Expert perspectives on the the *practice* of decadal prediction: from CMIP5 and beyond CMIP6

Alicia R. Karspeck, PhD
Jupiter Technologies

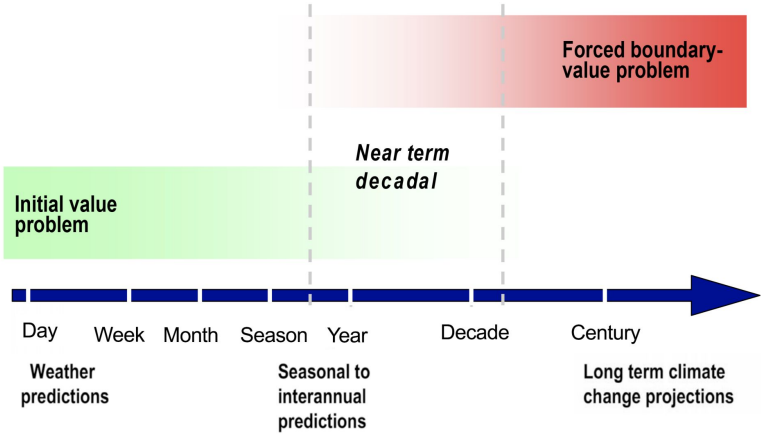
Second International Conference on Seasonal to Decadal Prediction (S2D)

Boulder, CO

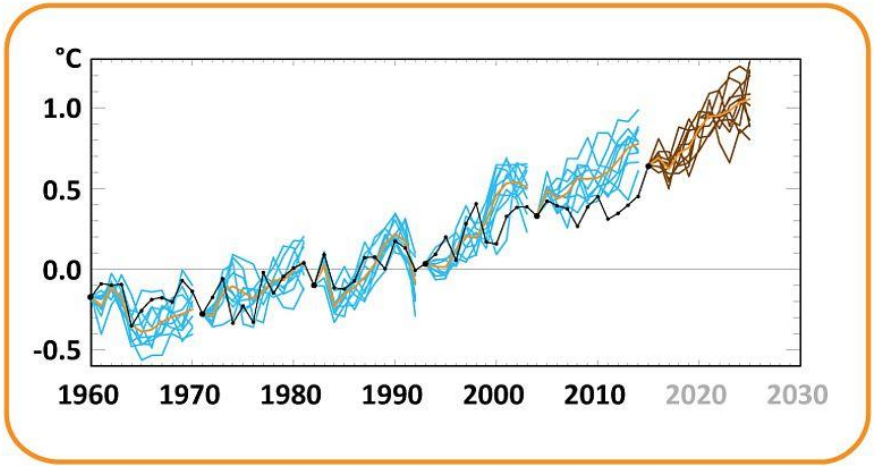
September 18-21, 2018

Thanks to:
Tom Delworth, Doug Smith, Bill Merryfield, Francisco Doblas-Reyes, Wolfgang Mueller, Masahide Kimoto. Noel Keenlyside, Alessio Bellucci, Ed Schneider. Emilia Sanchez, Steve Yeager, Gokhan Danabasoglu, Juliette Mignot, Bo Wu

Decadal climate prediction



Time scale



Schematic representation of retrospective forecasts

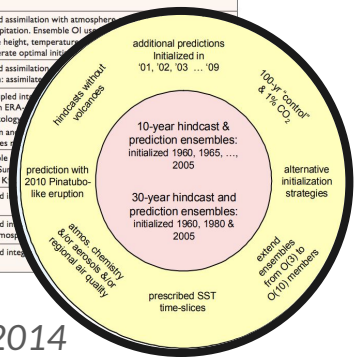
From: https://www.fona-miklip.de/decadal-forecast/background-information/how-are-decadal-climate-predictions-evaluated/

JUPITER

Circa 2012 (CMIP5) to current (CMIP6)

Table 1. (right) Brief summaries of initialization strategies used by (left) modeling groups for the CMIP5 decadal prediction experiments. Acronyms in this table are defined in Table 2.

