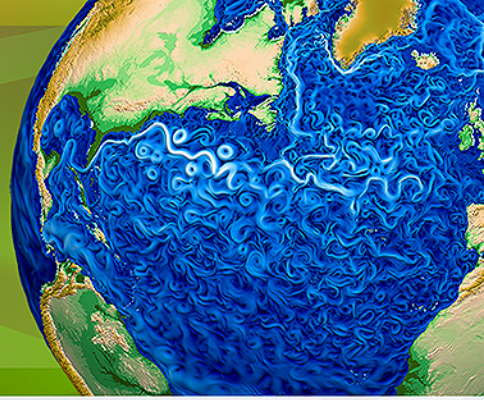
ACME
Accelerated Climate Modeling for Energy



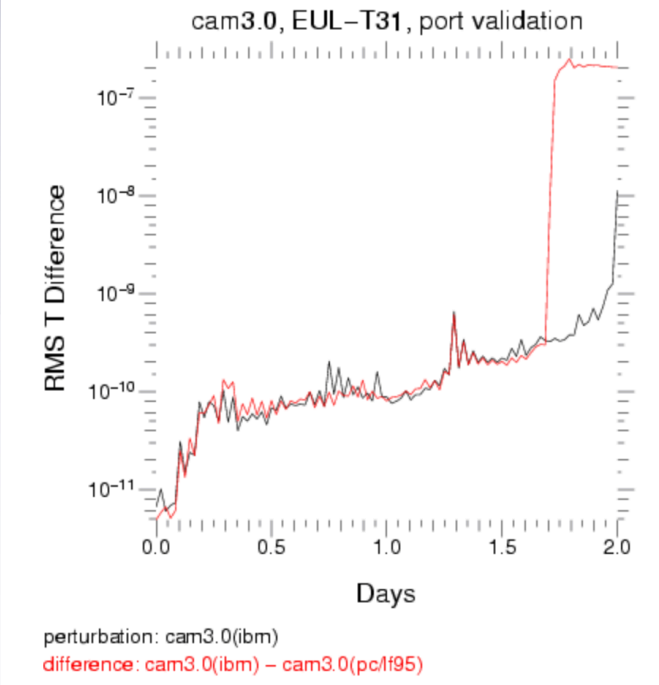
Non-BFB Solution Reproducibility:
A New Test Based on Time Step Convergence

Hui Wan, Kai Zhang, Phil Rasch, Balwinder Singh,
Xingyuan Chen (PNNL), and Jim Edwards (NCAR)

U.S. DEPARTMENT OF ENERGY

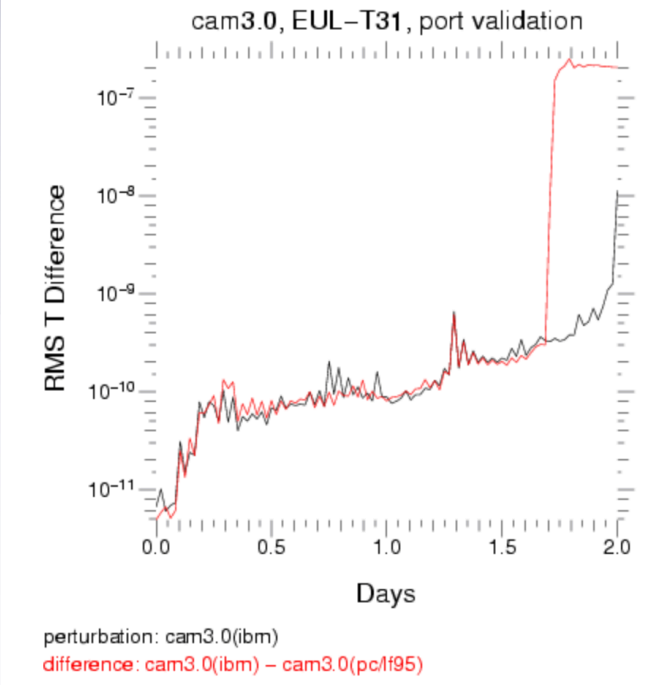
- Model results are no longer bit-for-bit identical to a reference run – has the climate changed as well?
- There used to be a simple and inexpensive way to find out…

- Model results are no longer bit-for-bit identical to a reference run – has the climate changed as well?
- There used to be a simple and inexpensive way to find out…



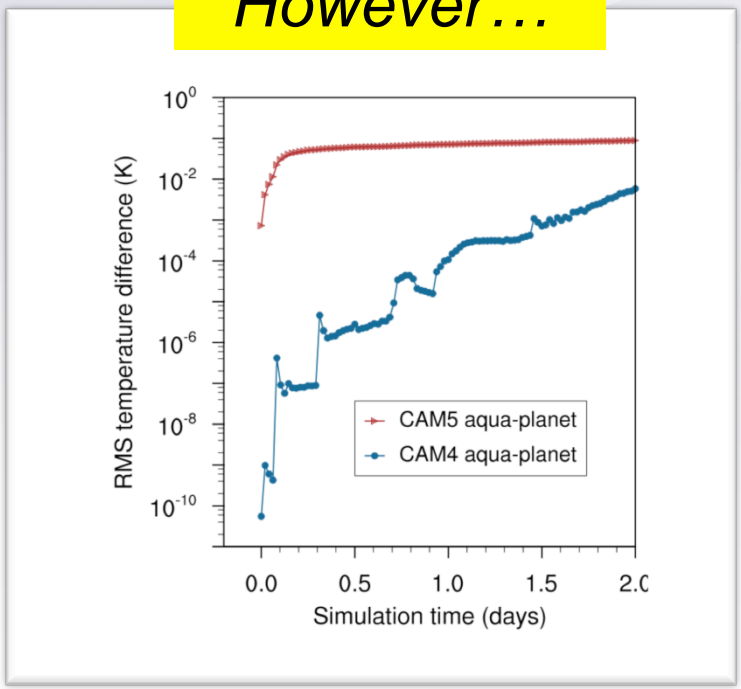
http://www.cesm.ucar.edu/models/atm-cam/port/

