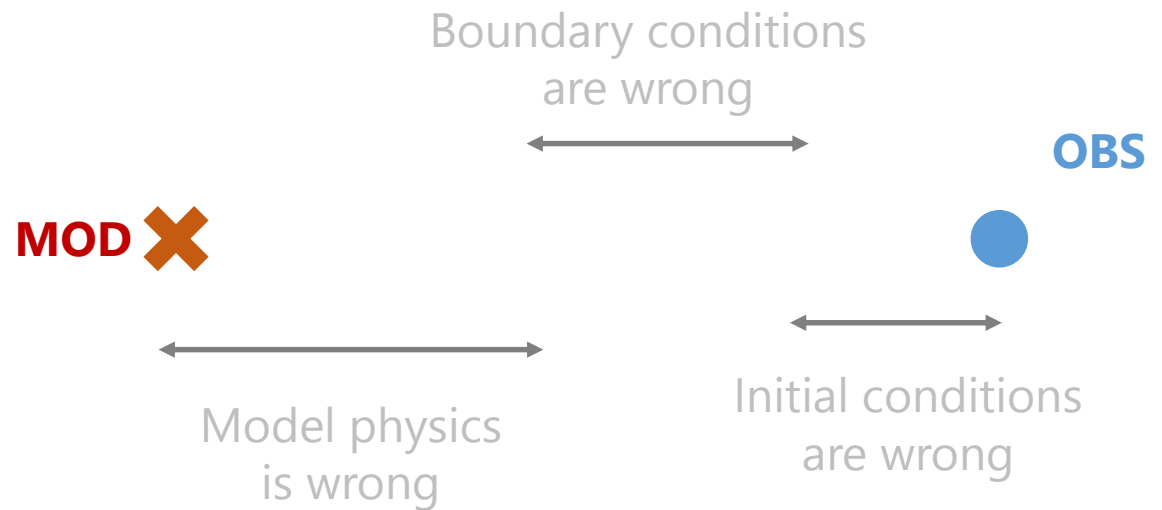
Identifying sources of model error



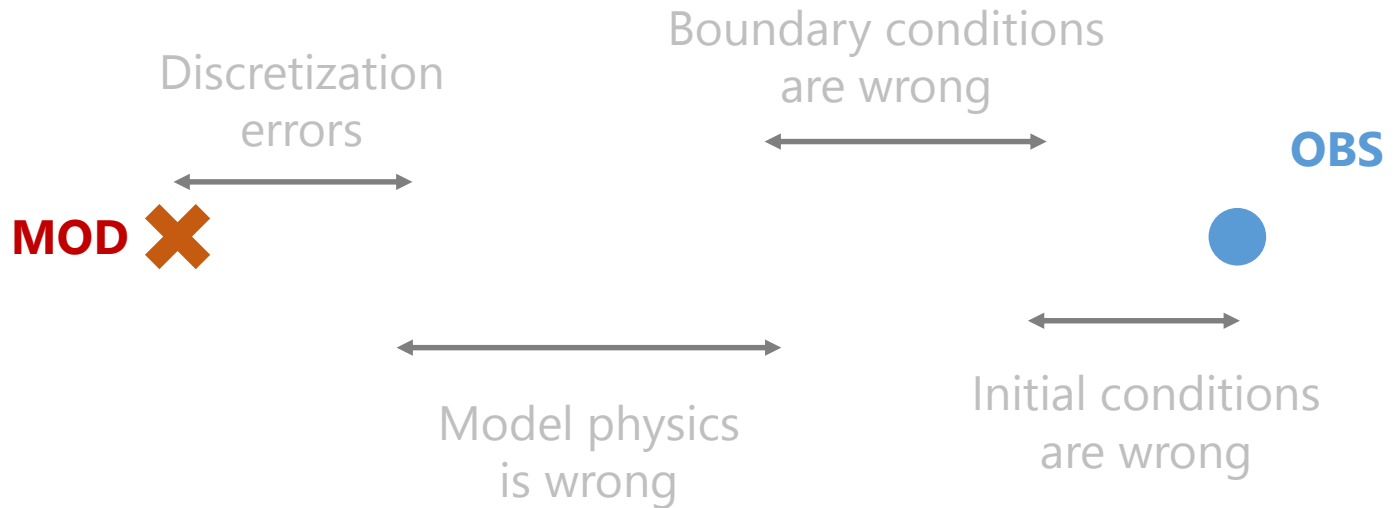
Identifying sources of model error



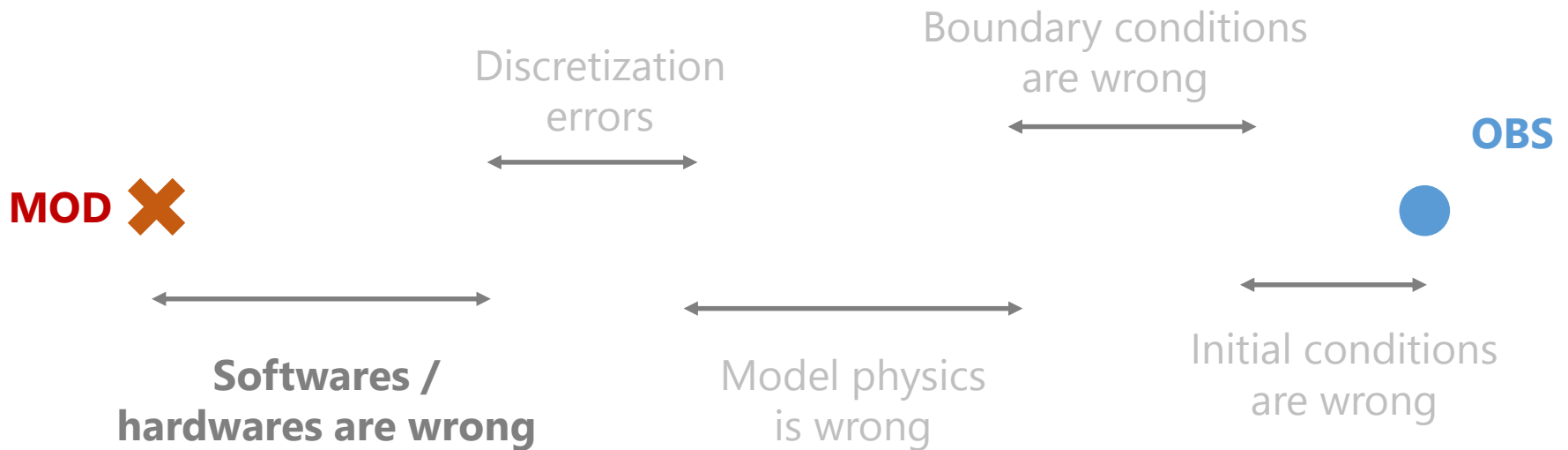
Identifying sources of model error



Identifying sources of model error



Identifying sources of model error



Hardware/software as sources of model error

Bit-reproducibility of EC-Earth

Clim-reproducibility of EC-Earth

Hardware/software as sources of model error

Bit-reproducibility of EC-Earth

Clim-reproducibility of EC-Earth

Software / hardware is a multiple and underestimated source of model error

Round-off errors and floating-point representation
The order matters: associativity is no longer valid

$$(\sqrt{2} \cdot \pi) \cdot e = \sqrt{2} \cdot (\pi \cdot e)$$

12.0771 \neq 12.0777

Software / hardware is a multiple and underestimated source of model error

Round-off errors and floating-point representation

The order matters: associativity is no longer valid

$$(\sqrt{2} \cdot \pi) \cdot e = \sqrt{2} \cdot (\pi \cdot e)$$

12.0771 \neq 12.0777

Aggressive optimization

Can degrade accuracy [Thomas et al. Wea. And Forecast., 2002]

Software / hardware is a multiple and underestimated source of model error

Round-off errors and floating-point representation

The order matters: associativity is no longer valid

$$(\sqrt{2} \cdot \pi) \cdot e = \sqrt{2} \cdot (\pi \cdot e)$$

12.0771 \neq 12.0777

Aggressive optimization

