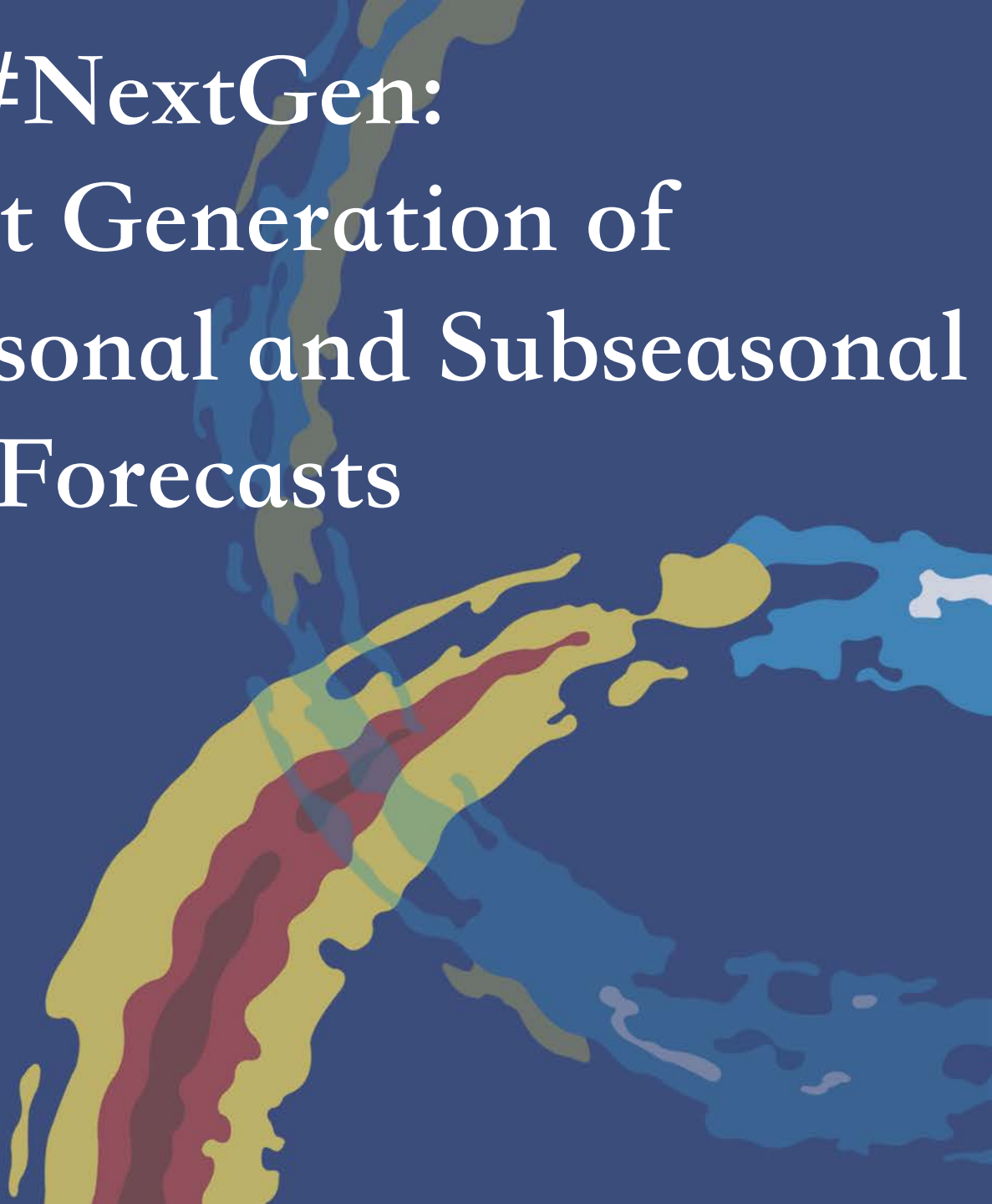
#NextGen: A Next Generation of Regional Seasonal and Subseasonal Forecasts

Á.G. Muñoz
(agmunoz@iri.columbia.edu),
A.W. Robertson,
S.J. Mason,
L. Goddard
D. Pons,
X. Chourio

International Research Institute
for Climate and Society
Earth Institute | Columbia University



Why, What, How and Where?