Model/modeling center	Initialization method
BCC-CM1.1 BCC, China	Full-field initialization. Coupled model integration with ocean *T* nudged to SODA ocean reanalysis product above 1500 m.
CanCM4 CCCma, Canada	Full-field initialization. Atmosphere: assimilate ERA-40 and Interim ECMWF Re-Analysis (ERA-Interim) with nudging to observed SSTs. Ocean: offline assimilation of SODA and GODAS subsurface ocean *T*, *S* adjusted to preserve model *T*–*S* relationship
CCSM4 NCAR, United States	Full-field initialization: (hd-i1) Ocean hindcast forced with CORE2 atmospheric dataset. (da-i2) Loosely coupled ocean–atmosphere Ensemble Kalman Filter (EnKF, see Table 2) analysis. Atmosphere: assimilates raw atmospheric observations, forced with observed SST dataset. Ocean: assimilates observations of subsurface *T* and *S*, forced with atmospheric EnKF analysis.
CFSv2–2011 NCEP, United States	Full-field initialization. Coupled (atmosphere, ocean, ice, land) three-dimensional variational data assimilation (3DVAR) using the NCEP CFSR. Atmosphere constrained by raw observations, ocean constrained by observed *T* and *S*, sea ice constrained by a combination of gridded and satellite-derived products. Additional nudging of ocean surface temperatures to SST reanalysis products (HADISST and Reynolds SST).
CFSv2–2011 COLA, United States	Full-field initialization. Ocean initialized using the NEMOVAR ocean reanalysis interpolated to the ocean model grid. Atmosphere, sea ice, and land initialized from CFSR reanalysis.
CMCC-CM CMCC, Italy	Full-field initialization. Atmosphere: uninitialized twentieth-century coupled model simulations. Ocean: three realizations of CMCC-INGV ocean synthesis of *T* and *S*.
CNRM-CM5 CNRM-CERFACS (France)	Full-field initialization. Coupled model integration with ocean *T* and *S* nudged to NEMOVAR ocean reanalysis product (NEMOVAR is a multivariate 3DVAR data assimilation of *T* and *S* observations into the NEMO ocean model).
EC-Earth (consortium)	Full-field initialization. Atmosphere and land: initialized from ERA-40 and ERA-Interim. Ocean: NEMOVAR ocean reanalysis product. Sea ice initial conditions come from a LIM2 run forced with the DFS4.3 atmospheric forcing.
FGOALS-g2 LASG-CESS, China FGOALS-s2 LASG-IAP, China	Full-field initialization. Atmosphere/land: none. Ocean: nudging to SST, *T*, and *S* with dynamic bias correction. Sea ice: none. Ensembles: perturbed atmosphere/ocean/land/sea ice with different initial time.
GEOS-5 NASA-GMAO, United States	Full-field initialization. Coupled assimilation with atmospheric reanalyses and observed precipitation. Ensemble OI us… subsurface *T* and *S* and surface height, temperatu… breeding method used to generate optimal initi…
GFDL-CM2.1 NOAA-GFDL (United States)	Full-field initialization. Coupled assimilation … atmospheric reanalysis. Ocean: assimilat…
HadCM3 Met Office Hadley Centre, United Kingdom	(i1) Anomaly initialization: coupled int… and HadCM3 climatology from ERA-… anomalies and HadCM3 climatolog… (i2) Full-field initialization: as in an… integration relaxed to full values …
IPSL-CMSA-LR IPSL (France)	Anomaly initialization. Ensemble … Extended Reconstructed Sea Sur… (nudging strength: ~40 W m⁻² K…
MIROC4h, MIROC5 MIROC, Japan	Anomaly initialization. Coupled i… dataset.
MPI-ESM-LR, MPI-ESM-MR MPI-M, Germany	Anomaly initialization. Coupled in… hindcast forced with NCEP atmos…
MRI-CGCM3, MRI, Japan	Anomaly initialization. Coupled integ… dataset.



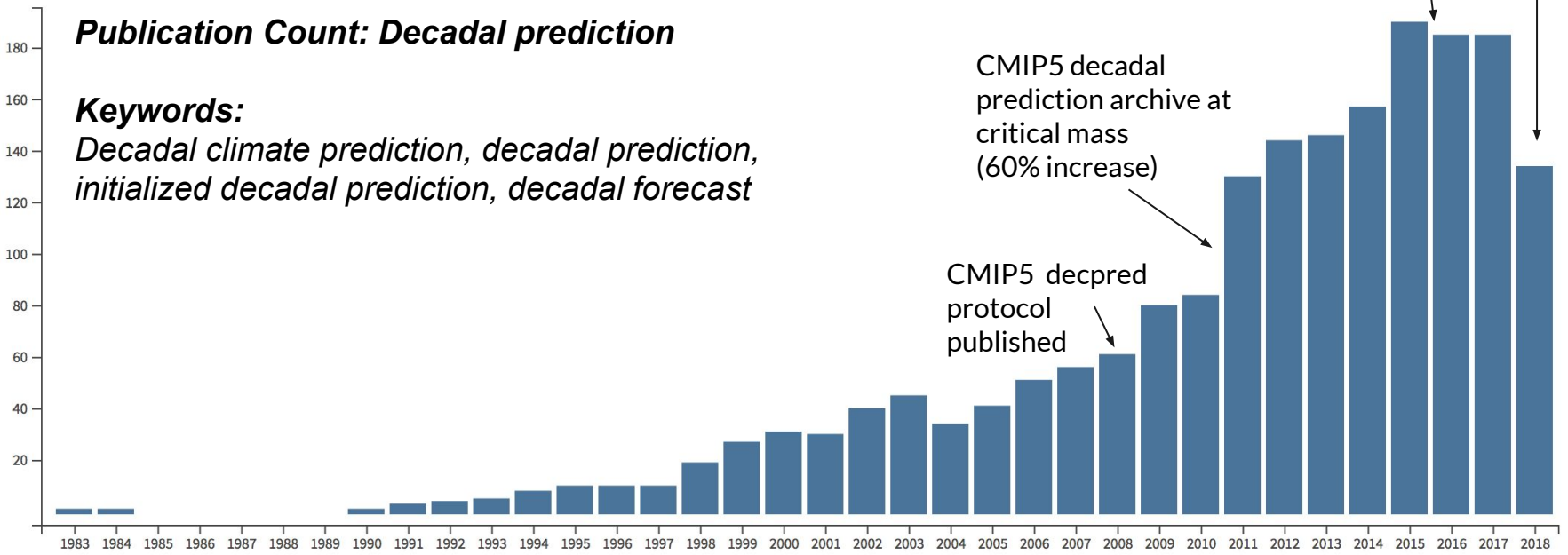
Meehl et al., BAMS 2014

Taylor et al, 2009

- CMIP5 represented one of the first attempts at a coordinated multi-model initialized decadal forecasting experiment.
- ~16 groups participated in CMIP5 decadal prediction experiments
- The at the core of these experiments was a suite of retrospective forecasts initialized every 5 years from 1960 to present.
- This protocol was published in 2009
- It galvanized efforts to evaluate and improve decadal prediction systems, mechanisms, and practices.

JUPITER

Decadal Prediction through time



Publication count from Web of Science

Time period of interest in this survey

JUPITER

CMIP6 Decadal Prediction protocol

Table 1. DCPP experiments.

