Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

What level of decadal prediction skill would be achievable given perfect initialisation?

Yiling Liu, Markus Donat, Andrea Taschetto, Paco Doblas-Reyes, Lisa Alexander, Matt England

EUCP GA, Venice 4-7 Feb 2019

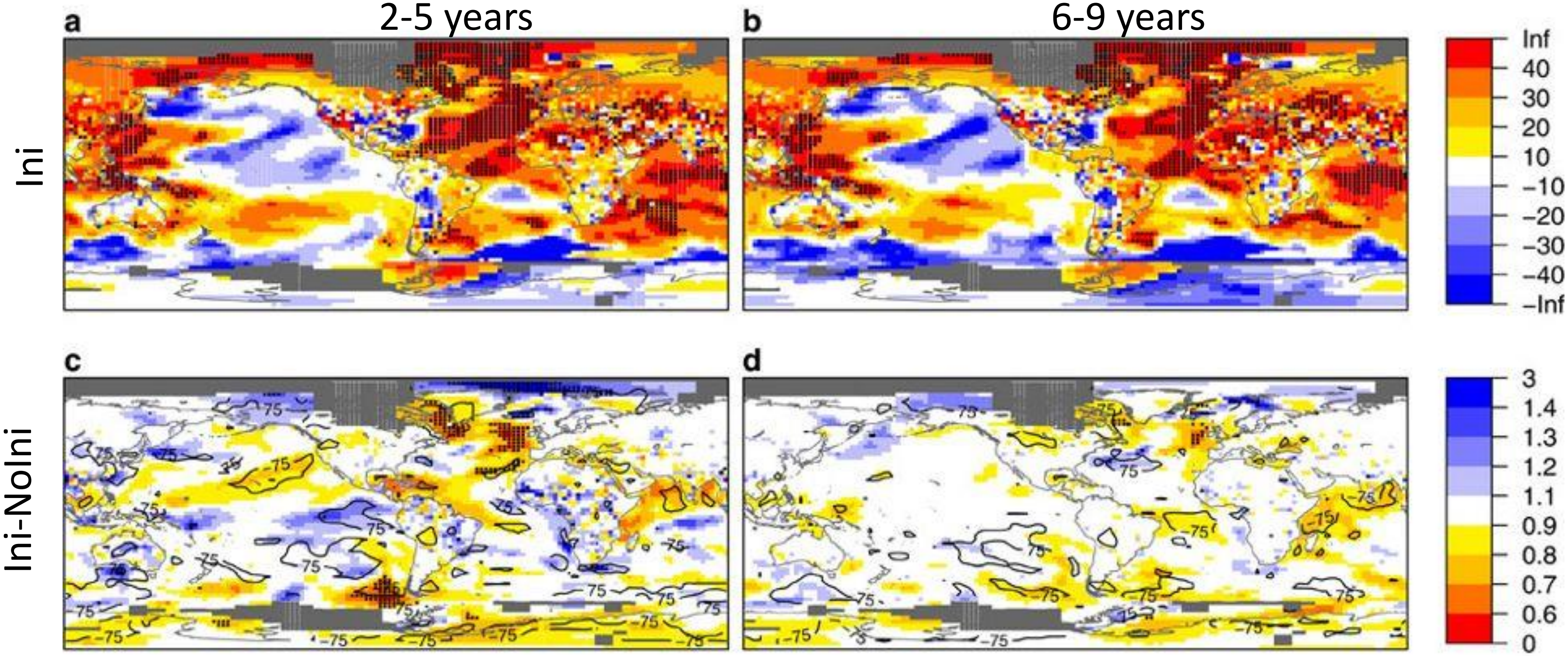
05/02/2019

EUCP is funded by the European Union under Horizon 2020 Grant Agreement no. 776613

EUCP
European Climate Prediction system

Forecast quality in decadal predictions – added value from initialization?

e.g. near-surface temperature in CMIP5 (RMSSS)



[Doblas-Reyes et al 2013]

→ Limited predictability from initialisation versus external forcing due to imperfect initialisation?

→ What level of skill may be achievable?

[Interesting scientific question, but also to important to manage expectations!]

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Perfect-model prediction experiments

How far can the model predict itself, starting from identical (almost) initial conditions?