Can degrade accuracy [Thomas et al. Wea. And Forecast., 2002]

Number of processors and their distribution

Processor topology defines order of operations [Thomas et al., Wea. And Forecast., 2002; Senoner et al., AIAA, 2008]

Software / hardware is a multiple and underestimated source of model error

Round-off errors and floating-point representation

The order matters: associativity is no longer valid

$$(\sqrt{2} \cdot \pi) \cdot e = \sqrt{2} \cdot (\pi \cdot e)$$

12.0771 \neq 12.0777

Aggressive optimization

Can degrade accuracy [Thomas et al. Wea. And Forecast., 2002]

Number of processors and their distribution

Processor topology defines order of operations [Thomas et al., Wea. And Forecast., 2002; Senoner et al., AIAA, 2008]

Compiler version

Different FORTRAN compilers can produce different outcomes [Lawrence et al., EOS, 1999]

Software / hardware is a multiple and underestimated source of model error

Round-off errors and floating-point representation
The order matters: associativity is no longer valid

$$(\sqrt{2} \cdot \pi) \cdot e = \sqrt{2} \cdot (\pi \cdot e)$$

12.0771 \neq 12.0777

Aggressive optimization

Can degrade accuracy [Thomas et al. Wea. And Forecast., 2002]

Number of processors and their distribution

Processor topology defines order of operations [Thomas et al., Wea. And Forecast., 2002; Senoner et al., AIAA, 2008]

Compiler version

Different FORTRAN compilers can produce different outcomes [Lawrence et al., EOS, 1999]