	Expmt	experiment_id	Tier	Years	Description
Component A: Decadal Hindcasts	A1	dcppA-hindcast	1	3000	Five-year hindcasts every year from 1960. Note that the first forecast year is 1961 from initialization toward the end of 1960.
	A2.1		2	3000	Extend A1 hindcast duration to 10 years
	A2.2	dcppA-historical	2	1700	Ensemble of uninitialized historical/future simulations
	A2.3	dcppA-assim	2	(60-600)	Ensemble of "assimilation" run(s) (if available). These are simulations used to incorporate observation-based data into the model in order to generate initial conditions for hindcasts. They parallel the historical simulations and use the same forcing. The number of years depends on the number of independent assimilation runs.
	A3.1	dcppA-hindcast	3	300m	Increase ensemble size by *m* for A1
	A3.2		3	300m	Increase ensemble size by *m* for A2.1
	A4.1	dcppA-hindcast-niff	4	3000	As A1 but no forcing information from the future (niff) with respect to the hindcast. Forcing from persistence or other estimate.
	A4.2	dcppA-historical-niff	4	3000	As A4.1 but initialized from historical simulations
Component B: Decadal Forecasts	B1	dcppB-forecast	1	50	Ongoing near-real-time forecasts
	B2.1		2	5m	Increase ensemble size by *m* for B1
	B2.2		2	50	Extend forecast duration to 10 years for B1
Component C: Hiatus+	C1.1	dcppC-atl-control	1	250	Idealized Atlantic control
	C1.2	dcppC-amv-pos	1	250	Idealized impact of AMV+
	C1.3	dcppC-amv-neg	1	250	Idealized impact of AMV-
	C1.4	dcppC-pac-control	1	100	Idealized Pacific control
	C1.5	dcppC-ipv-pos	1	100	Idealized impact of IPV+
	C1.6	dcppC-ipv-neg	1	100	Idealized impact of IPV-
	C1.7	dcppC-amv-ExTrop-pos	2	500	Idealized impact of extratropical AMV+

Boer et al. (2016) GMD

- Lessons learned in CMIP5 were incorporated into a revised experimental design for CMIP6
- CMIP6 decpred protocol published in 2016
  - More frequent start dates for retrospective predictions
  - larger number ensembles
  - quasi-operational experimental decadal forecasts
  - Targeted idealized experiments
- CMIP6 experiments are scheduled to commence this year

JUPITER

Expert interviews

- I contacted experts from the original modeling centers that contributed to the CMIP5 DecPred experiments (roughly one per modeling center)

The goal was to hear and synthesize expert narratives/perspectives

on the practice of decadal prediction

as it evolved from CMIP5 to present. …

Expert participants
(14 experts from 13 groups)

Survey Participant	Institution	Model
Tom Delworth	GFDL	SPEAR
Doug Smith	UK Met Office	DePresSys3
Bill Merryfield	Environment Canada	CANESM5
Francisco Doblas-Reyes	Barcelona Super-Computing Centre	EC-EARTH
Wolfgang Mueller	Max Plank Institute/ German Weather Service	MPI-ESM
Masahide Kimoto	University of Tokyo/MIROC	MIROC-ESM
Noel Keenlyside	U. Of Bergen	NorCPM/ESM
Alessio Bellucci	CMCC	CMCC-CM
Ed Schneider	COLA	CFSv2
Emilia Sanchez	CERFACS	CNRM-CM6
Steve Yeager	NCAR	CESM2
Gokhan Danabasoglu	NCAR	CESM2
Juliette Mignot	IPSL	IPSL-CM6
Bo Wu	LASG/IAP	FGOALS

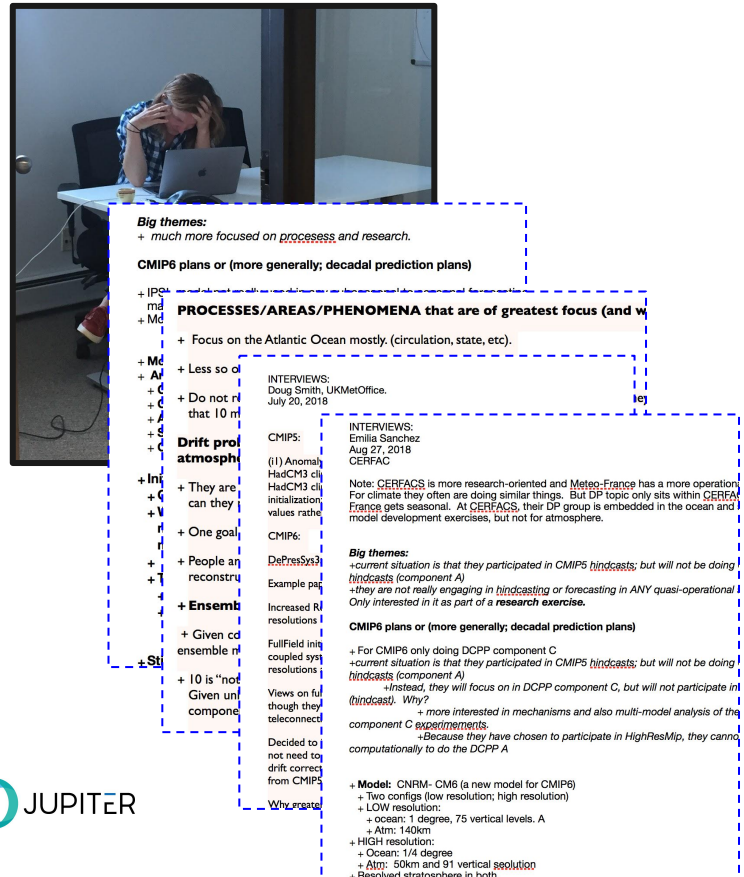
Who is missing here?

existing efforts:
- BCC, China
- GEOS-4 NASA, US

new efforts:
- Swedish Met Office?
- Danish Met Office?
- U. Miami, US?