- Model results are no longer bit-for-bit identical to a reference run – has the climate changed as well?
- There used to be a simple and inexpensive way to find out…



http://www.cesm.ucar.edu/models/atm-cam/port/

However…



Wan et al. (2016, submitted)

NCAR's New Climate Consistency Test (Baker et al., 2015, GMD)

- Test ensemble contains 3 one-year simulations
- Reference ensemble contains 151 one-year simulations
- Analysis uses global and annual mean values of h0 output

NCAR's New Climate Consistency Test (Baker et al., 2015, GMD)

- Test ensemble contains 3 one-year simulations
- Reference ensemble contains 151 one-year simulations
- Analysis uses global and annual mean values of h0 output

- Expensive
- Can miss differences in small-scale features

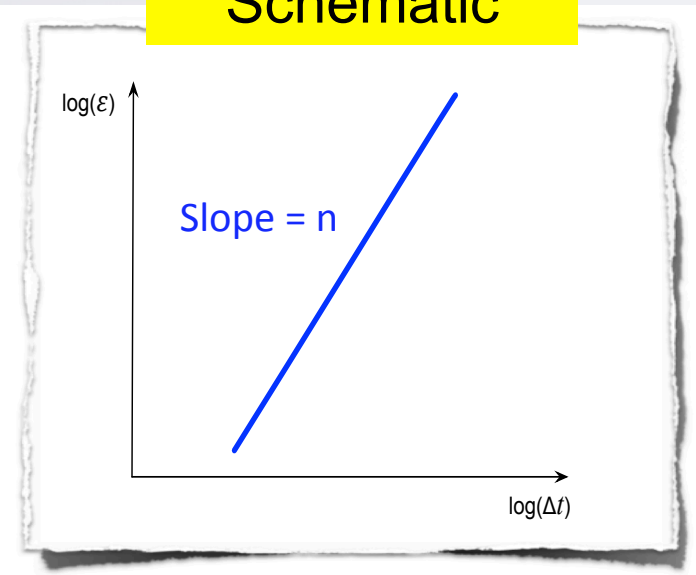
- Can we find a less expensive and more sensitive method?

Introduction to Convergence

- Differential equation $\frac{dy}{dt} = F(y,t)$
- Discretization: time step Δt , total simulation length τ
- Numerical error $\mathcal{E} = C\left(\frac{\Delta t}{\tau}\right)^{n} + \cdots$