Unpredictable hardware failures

[Düben and Palmer, Mon. Wea. Rev., 2014]

Hardware/software as sources of model error

This aspect has been overlooked but is of non-negligible importance

Bit-reproducibility of EC-Earth

Clim-reproducibility of EC-Earth

Hardware/software as sources of model error

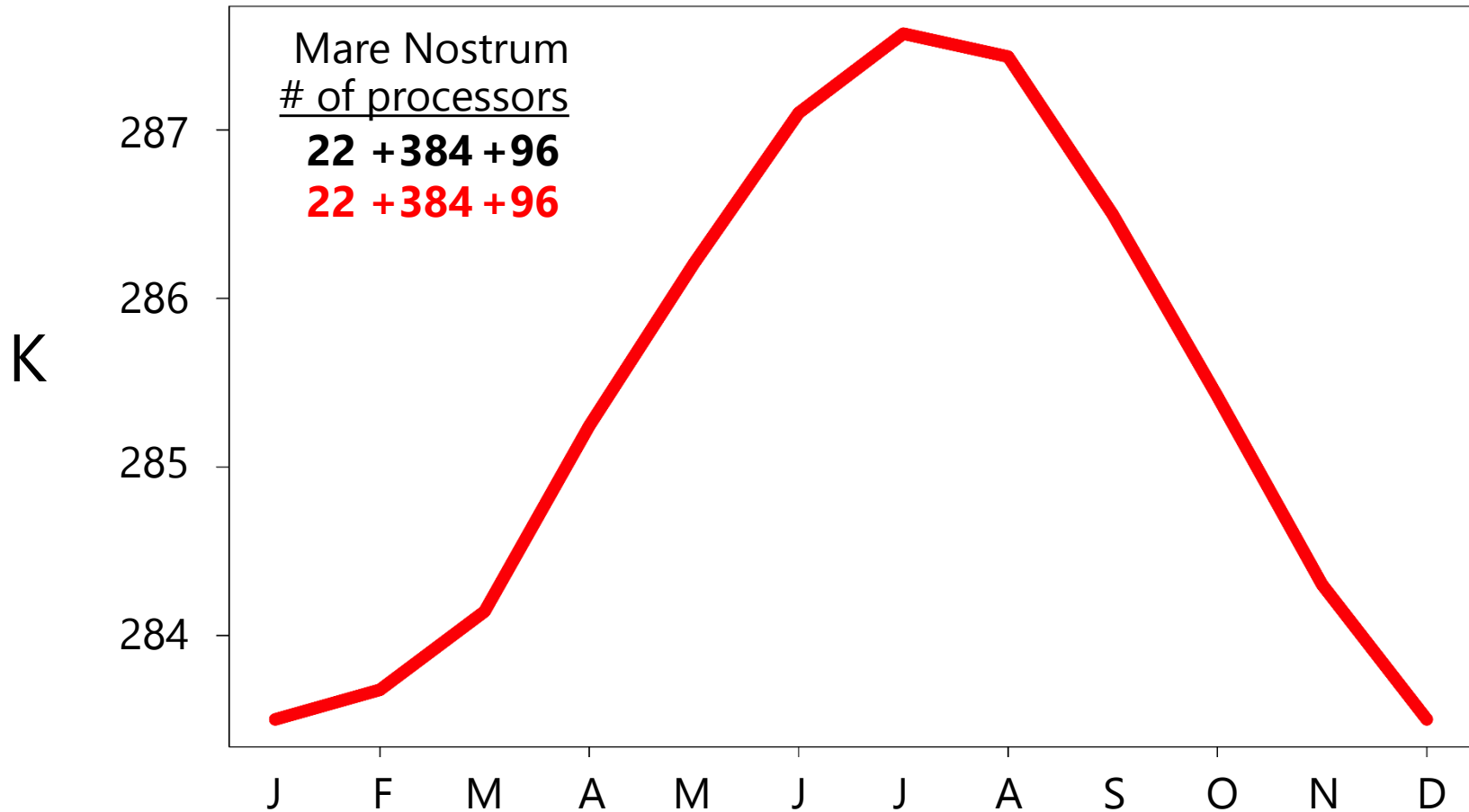
This aspect has been overlooked but is of non-negligible importance