JUPITER

Format of the interview



- 1-on-1 phone or video-calls -- written notes by AK
- 1-2 interviews per week since late July 2018
- Roughly 1 hour (often longer, sorry!)
- Semi-structured questions (next slide)
  - *original intent was to focus on DP "initialization"*
  - *Scope broadened to include other aspects of the "practice" of decadal prediction*
  - *Similar questions topics posed to each expert.*

Caveat: Good for hearing subjective perspectives, letting experts express their areas of interest and concern, hearing about complexities. ***Not a controlled survey!***

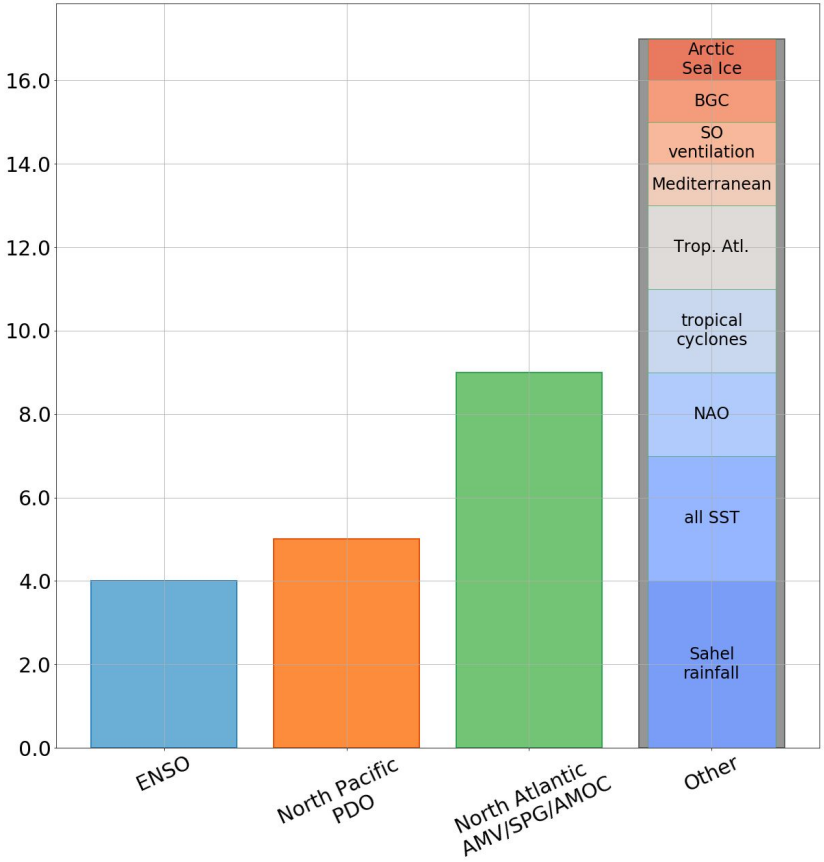
JUPITER

Guiding themes in the expert interviews

Word-cloud from AK's written notes

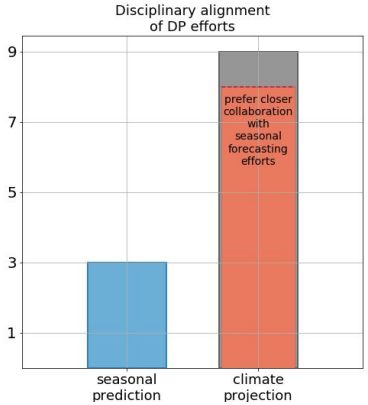
- Scientific focus areas (process and/or geographic)
- Disciplinary alignment (seasonal v. climate projection; research v. operational)
- Horizontal model resolution
- Initialization practice
- Ensemble size
- Drift
- Lessons learned
- Best prospects for the future of DP

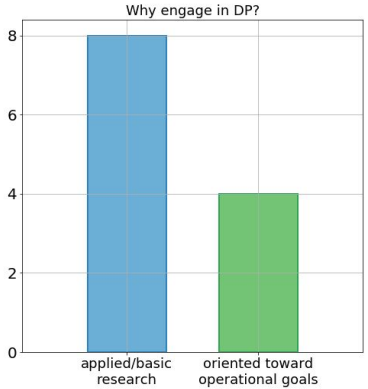
Decadal Prediction Areas of Focus



- Most respondents listed multiple geographic or process areas of interest
- “North Atlantic” (region and associated processes) was the most common answer
- … followed by N. Pacific/PDO
- ENSO was an area of focus for groups that had a natural alignment with seasonal and 2-5 year timescales