CESM1, consistent set-up to decadal hindcasts (just replacing real-world observations with a historical simulation, for both initialisation and evaluation):

- decadal simulations started from a historical reference run, 1^{st} Jan each year 1961-2005
- 5 ensemble members (perturbing air temperature with Gaussian noise order of magnitude 10^{-5}K)
- Historical runs (5 members CMIP5 hist) as un-initialised counterpart
- Compare skill for initialised/uninitialised runs, and perfect-model/real-world predictions (real-world hindcasts from CMIP5, *Yeager et al 2012*)
- (focus on near-surface temperature to illustrate the framework)

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Comparing the different prediction experiments

	The initialized perfect model predictions	The initialized real-world decadal hindcasts	The uninitialized climate predictions
Near surface temperature observations		The skill of real-world decadal hindcasts (iii)	The skill of the uninitialized real-world climate predictions (iv)
Perfect model temperature reference	The prediction skill of initialized perfect model (i)		The prediction skill of the uninitialized perfect model (ii)

Skill measures

Forecast accuracy: Mean Squared (error) Skill Score (MSSS)

$$\mathrm{MSSS}(\mathrm{H}, \bar{\mathrm{O}}, \mathrm{O}) = \mathrm{r}_{\mathrm{HO}}^2 - \left[\mathrm{r}_{\mathrm{HO}} - \frac{\mathrm{s}_{\mathrm{H}}}{\mathrm{s}_{\mathrm{O}}}\right]^2 - \left[\frac{\bar{\mathrm{H}} - \bar{\mathrm{O}}}{\mathrm{s}_{\mathrm{O}}}\right]^2$$

r_{HO}: sample correlation between the hindcasts and the observations
$\mathrm{S_H}^2$; $S_O{}^2$: the sample variances of the ensemble mean hindcasts and observations,
$\bar{O} = \sum_{j=1}^{n} O_j$: climatological forecast (where *Oj* represents the observations, or perfect-model reference respectively, over *j* = 1,...,*n* start times), $\bar{\mathrm{H}}$ the mean hindcast

Ensemble dispersion: Logarithmic Ensemble Spread Score (LESS)