order of accuracy
a.k.a. **convergence rate**

Schematic

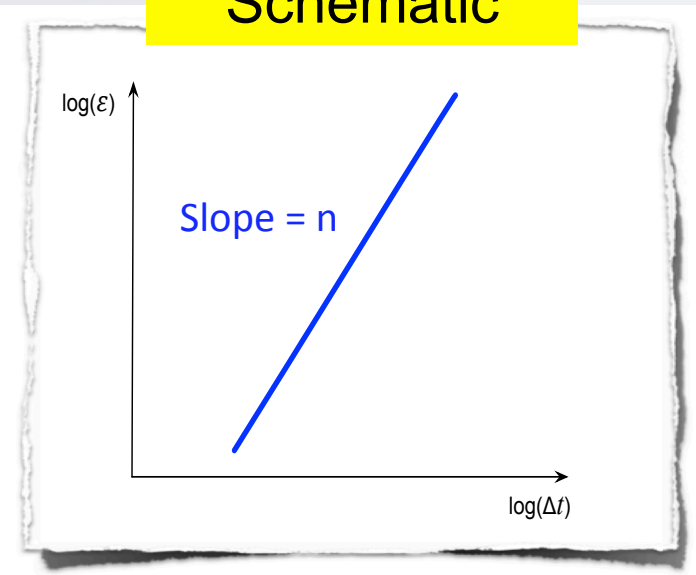


Introduction to Convergence

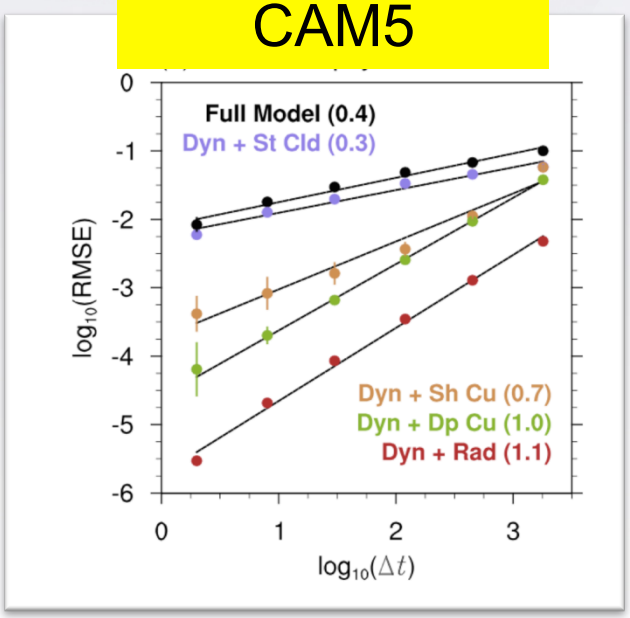
- Differential equation $\frac{dy}{dt} = F(y,t)$
- Discretization: time step Δt , total simulation length τ
- Numerical error $\mathcal{E} = C\left(\frac{\Delta t}{\tau}\right)^{n} + \cdots$

order of accuracy
a.k.a. **convergence rate**

Schematic



CAM5



Wan et al. (2015, JAMES)

An Important Note

- When assessing convergence, the numerical error should be calculated against the "true" solution that the discrete model is trying to solve for!

- dy/dt = F(y,t) => true solution is y = A(t)
- dx/dt = G(x,t) => true solution is x = B(t)

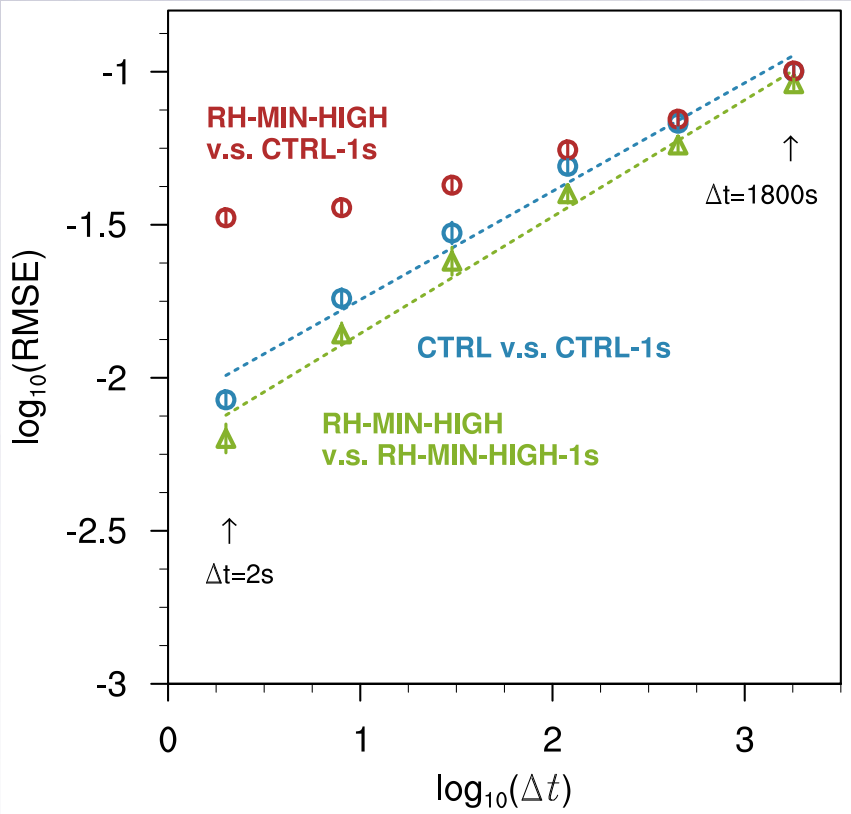
An Important Note

- When assessing convergence, the numerical error should be calculated against the "true" solution that the discrete model is trying to solve for!

- dy/dt = F(y,t) => true solution is y = A(t)
- dx/dt = G(x,t) => true solution is x = B(t)

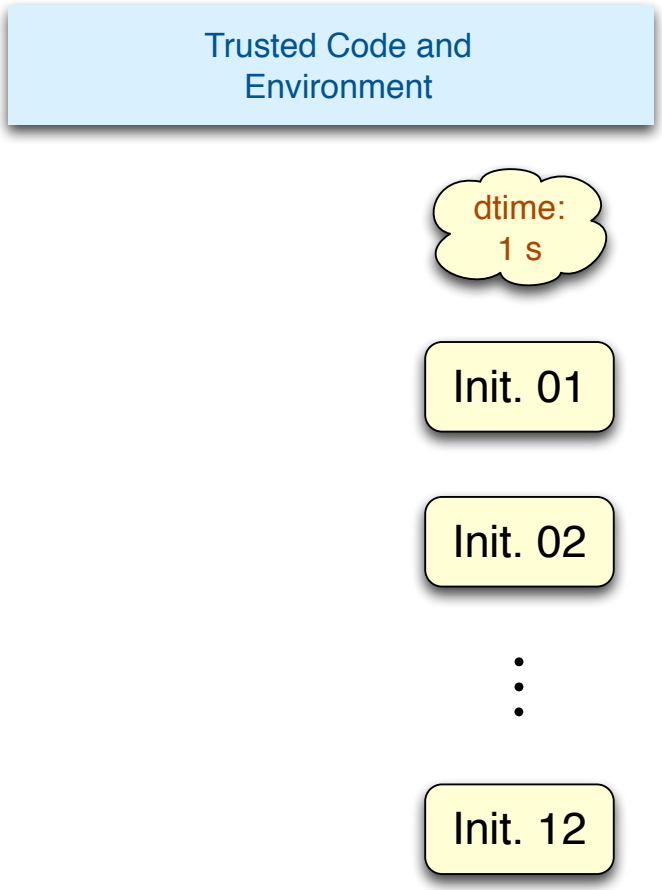
- A change in model parameter, a new parameterization scheme, or the introduction of a bug, will result in a new "true solution" in the mathematical sense.