Disciplinary alignment

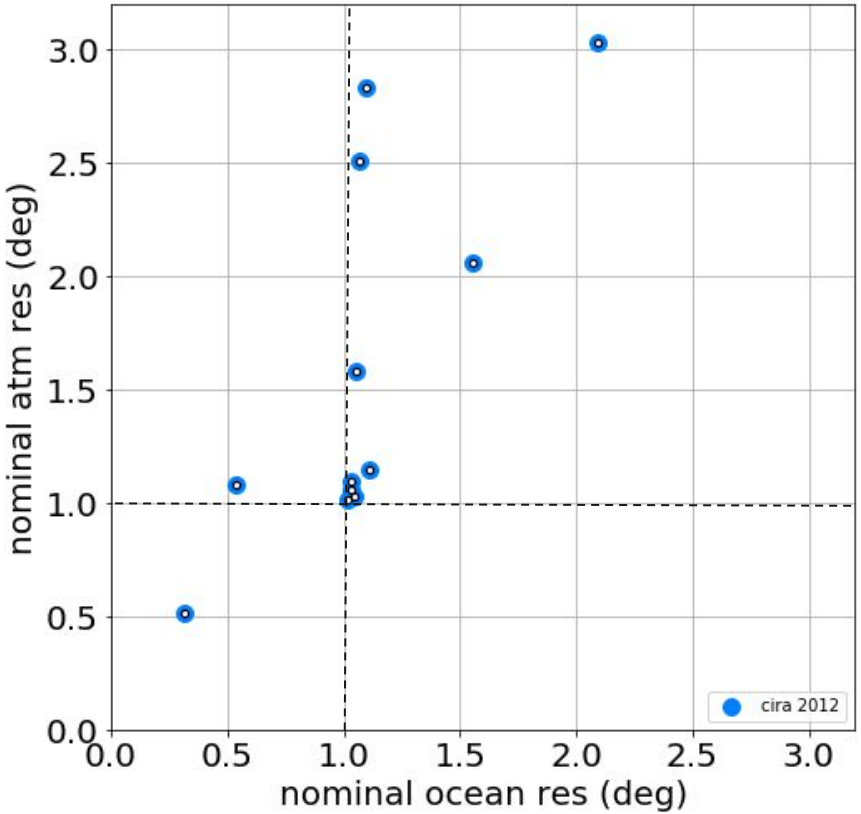




- Majority of groups considered DP a **research exercise** that was primarily aligned with the climate-projection community
- However, nearly everyone mentioned a desire for **closer collaboration with the seasonal prediction community**
  - *"The DP community can learn from the seasonal -- they have greater expertise in the mechanics of forecasting and post-processing"*
  - *A lack of skill in the tropical Pacific can adversely affect forecasts globally*
- Every expert expressed a **need for greater process understanding**, even those groups oriented toward operational goals.
  - *"We don't always know where hindcast skill is coming from"*
  - *"... easy to get lost in the mechanics [of forecasting] and loose sight of what is going on physically"*
- Decadal Prediction Exchange (now ADCP) frequently lauded as an important step toward operationalizing decadal prediction and learning what is useful to end-users.

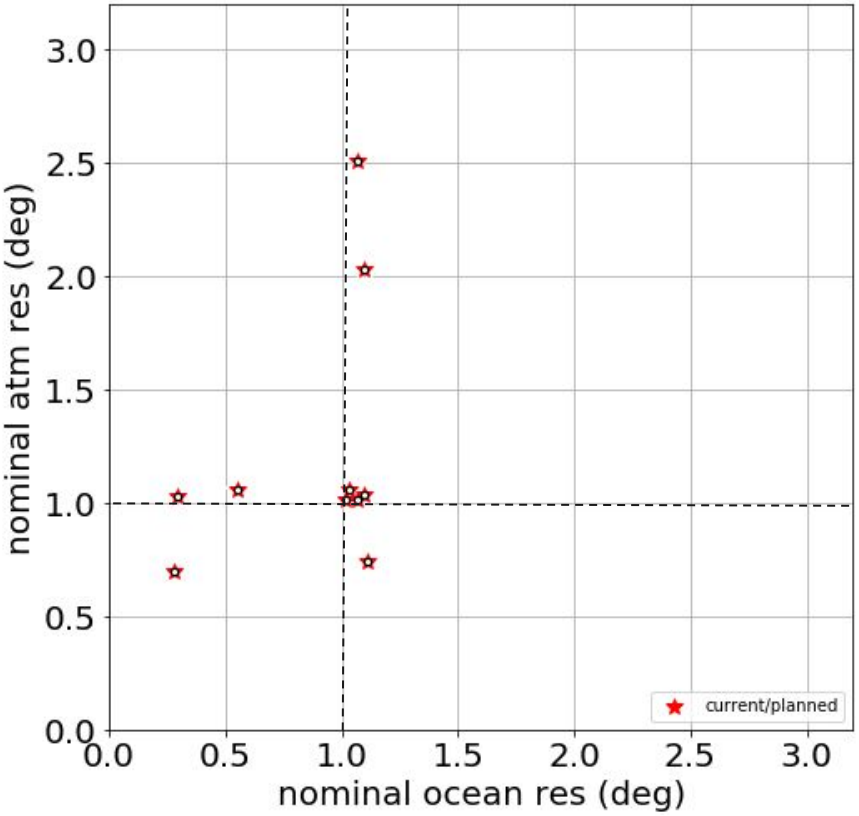
Caveat: Responses are consistent with a set of experts with an Atlantic and CMIP focus

Changes in resolution for DP hindcast/forecast



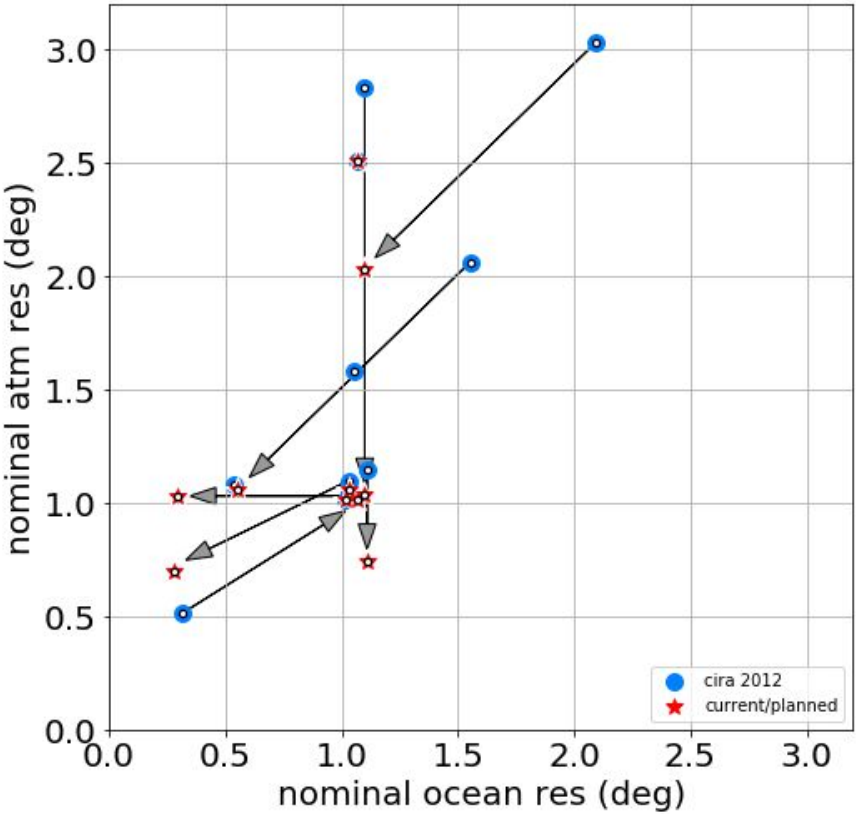
- Circa 2012 (2) groups were using sub 1 deg ocean models
- … both are choosing not to repeat