Bit-reproducibility of EC-Earth

Clim-reproducibility of EC-Earth

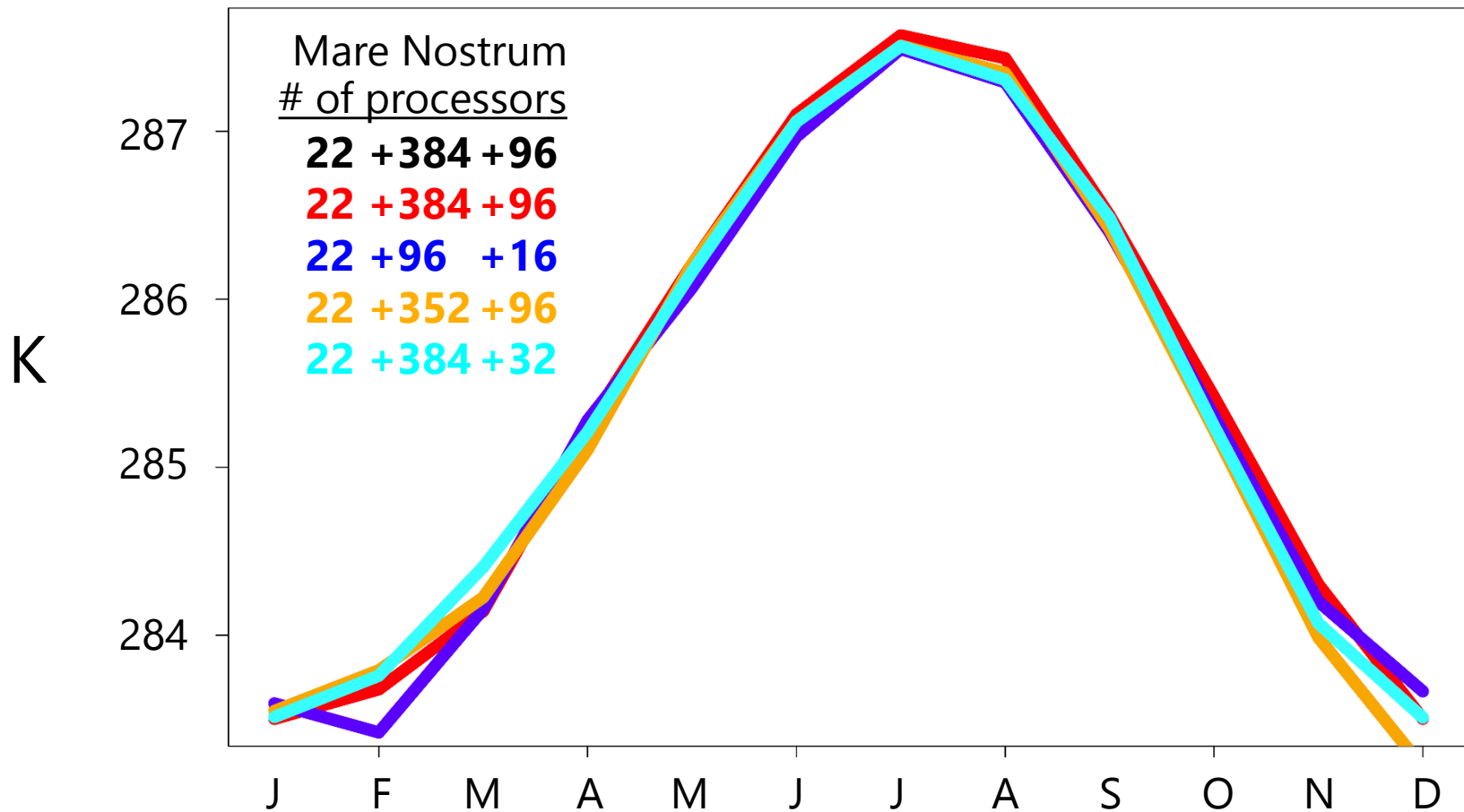
All things being equal, EC-Earth
is reproducible bitwise

Near-surface global temperature



All things being equal, EC-Earth is sensitive to processor distribution

Near-surface global temperature



Hardware/software as sources of model error

This aspect has been overlooked but is of non-negligible importance

Bit-reproducibility of EC-Earth

- Two identical runs give exactly matching output
- No reproducibility for different processor (IFS or NEMO) distribution