$$LESS = \ln\left(\frac{\overline{\sigma_{\hat{H}}^2}}{\sigma_R^2}\right)$$

in which,

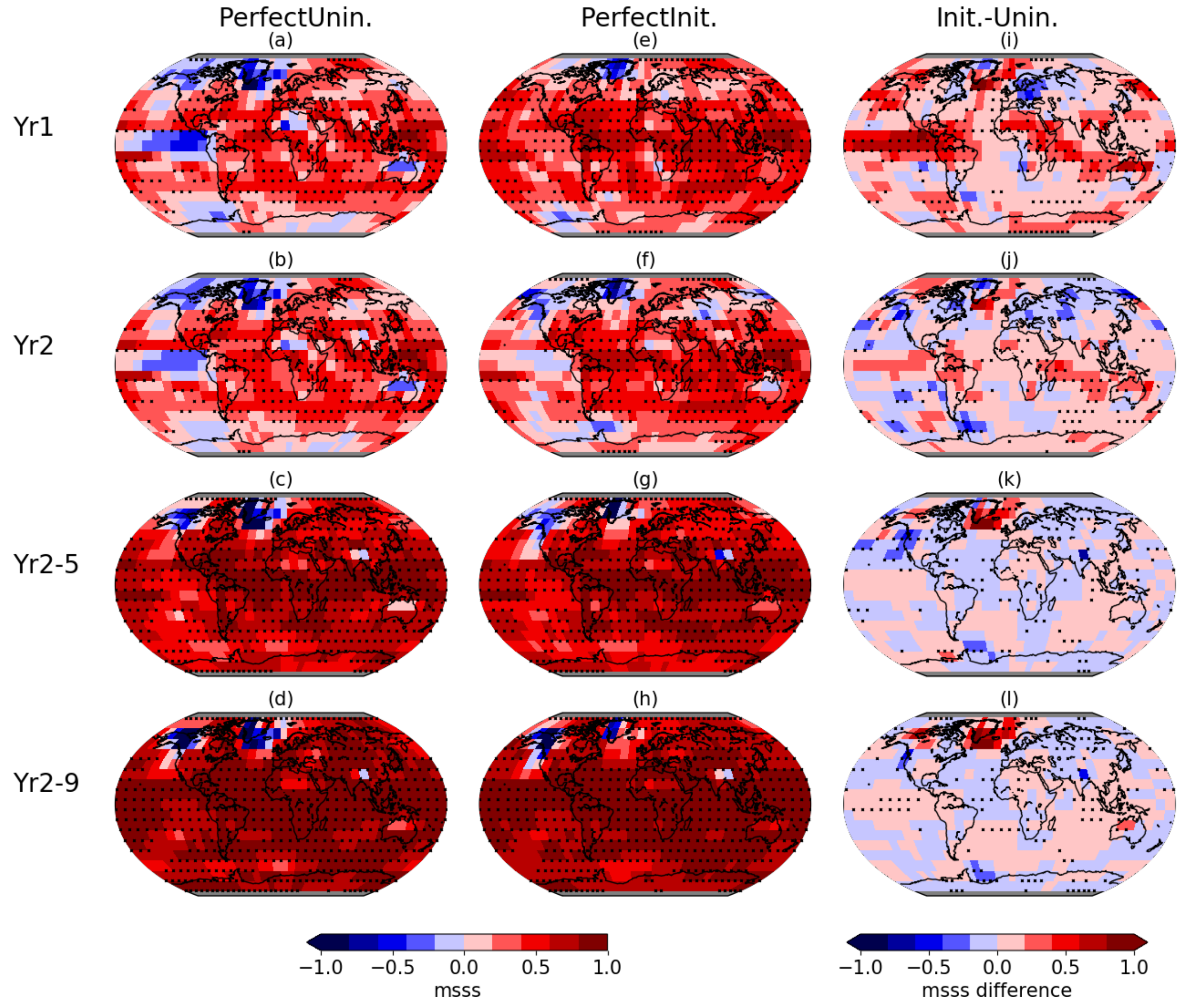
$\overline{\sigma_{\hat{H}}^2} = \frac{1}{n}\sum_{j=1}^{n} \frac{1}{m-1}\sum_{i=1}^{m}(\hat{H}_{ij} - \hat{H}_j)^2$ (average ensemble variance)

$\sigma_R^2 = \frac{1}{n-2}\sum_{j=1}^{n}\left(\hat{H}_j - O_j\right)^2$ (reference mean square error)

LESS <0: the ensemble is under-dispersive (i.e. overconfident)
LESS >0: the ensemble is over-dispersive (i.e. under-confident)