Changes in resolution for DP hindcast/forecast



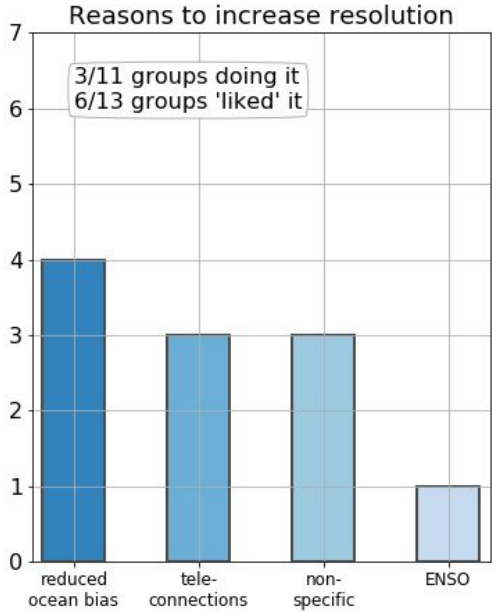
- Circa 2012 (2) groups were using sub 1 deg ocean models
- … both are choosing not to repeat
- Currently 3 (new) groups reported plans to use sub 1 degree models for hindcast DP.
- ... no groups planning on oceans coarser than 1 deg.

Changes in resolution for DP hindcast/forecast



- Circa 2012 (2) groups were using sub 1 deg ocean models
- … both are choosing not to repeat
- Currently 3 (new) groups reported plans to use sub 1 degree models for hindcast DP.
- ... no groups planning on oceans coarser than 1 deg.
- All but (1) group is increasing resolution or remaining the same

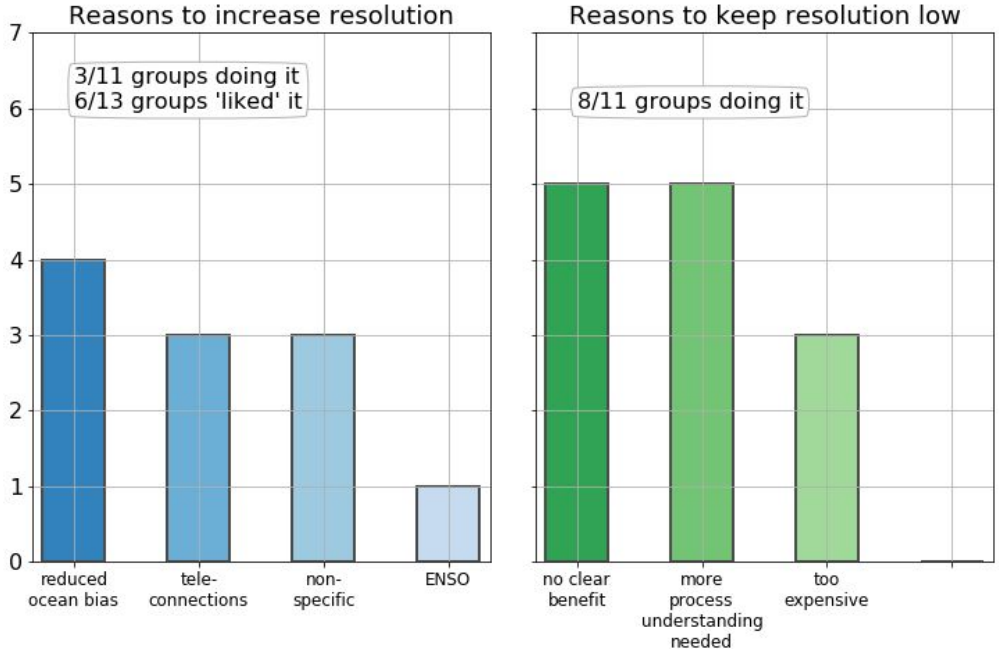
View on increasing resolution … or keeping it low



Why increase resolution?

- Ocean-atm coupling will be more realistic
  → better teleconnections/better skill overland
  → reduced ocean bias
- More realistic GulfStream and NA current
  → reduced ocean bias
- A process-agnostic sense that higher-resolution is where the community should be heading
- Consistency w/ seasonal systems

View on increasing resolution … or keeping it low



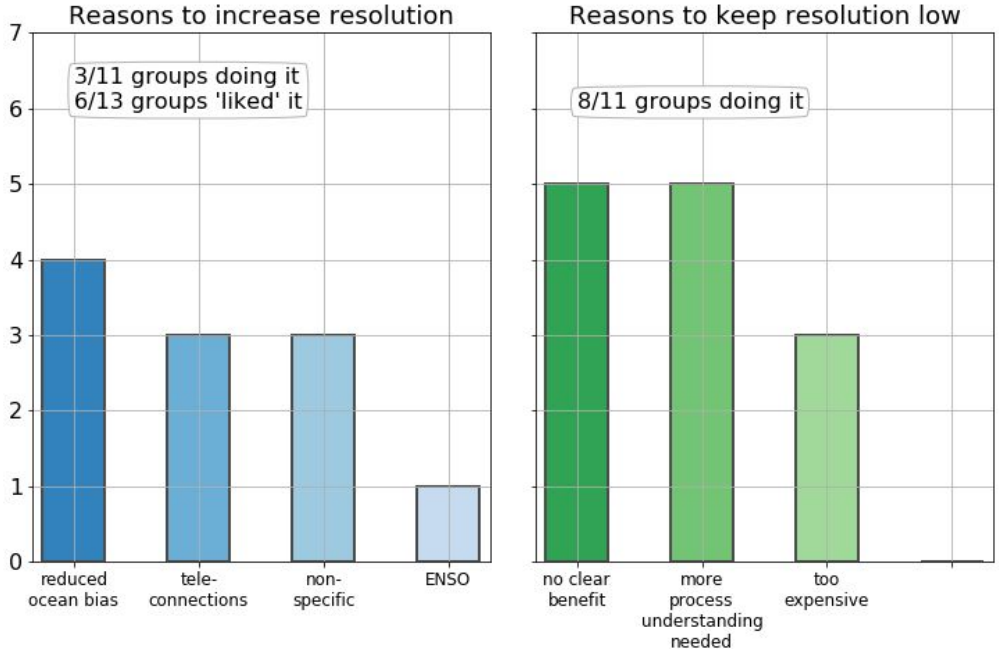
Why increase resolution?