A Real Example from CAM5.3

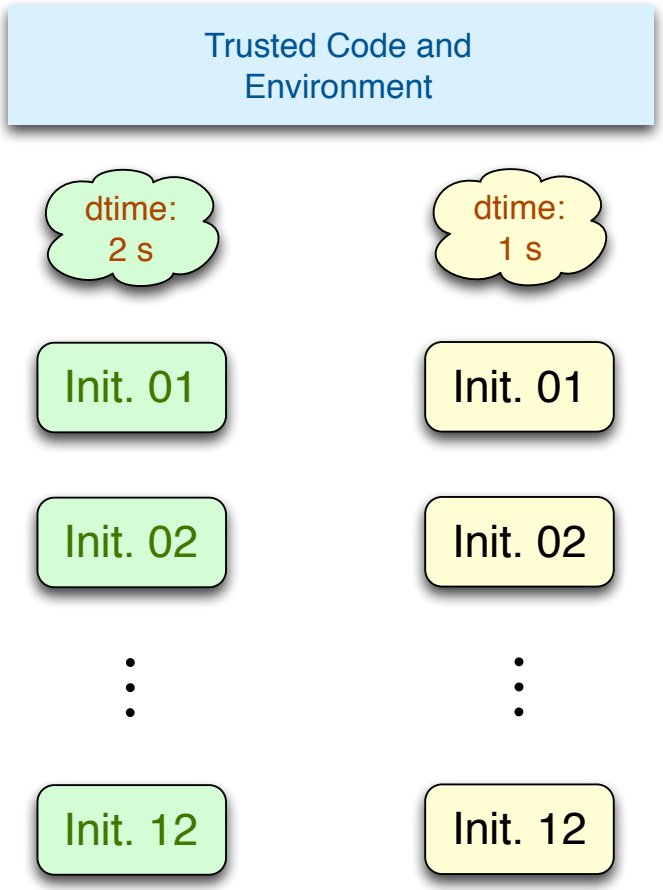


Wan et al. (2016, submitted)

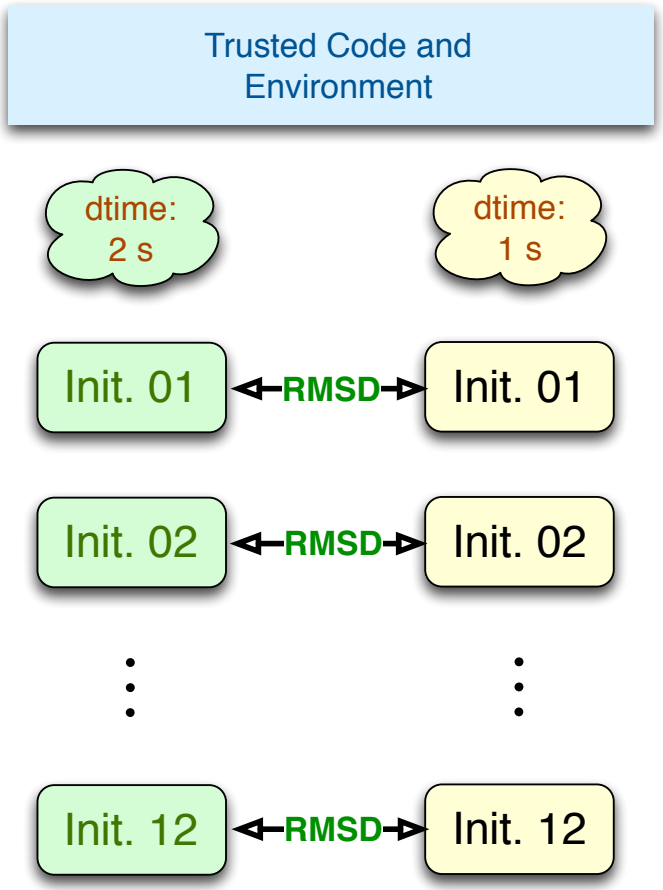
The New Test Method (TSC1.0)