Clim-reproducibility of EC-Earth

Hardware/software as sources of model error

This aspect has been overlooked but is of non-negligible importance

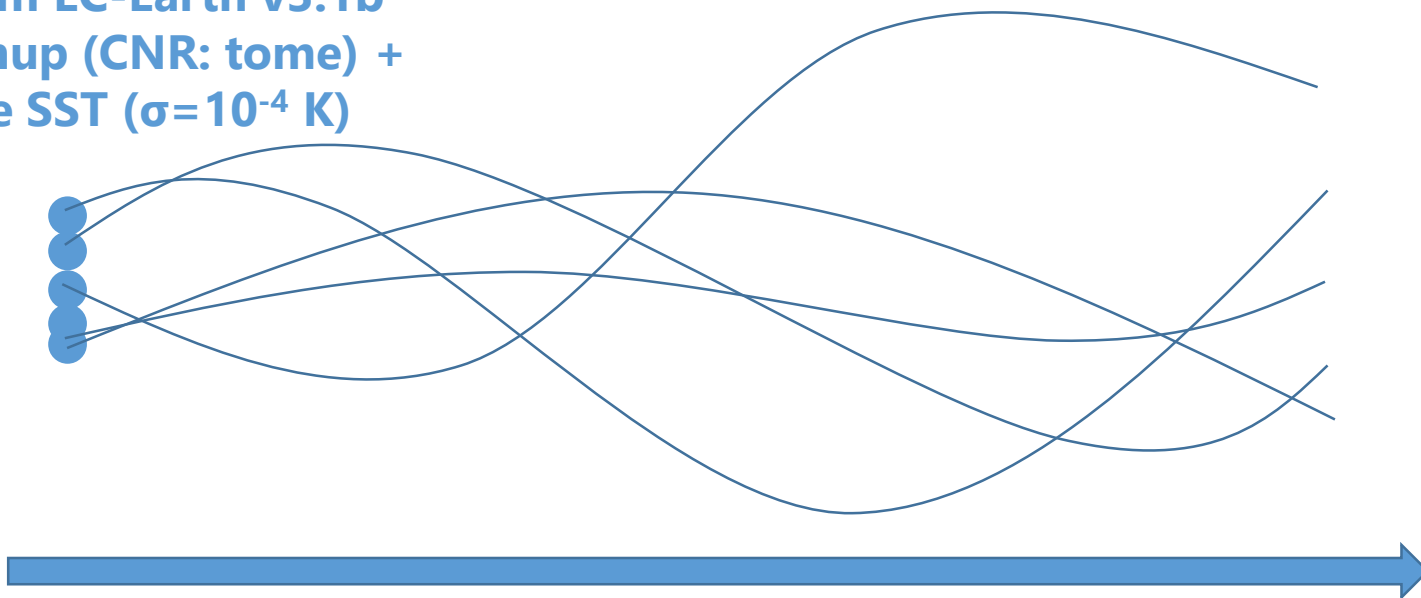
Bit-reproducibility of EC-Earth

- Two identical runs give exactly matching output
- No reproducibility for different processor (IFS or NEMO) distribution

Clim-reproducibility of EC-Earth

Restart from EC-Earth v3.1b
500 yr-spinup (CNR: tome) +
white noise SST ($\sigma=10^{-4}$ K)