Why, What, How and Where?

[The name? That was Francesco!]

Why NextGen?

+ The new availability of global ensemble forecast system products has the potential to transform climate forecasting at regional and national levels.

+ Enable these global model outputs to be tailored to local/regional data and user-relevant variables and threshold-exceedances.

+ Routine, automated operational rolling forecasts at national and regional level are now possible, using standard CPT tools

Why NextGen?

+ NMHs in developing countries tend to use SSTs in the Pacific as the only candidate predictor!! The atmosphere-ocean coupling is normally not considered in their empirical forecast approach.

+ Most NMHs in developing countries have not adopted yet an *objective* seasonal (or subseasonal) forecast approach.

Why NextGen?

+ WMO Executive Council, 69th Session (EC69, 2017):
  "…consider the adoption of objective sub-seasonal and seasonal forecasts as an overarching technical strategy, particularly at regional and national levels, promoted through RCOFs, …" *[WMO Guidance Doc on Seasonal Forecasts will be available later this year]*

+ The RCOF Review (2017), recommended
  "RCCs to access digital forecast and hindcast data from the WMO LC-LRFMME and produce an objectively consolidated forecast product combining information of various GPCs-LRF to be used as a first estimate for RCOF discussions."

Why NextGen?

+ Decision makers require tailored information:
  - a wide range of predictands, well beyond rainfall totals and temperatures. E.g., onset, duration and demise of the rainy/monsoon/dry season(s), crop yield, planting *dates*, energy consumption, …
  - IRI's flexible format approach (entire PDF — forget about terciles until they're really needed).

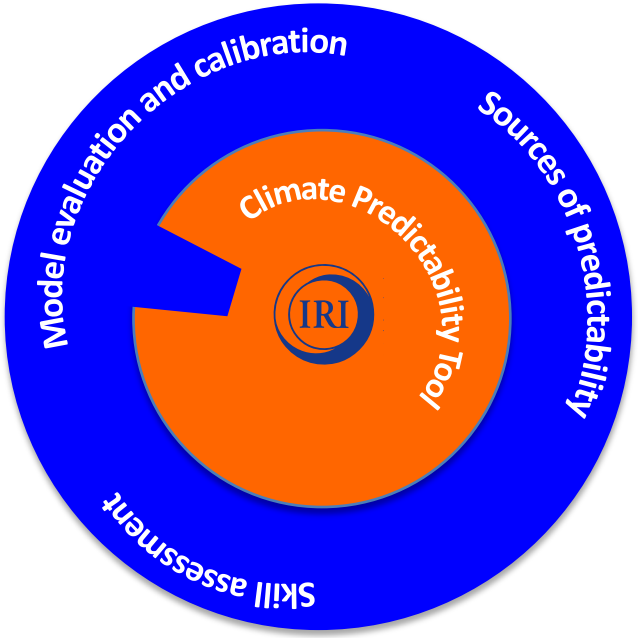
#NextGen

NextGen is a system that involves:

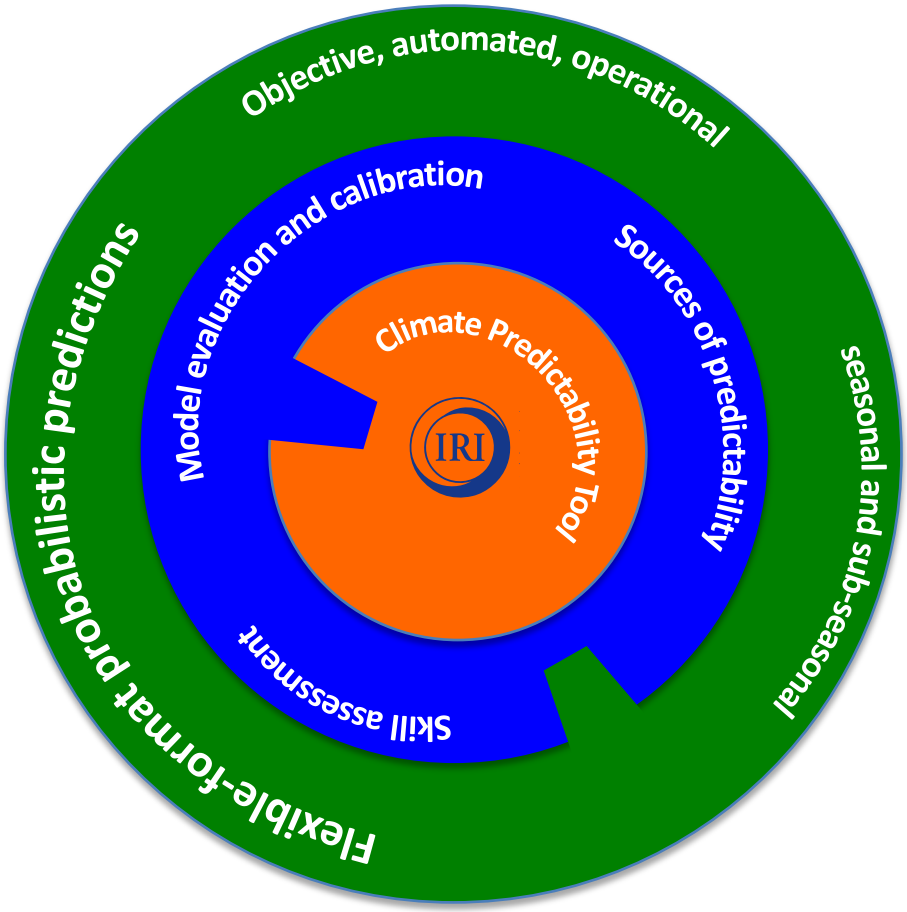
1. Key variables (predictands) selection;
2. Work with local experts to advance the identification of sources of predictability at multiple-timescales (candidate predictors);
3. Local expert evaluation of performance of publicly available GCMs (NMME, C3S, LC-LRFMME), considering different types of biases, including model biases of spatial patterns [conditional bias];
4. Pattern-based calibration of the selected GCMs.
5. Ensemble construction (in the "probability space", but deterministic ensemble is also computed);
6. Forecasts in flexible format, on the IRI (or local) Data Library.

What is #NextGen?

What is #NextGen?



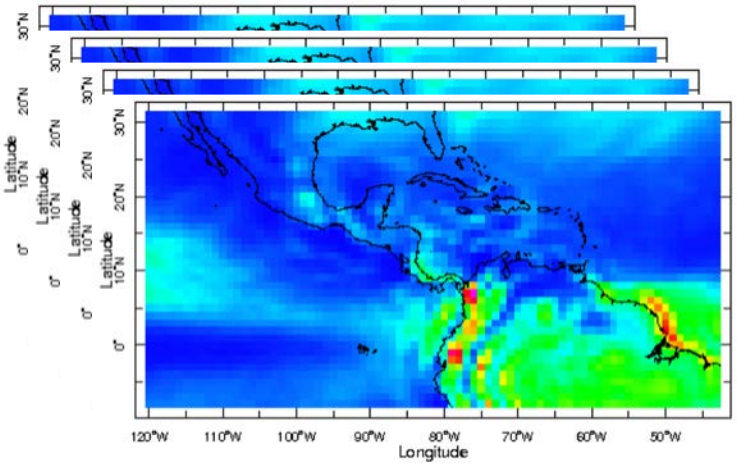
What is #NextGen?