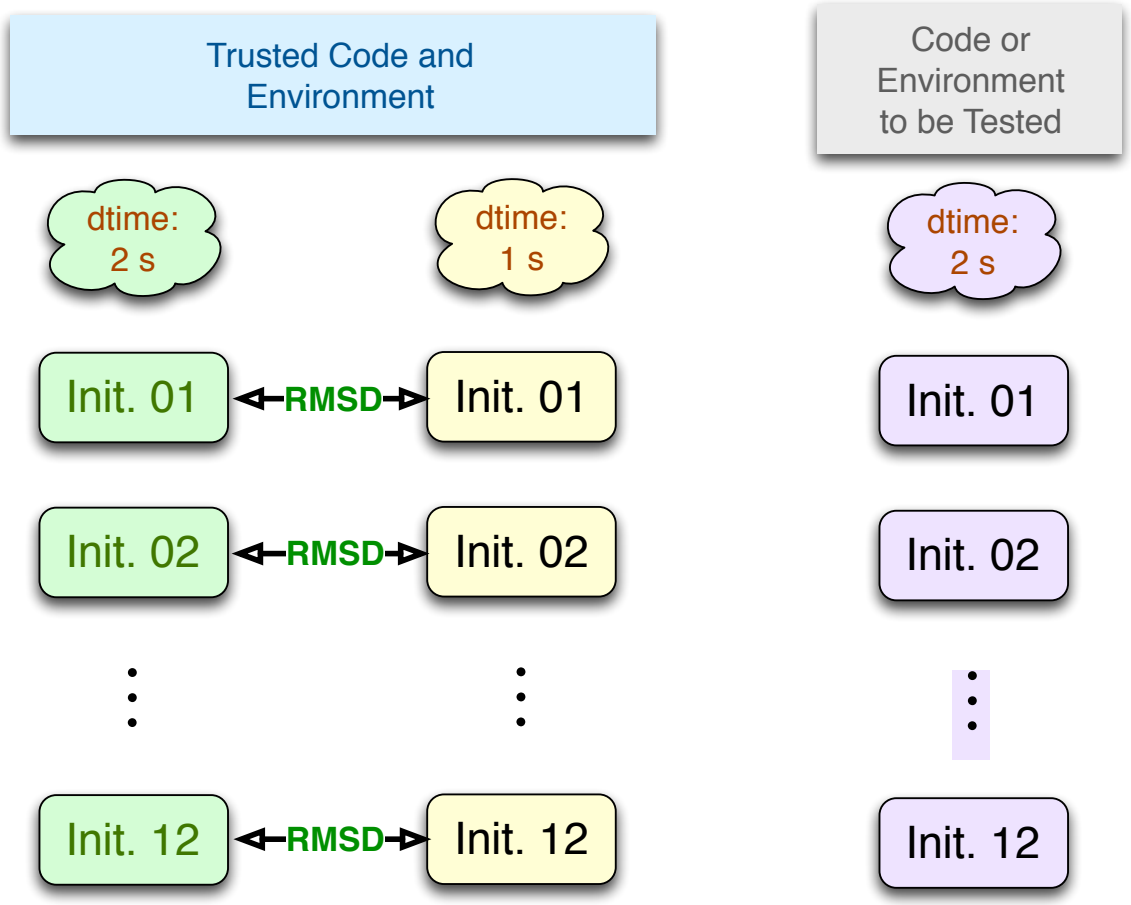
The New Test Method (TSC1.0)



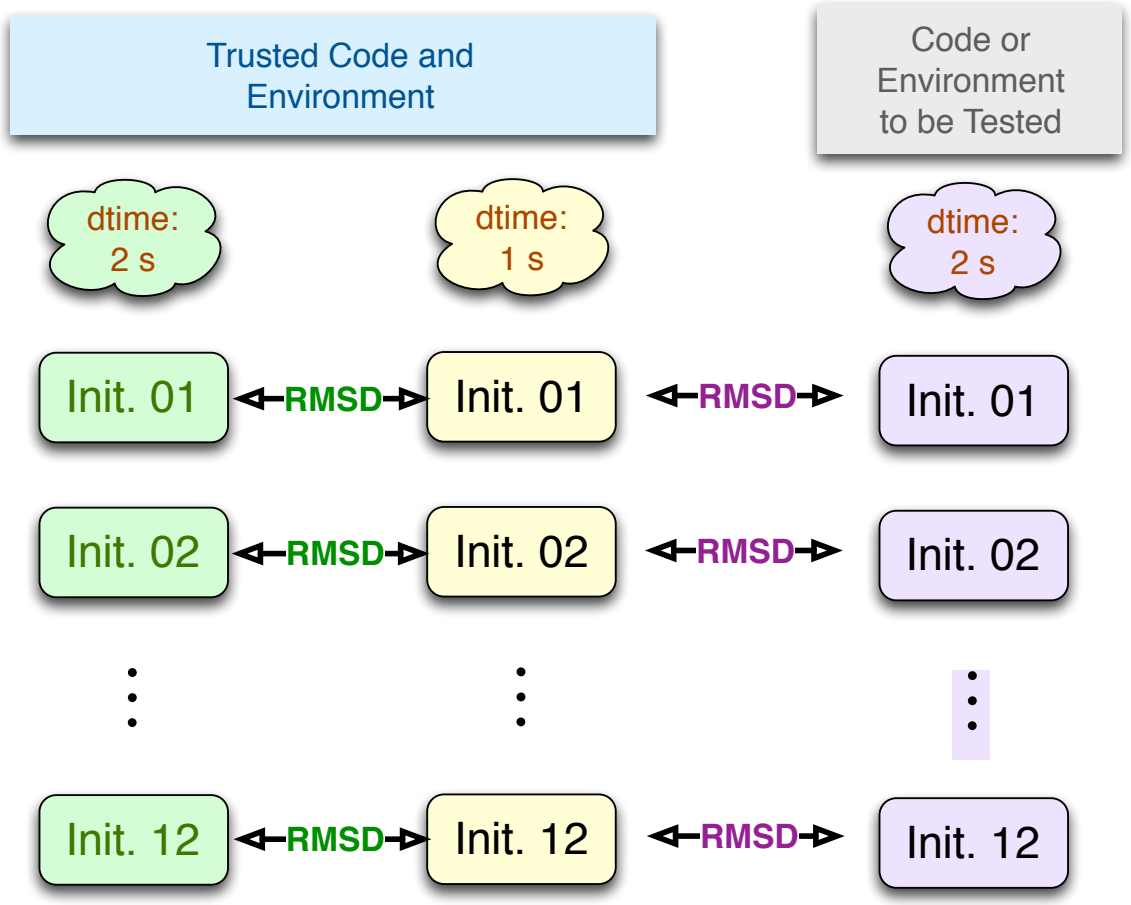
The New Test Method (TSC1.0)