5 members









1850

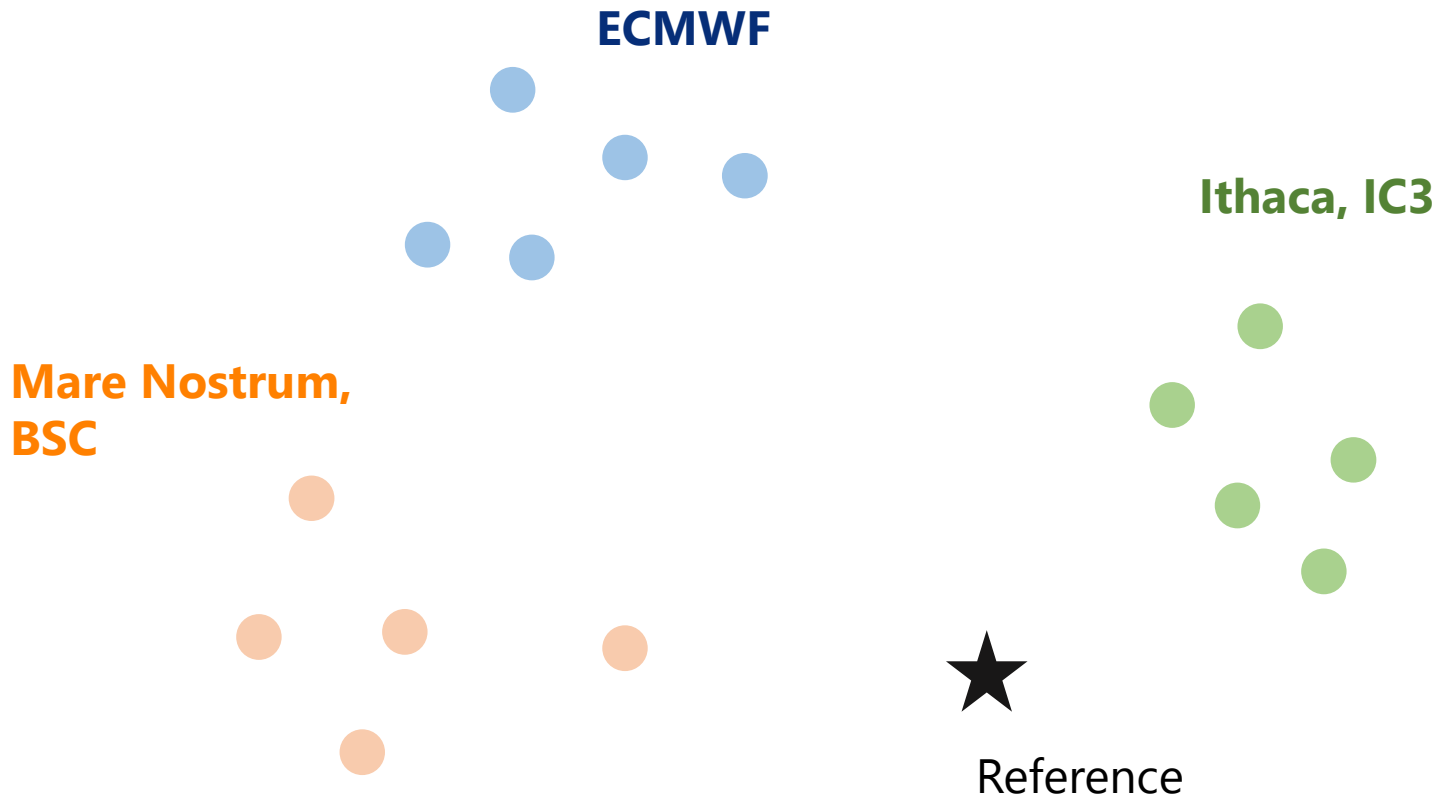
1860

1870

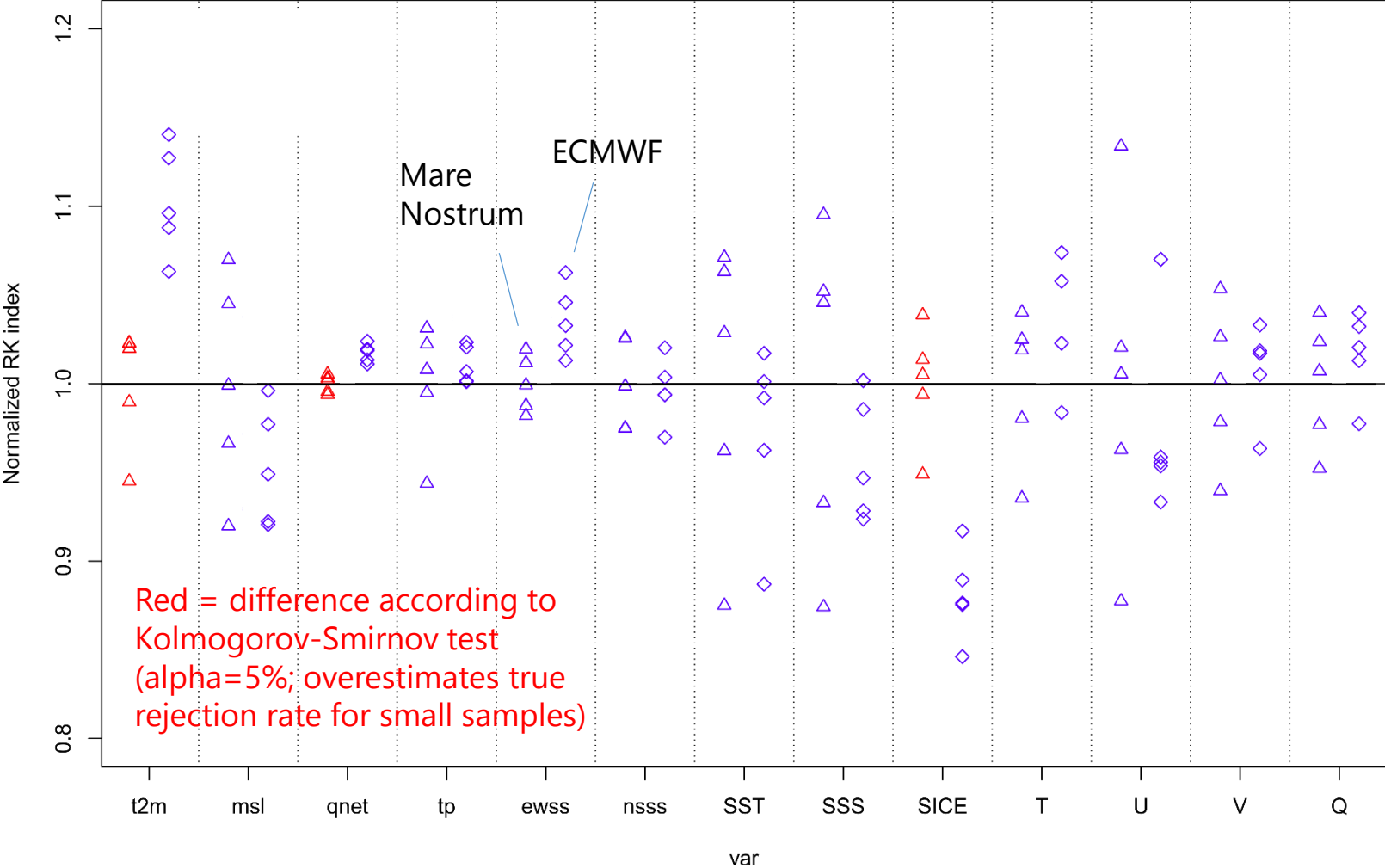
Forcing: pre-industrial

	Machine 1 (Mare Nostrum, BSC)	Machine 2 (ECMWF)	Machine 3 (Ithaca, CFU)
Motherboard			
Operating system			
Compilation flags	Identical	Identical	Identical