- Ocean-atm coupling will be more realistic
  → better teleconnections/better skill overland
  → reduced ocean bias
- More realistic GulfStream and NA current
  → reduced ocean bias
- A process-agnostic sense that higher-resolution is where the community should be heading
- Consistency w/ seasonal systems

Why not?

- No clear performance benefit
- DP process understanding not mature enough to take on the increased complexity
- Practically impossible (too expensive)

View on increasing resolution … or keeping it low



Why increase resolution?

- Ocean-atm coupling will be more realistic
  → better teleconnections/better skill overland
  → reduced ocean bias
- More realistic GulfStream and NA current
  → reduced ocean bias
- A process-agnostic sense that higher-resolution is where the community should be heading
- Consistency w/ seasonal systems

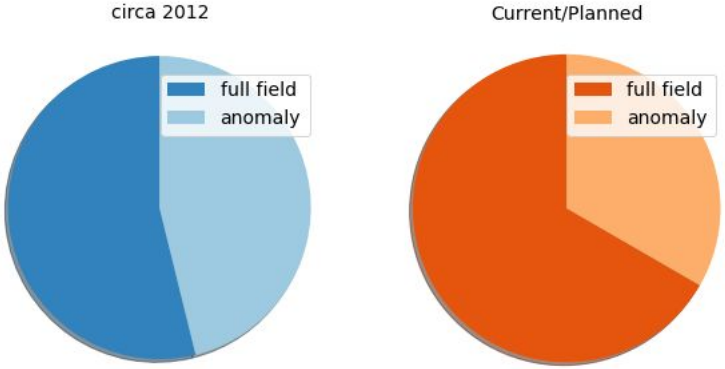
Why not?

- No clear performance benefit
- DP process understanding not mature enough to take on the increased complexity
- Practically impossible (too expensive)

Key points:

+ AK's observation: most experts had a clear "view" on this question
+ Getting a stable climate (TOA radiative balance/'reasonable" AMOC) in higher resolution models is a significant issue.
+ Resolution/model version decisions are often made by development groups separate from the DP science/implementation groups -- DP "piggy-backing", not driving these decisions

Initialization

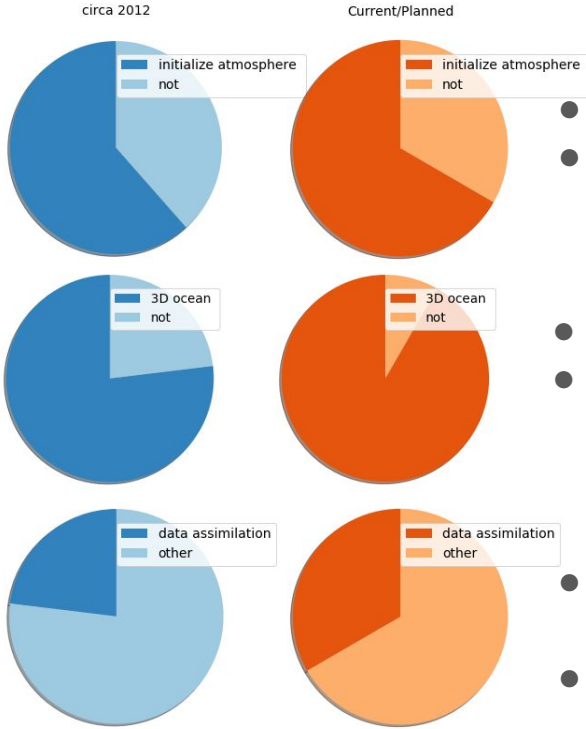


- Questions of **anomaly vs. full-field** of greatest interest to this group
- More groups are doing full-field (as opposed to anomaly) initialization
- Trending up (+2 groups for full-field)

Key points raised about initialization:

- **Most groups reported spending significant time** developing/improving their initialization practice,
- However for groups that tested multiple "working" initialization systems - **quantitative skill metrics did not tend to yield meaningful differences**.
- As a result, decisions to use one method vs. another were typically based on
  - ***Practical factors*:** computational expense, legacy practice, desire for consistency with seasonal prediction, availability of reanalysis datasets
  - ***Avoidance of spurious behavior***: e.g. "the forecast drift was too large to trust results", "the AMOC was not realistic", "triggered El Ninos"
  - ***Epistemic considerations:*** e.g. "anomaly initialization muddles our diagnosis of model bias."

Initialization (cont.)



Initialized the atmosphere?

- More than half of groups are initializing the atmosphere
- … no meaningful change

Initialized the 3-D ocean state?