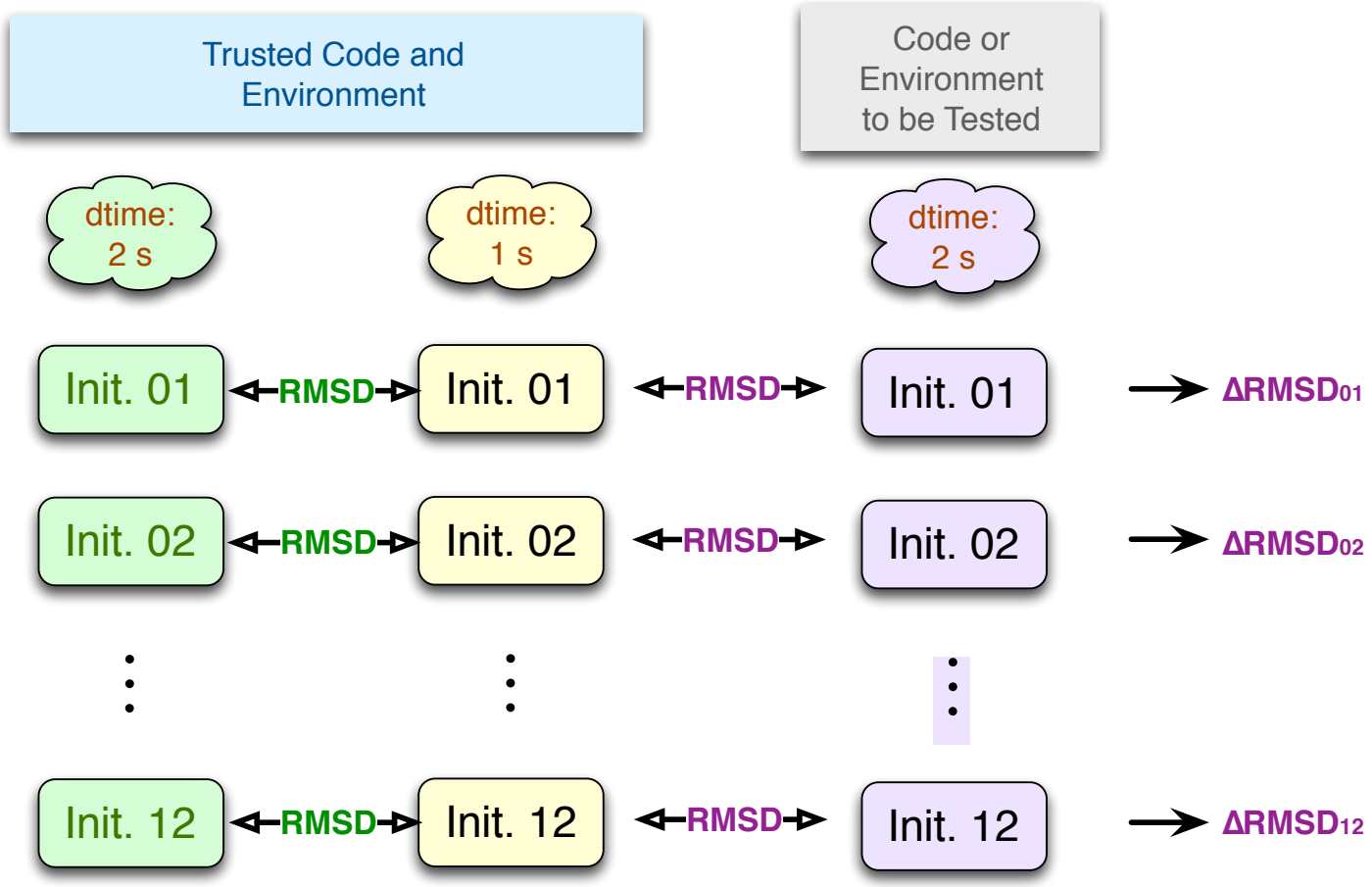
The New Test Method (TSC1.0)