NetCDF, GRIB, HDF5 libraries	Different	Different	Different
# of processors	22 + 384 + 96	22 + 480 + 96	22 + 32 + 16
	Autosubmit ensures identical configurations		

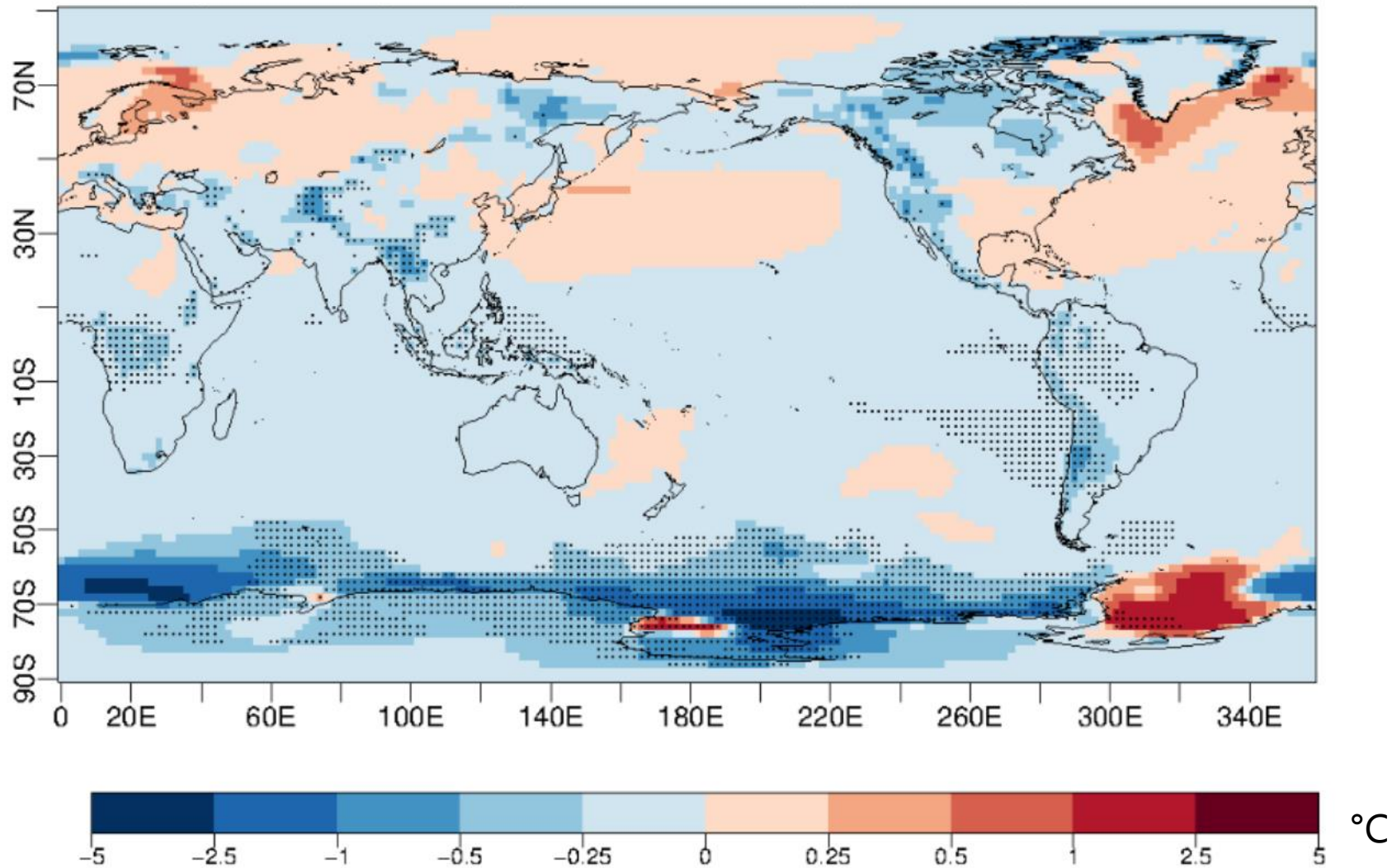
The simulations are evaluated against the same, static data set