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

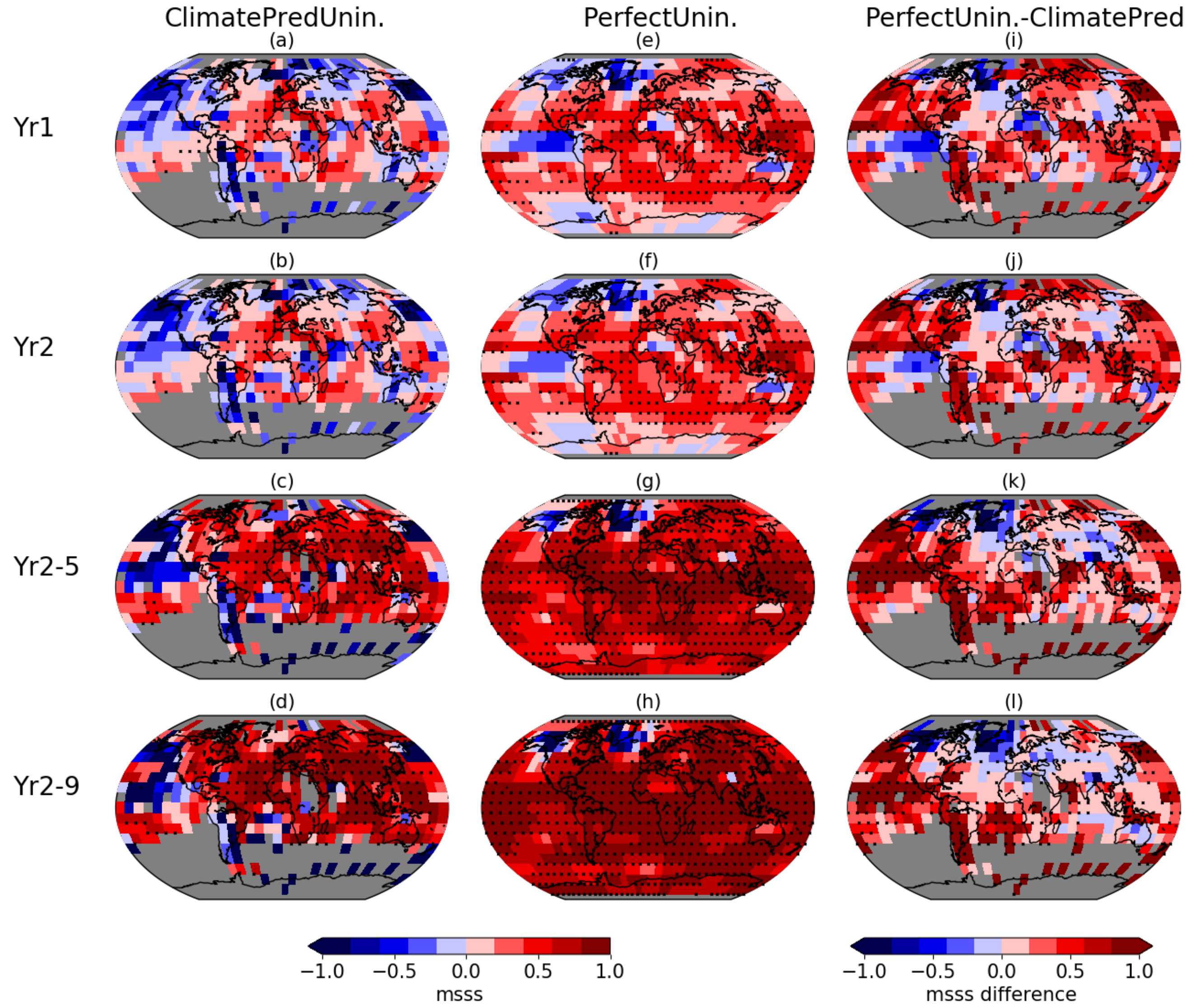
MSSS in initialized and un-initialized perfect-model predictions



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

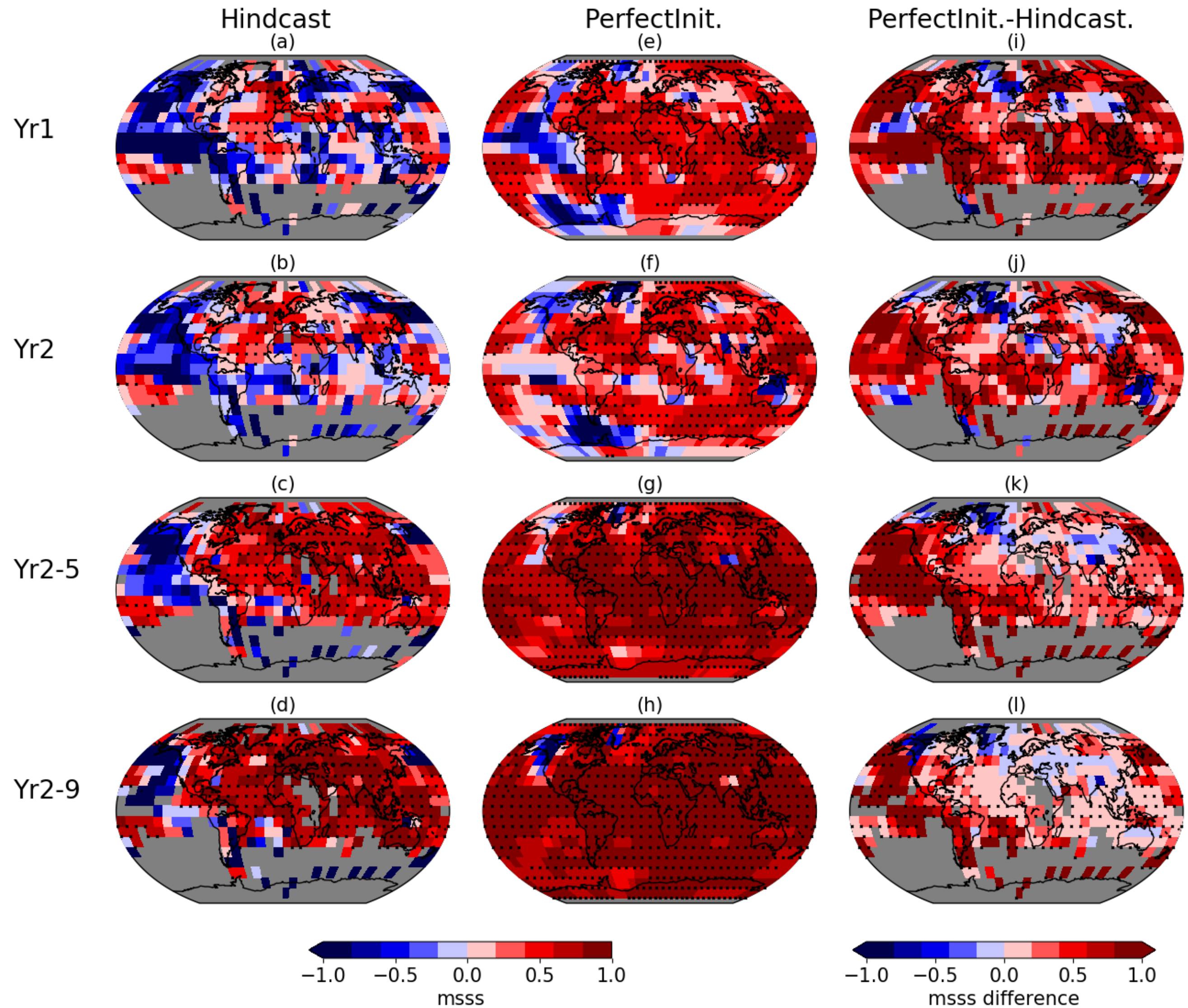
[added value from initialization?]