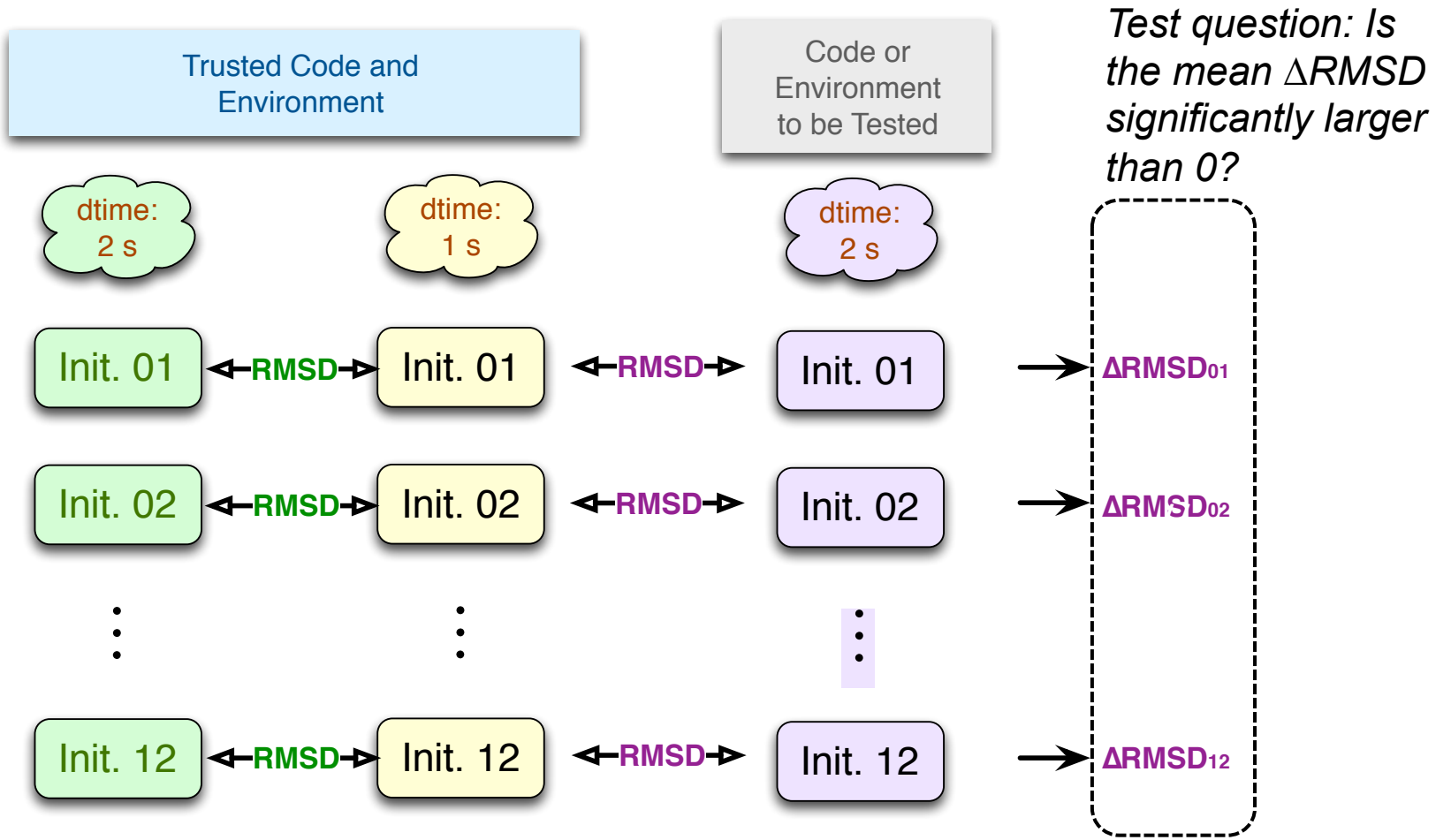
The New Test Method (TSC1.0)



The New Test Method (TSC1.0)