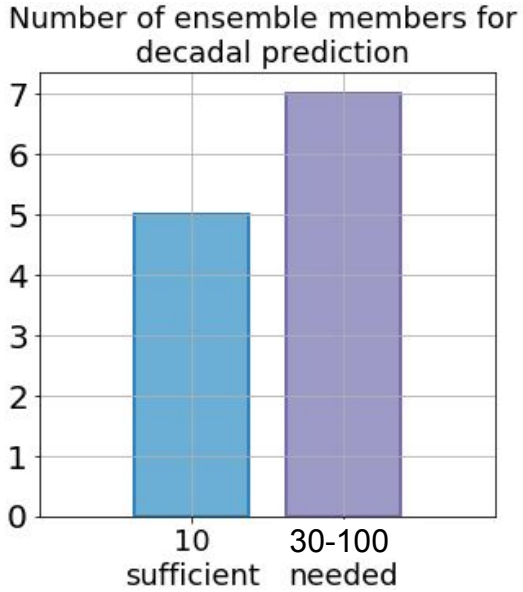
- Most groups are initializing the full 3-D ocean state (T/S)
- …increasing proportion

Data assimilation?

- Very few groups are doing in-house data assimilation to support initialization
- …slightly increasing proportion.

Very few experts thought investing in state-of-the-science DA was worthwhile for DP initialization

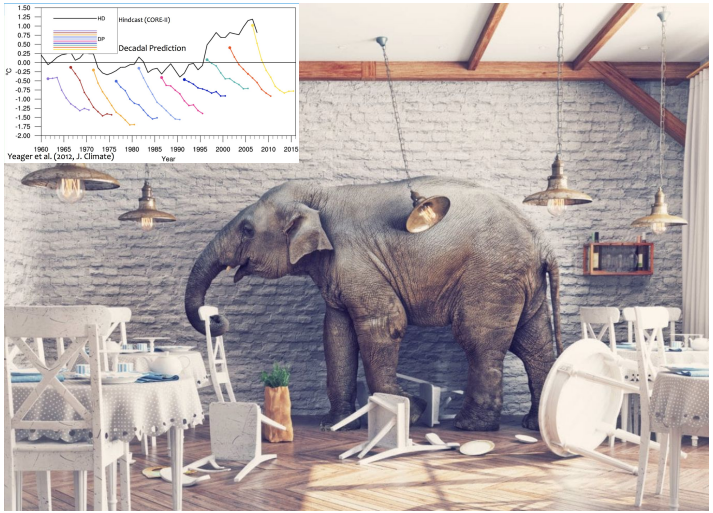
Ensemble sizes



- Many experts noted that the answer is process dependent.
- 30-100 tended be the sizes needed for “impacts” and overland processes requiring the resolution of atmospheric teleconnections
- 10 being “sufficient” for basin-scale ocean processes
- “More is better” was a common sentiment.

- **General sense that annual start dates are sufficien**t, but a few experts “curious” about whether skill was dependent on season of initialization.
- This question i**nspired comments about increasing focus on the 2-5 year time horizon** as a way to afford more ensembles.
- Questions of ensemble size should be addressed in idealized experiments (like DCPP component C)

Considering drift



- **All groups planning to do some form of drift correction**
  - All groups reported plans to follow CMIP protocol for lead-dependent drift correction.
  - 3 groups working on more advanced calibration of forecast (trend correction, spread calibration)
- Many groups noted that drift is present regardless of whether full field or anomaly initialization is used.
- Multiple experts offered that **drift and forecast transients should be studied more systematically** to help understand and communicate model deficiencies
- **A clear point of frustration -- with very little optimism**

"the elephant in the room"

"The dirty laundry of decadal prediction"

" the community knows very little about how to deal with this "

" a necessary evil "

"... so severe that prediction seemed pointless"

"A 'nasty thing' that we sweep under the rug"

Lessons learned

- **There is value in the practice of initialized decadal prediction (7)**
  - "[The practice itself] is necessary to move the science forward"
  - "... forecasting creates a vehicle for observational and modeling communities to engage with one another"
  - "[The practice] exposes model error"
  - The real world may be more predictable than models suggest ("signal to noise paradox")
  - Contrary view: The community engaged in the practice of prediction before the processes understanding was mature enough (1)
- **Initialization practice requires a great deal of attention -- yet does not *easily* yield quantifiable improvement (6)**
- **More ensembles increase skill (4)**
- At this point in the maturity of the science, **process understanding is more important than skill-scores (4)**

AK: This was the most interesting and dynamic part of the interviews, also the hardest to synthesize

Lessons learned (cont)

- **...A huge effort that yields only marginal improvement over climate projection (2)**
- Impacts community wants decadal information, but the science is not mature enough yet (2)
- Increased resolution is “very painful” with “no clear benefit” (2)
- **“Simple” initialization may be good enough (2)**

- **Academic and research labs need to consider what role they want to play in contributing to public-facing climate services (1)**
- **Separation of prediction and model development groups/efforts hinders progress (1)**
- “Drift is massive”(1)
- Aspects of ocean biogeochemistry may be predictable -- could be useful for carbon cycle (1)
- Calibration and post-processing are important (1)

AK: This was the most interesting and dynamic part of the interviews, also the hardest to synthesize

Best prospects for advancing the field of DP

- **Better models (9)**
  - Greater focus on processes
  - More coordinated idealized experiments (e.g. DCPP component C)
  - Use collective understanding of drift to inform model developers
- **More formalization of forecast dissemination activities (5)**
  - e.g Decadal Prediction Exchange

- Larger ensembles and more start dates (2)
- **More interaction with seasonal community(1)**
- **More researchers working on the problem (more eyes, not more compute)(1)**
- More advanced data assimilation (1)
- Higher resolution (1)
- Reduced focus on “gratuitous publication”(1)
- Fewer models (1)
- Greater focus on skill over the continents (1)
- Increased interaction with ocean-reanalysis community (1)

END

Thank you to all the experts for their time

- Because conversational structure was generally fluid, some subjective interpretation (on my part) went into creating syntheses. My intent was always to capture the sentiment of the experts.
    - e.g. if a “lesson learned” or “best prospect” came up while discussing other topics, I attempted reflect it in the way I thought it was intended.
- Not all the topics that we discussed are in this presentation.
- Happy to share original notes (with permission from experts).