MSSS un-initialized real-world and perfect-model predictions



[inconsistencies model vs. obs?]

MSSS initialized real-world hindcasts and perfect-model prediction



[total margin of possible improvement?]

Compare areas with increased/decreased skill

	Lead year	Significant improvement ($p<0.1$)	Significant negative difference
Initialized perfect model vs 'uninitialized' perfect model	1	35%	--
	2	15%	5%
	2-5	--	12%
	2-9	13%	9%
Uninitialized perfect model vs uninitialized climate predictions	1	29%	--
	2	27%	--
	2-5	45%	--
	2-9	44%	5%
Initialized perfect model vs decadal hindcasts	1	48%	--
	2	29%	--
	2-5	61%	--
	2-9	54%	--

Effect of initialisation

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Compare areas with increased/decreased skill

	Lead year	Significant improvement ($p<0.1$)	Significant negative difference
Initialized perfect model vs 'uninitialized' perfect model	1	35%	--
	2	15%	5%
	2-5	--	12%
	2-9	13%	9%
Uninitialized perfect model vs uninitialized climate predictions	1	29%	--
	2	27%	--
	2-5	45%	--
	2-9	44%	5%
Initialized perfect model vs decadal hindcasts	1	48%	--
	2	29%	--
	2-5	61%	--
	2-9	54%	--