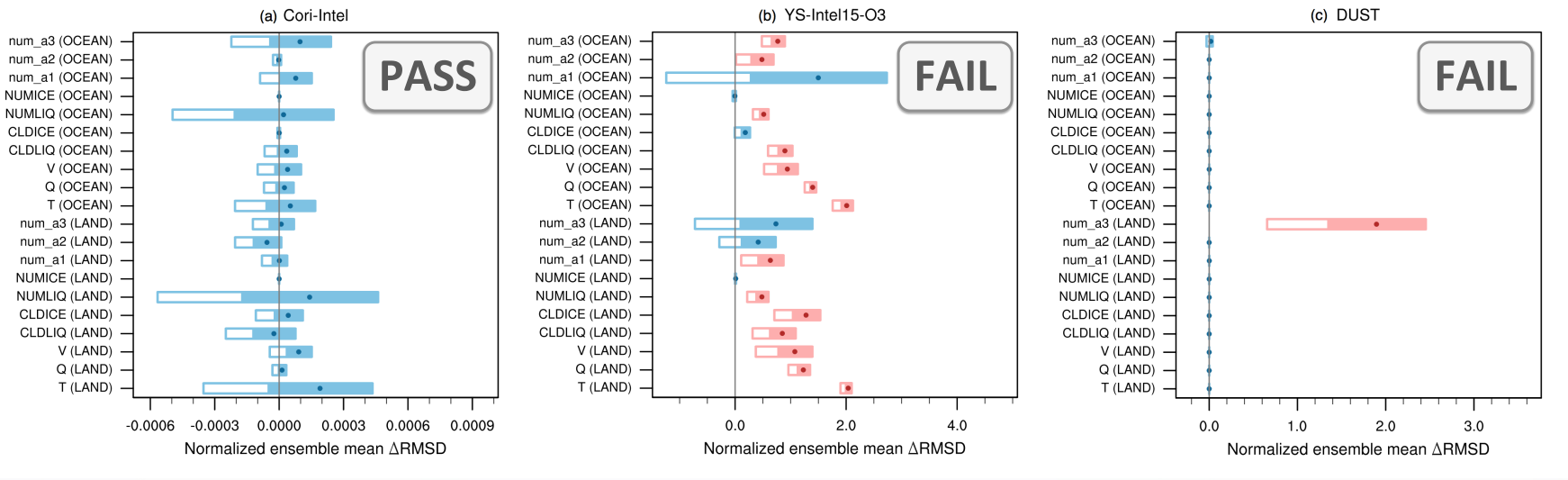


The New Test Method (TSC1.0)



The New Test Method (TSC1.0)

- ΔRMSD is calculated for 10 prognostic variables and for land and ocean separately.
- If the probability of the ΔRMSD being zero is smaller than 0.05% for any of the 20 variables/domains, the test result is a “fail”



But How is This Related to the "Climate-Changing or Non-Climate-Changing" Question?

Category	Experiment	Expected Results	TSC Results
E	Titan-PGI	Pass	Pass
E	Constance-Intel	Pass	Pass
E	Cori-Intel	Pass	Pass
E	Yellowstone O2	Pass	Pass
E	Yellowstone O3	Fail	Fail
P	DUST	Fail	Fail
P	FACTB	Fail	Fail
P	FACTIC	Fail	Fail
P	RH-MIN-LOW	Fail	Fail
P	RH-MIN-HIGH	Fail	Fail
P	CLDFRC-DP	Fail	Fail
P	UW-SH	Fail	Fail
P	CONV-LND	Fail	Fail
P	CONV-OCN	Fail	Fail
P	NU-P	Fail	Fail
P	NU	Fail	Fail
P	QSMALL	Fail	Fail

E: computer/compiler change; P: parameter perturbation.

But How is This Related to the "Climate-Changing or Non-Climate-Changing" Question?

Category	Experiment	Expected Results	TSC Results
E	Titan-PGI	Pass	Pass
E	Constance-Intel	Pass	Pass
E	Cori-Intel	Pass	Pass
E	Yellowstone O2	Pass	Pass
E	Yellowstone O3	Fail	Fail
P	DUST	Fail	Fail
P	FACTB	Fail	Fail
P	FACTIC	Fail	Fail
P	RH-MIN-LOW	Fail	Fail
P	RH-MIN-HIGH	Fail	Fail
P	CLDFRC-DP	Fail	Fail
P	UW-SH	Fail	Fail
P	CONV-LND	Fail	Fail
P	CONV-OCN	Fail	Fail
P	NU-P	Fail	Fail
P	NU	Fail	Fail
P	QSMALL	Fail	Fail

E: computer/compiler change; P: parameter perturbation.

But How is This Related to the "Climate-Changing or Non-Climate-Changing" Question?

From Baker et al. (2015)

Category	Experiment	Expected Results	TSC Results
E	Titan-PGI	Pass	Pass
E	Constance-Intel	Pass	Pass
E	Cori-Intel	Pass	Pass
E	Yellowstone O2	Pass	Pass
E	Yellowstone O3	Fail	Fail
P	DUST	Fail	Fail
P	FACTB	Fail	Fail
P	FACTIC	Fail	Fail
P	RH-MIN-LOW	Fail	Fail
P	RH-MIN-HIGH	Fail	Fail
P	CLDFRC-DP	Fail	Fail
P	UW-SH	Fail	Fail
P	CONV-LND	Fail	Fail
P	CONV-OCN	Fail	Fail
P	NU-P	Fail	Fail
P	NU	Fail	Fail
P	QSMALL	Fail	Fail

E: computer/compiler change; P: parameter perturbation.

But How is This Related to the "Climate-Changing or Non-Climate-Changing" Question?

From Baker et al. (2015)

Category	Experiment	Expected Results	TSC Results
E	Titan-PGI	Pass	Pass
E	Constance-Intel	Pass	Pass
E	Cori-Intel	Pass	Pass
E	Yellowstone O2	Pass	Pass
E	Yellowstone O3	Fail	Fail
P	DUST	Fail	Fail
P	FACTB	Fail	Fail
P	FACTIC	Fail	Fail
P	RH-MIN-LOW	Fail	Fail
P	RH-MIN-HIGH	Fail	Fail
P	CLDFRC-DP	Fail	Fail
P	UW-SH	Fail	Fail
P	CONV-LND	Fail	Fail
P	CONV-OCN	Fail	Fail
P	NU-P	Fail	Fail
P	NU	Fail	Fail
P	QSMALL	Fail	Fail

E: computer/compiler change; P: parameter perturbation.

Our New Method

- Is objective
- Is easy to implement (does not require any code change)
- Is much cheaper than NCAR's climate consistency test
- Can be used for debugging

What's Next

- Try it with the V1 model
- Evaluate the method using more scenarios (e.g. code bugs)
- Further optimization (ensemble size, statistical testing)

We believe the method is applicable to other model components - as long as they have appreciably positive convergence rates.