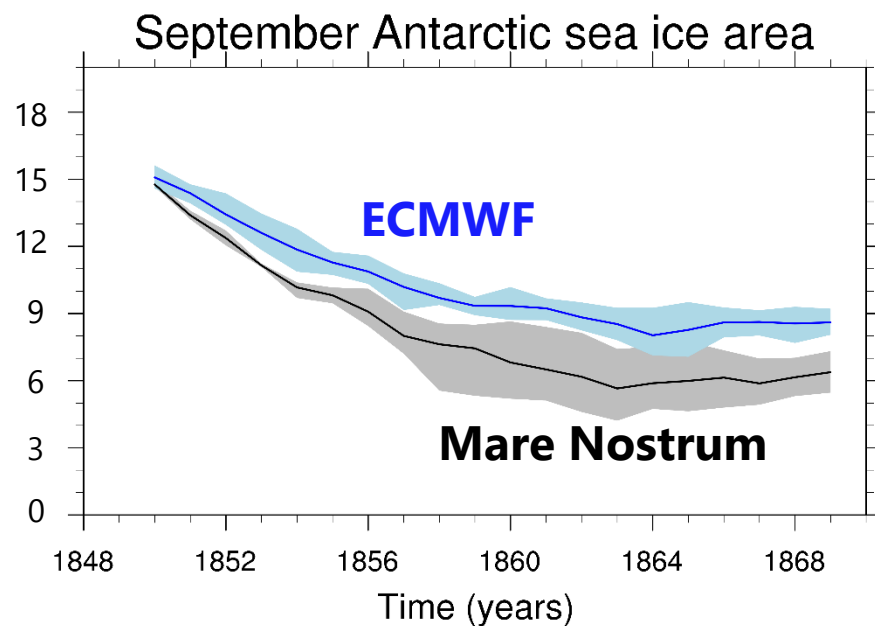
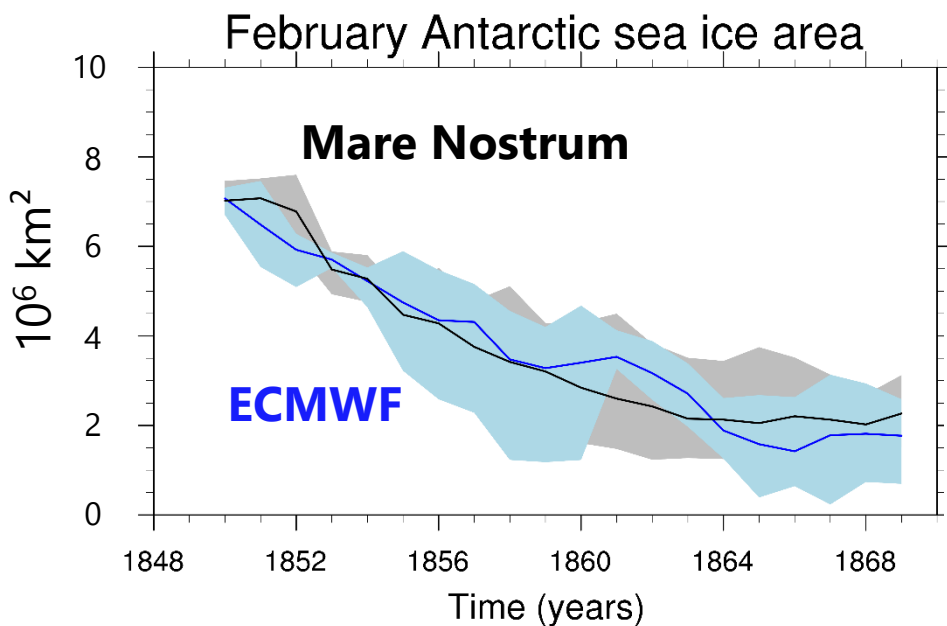
Statistically significant differences are found in the performance indices



Difference (ECMWF – Mare Nostrum) near-surface temperature



Differences originate in winter.
Associated to deep oceanic convection
and parameterizations?



Hardware/software as sources of model error

This aspect has been overlooked but is of non-negligible importance

Bit-reproducibility of EC-Earth

- Two identical runs give exactly matching output
- No reproducibility for different processor (IFS or NEMO) distribution

Clim-reproducibility of EC-Earth

EC-Earth is not climate-reproducible on different platforms

Take home messages and implications

1. Machines introduce an additional, non negligible source of error in climate simulations

The results are supported by several published studies, unpublished documents, colloquial discussions, good practices

2. Following the precautionary principle, CMIP6 simulations should be centralized...

... unless a benchmark experiment is produced to re-assess reproducibility with the next EC-Earth version

3. Model evaluation should account for dependency of results on software/hardware

Machines sample uncertainty just as members sample internal variability

4. Control and sensitivity experiments must be designed on the same machine

Thank you!

francois.massonnet@uclouvain.be

www.climate.be/u/fmasson



@FMassonnet