Inconsistencies
Model - observations

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

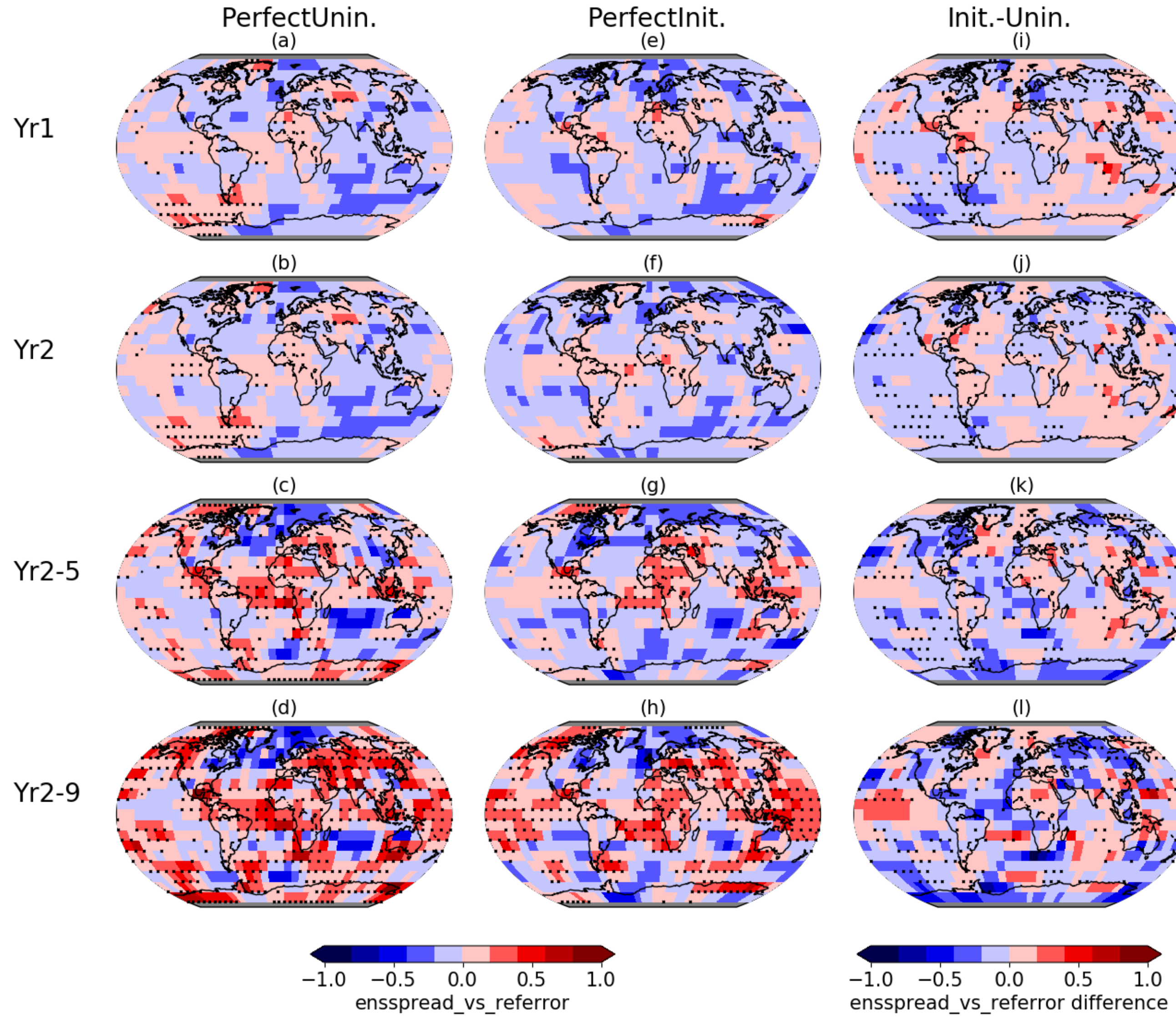
Compare areas with increased/decreased skill

	Lead year	Significant improvement ($p<0.1$)	Significant negative difference
Initialized perfect model vs 'uninitialized' perfect model	1	35%	--
	2	15%	5%
	2-5	--	12%
	2-9	13%	9%
Uninitialized perfect model vs uninitialized climate predictions	1	29%	--
	2	27%	--
	2-5	45%	--
	2-9	44%	5%
Initialized perfect model vs decadal hindcasts	1	48%	--
	2	29%	--
	2-5	61%	--
	2-9	54%	--

Overall possible improvement
If perfect-model skill was achievable

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

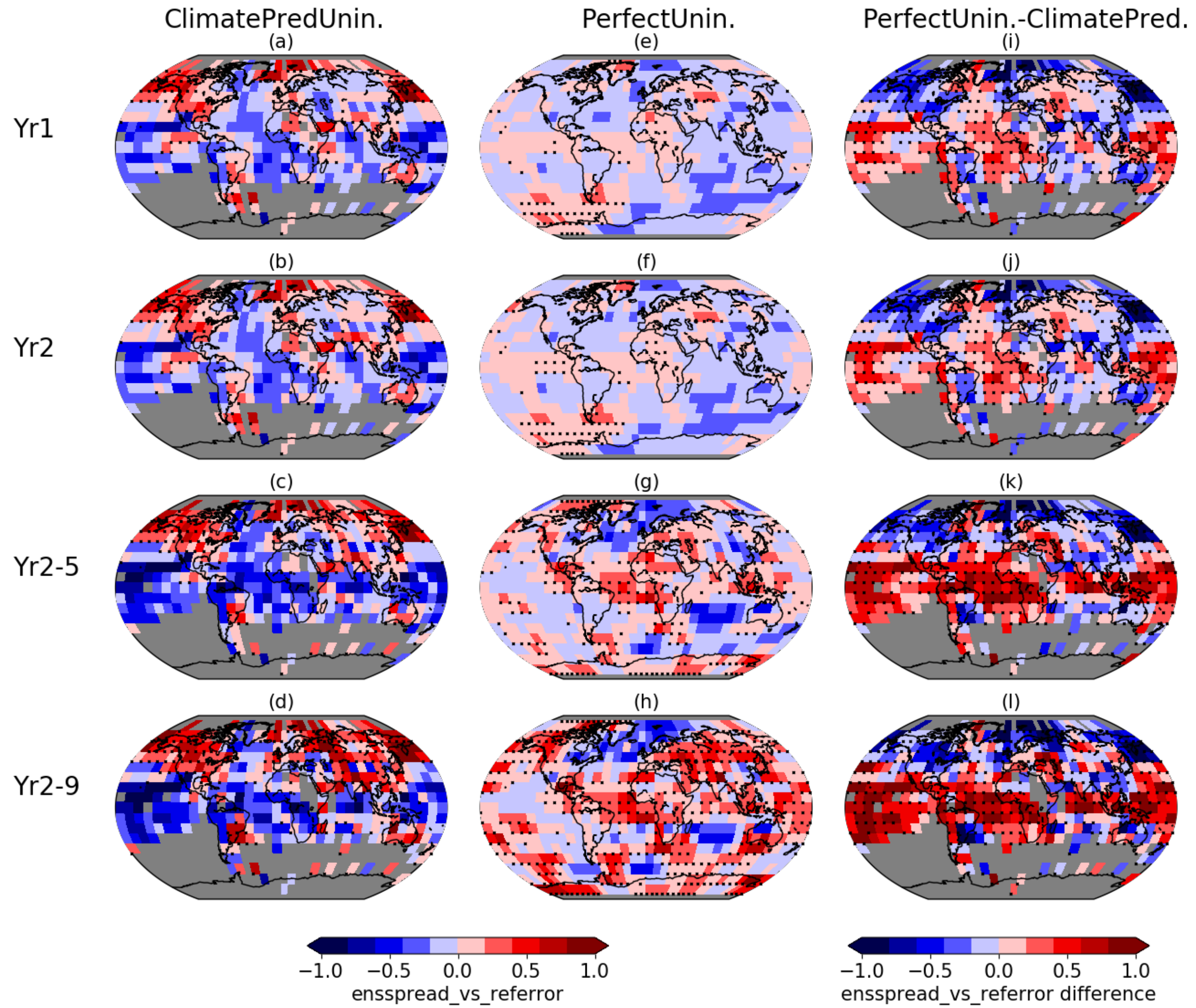
LESS in initialized and un-initialized perfect-model predictions



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

[no clear effect from initialization!]

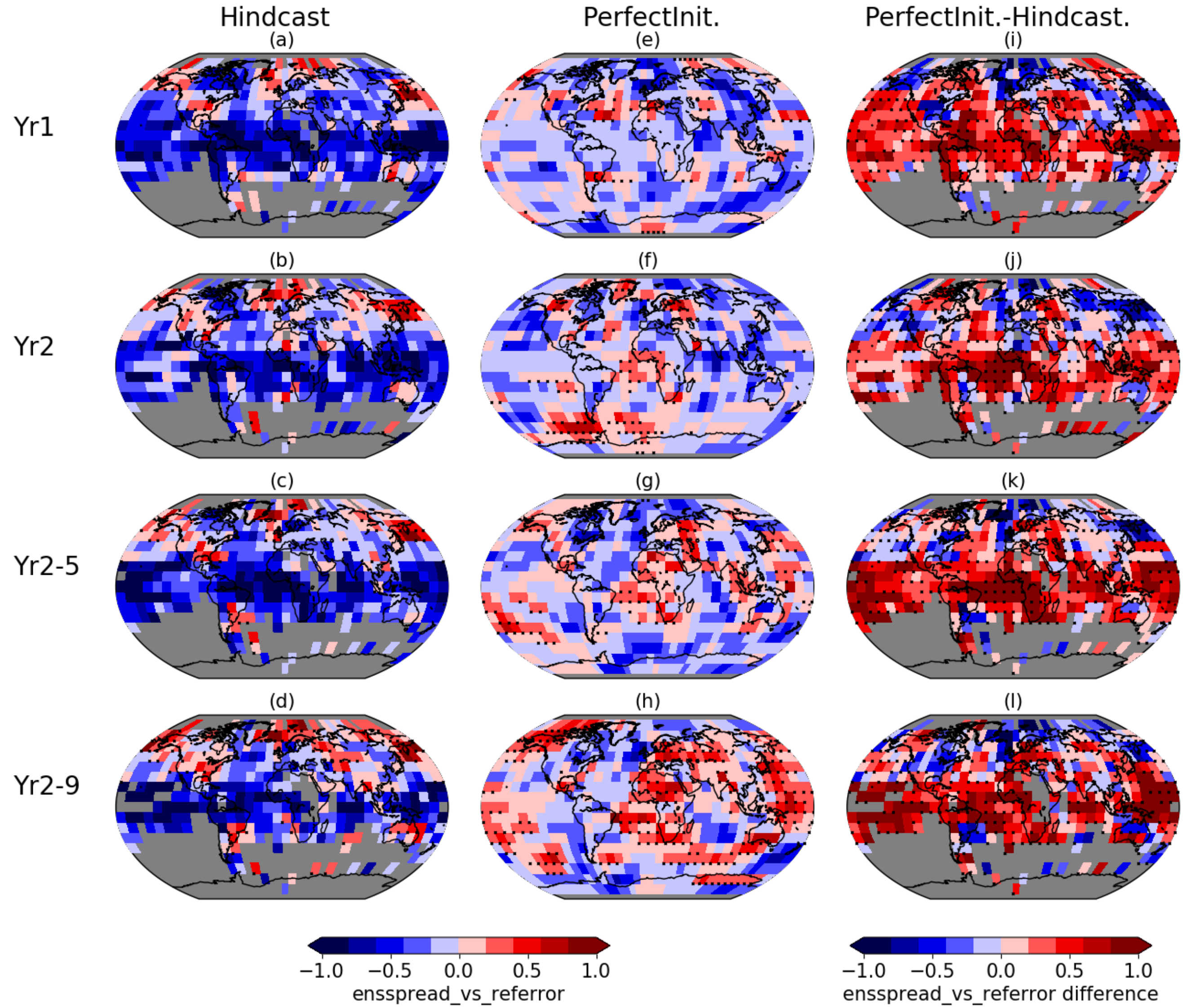
LESS un-initialized real-world and perfect-model predictions



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

[inconsistencies model vs. obs?]

LESS initialized real-world hindcasts and perfect-model prediction



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

[pattern/differences very similar between initialized/un-initialized]

Summary + Concluding thoughts

- perfect-model experiments can be useful to determine the limits of achievable prediction skill (illustrated for tas, but similarly applicable to other variables…)
- Helps to better understand predictability (initialisation versus external forcing), and to manage expectations
- Added value from initialisation until ~2-3 years forecast time
- (ideal) initialisation does not appear to affect ensemble dispersion

- Run similar experiments with different models to understand in how far predictability patterns are model-dependent

→ run such perfect-model predictions complementary to DCPP-A real-world hindcasts??!

[Caveat: predictability of real world may be different to predictability within model missing key processes? Improved models more/less predictable?
→ Perfect-model benchmark informs us what skill is achievable with our existing models (the same used to make predictions!) given ideal initialisation]

(Larger ensemble sizes to e.g. better estimate robustness of skill differences, reliability, etc.)

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Thank you

Liu, Donat, Taschetto, Doblas-Reyes, Alexander, England: A framework to determine the limits of achievable skill for interannual to decadal climate predictions, *JGR-Atmosphere* (revised)

markus.donat@bsc.es

EUCP is funded by the European Union under Horizon 2020 Grant Agreement no. 776613

EUCP
European Climate Prediction system