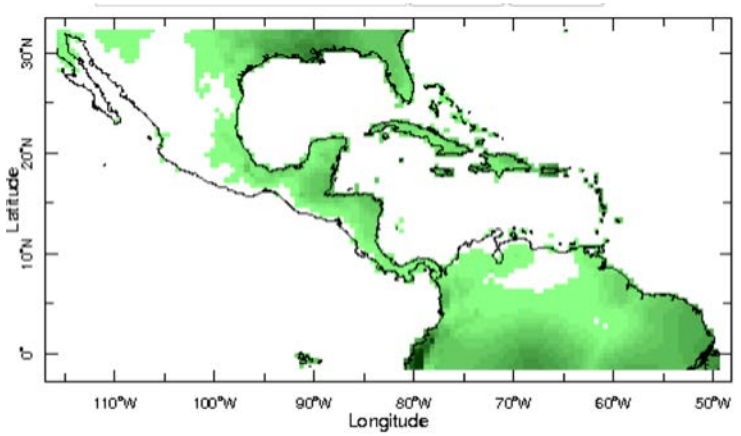
Multi-model calibrated ensemble + probabilistic flexible format

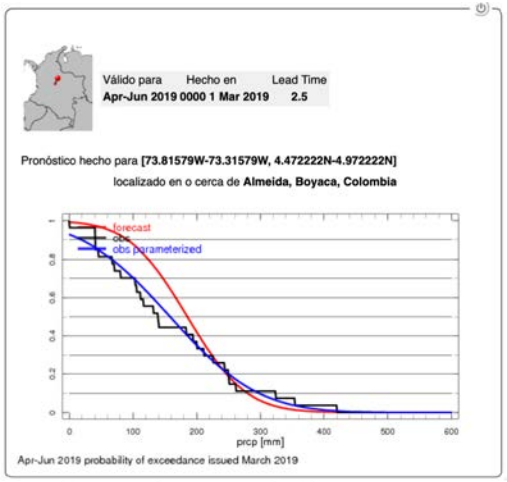
Raw Model Output (Predictors)

Multiple (C)GCMs
[NMME, C3S, LC-LRFMME]

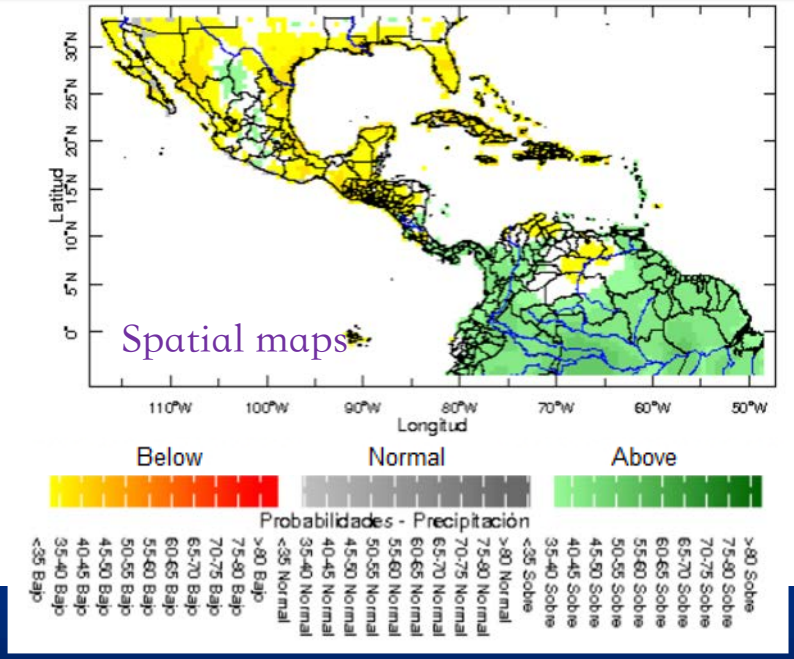


Gridded/Station Obs (Predictand)



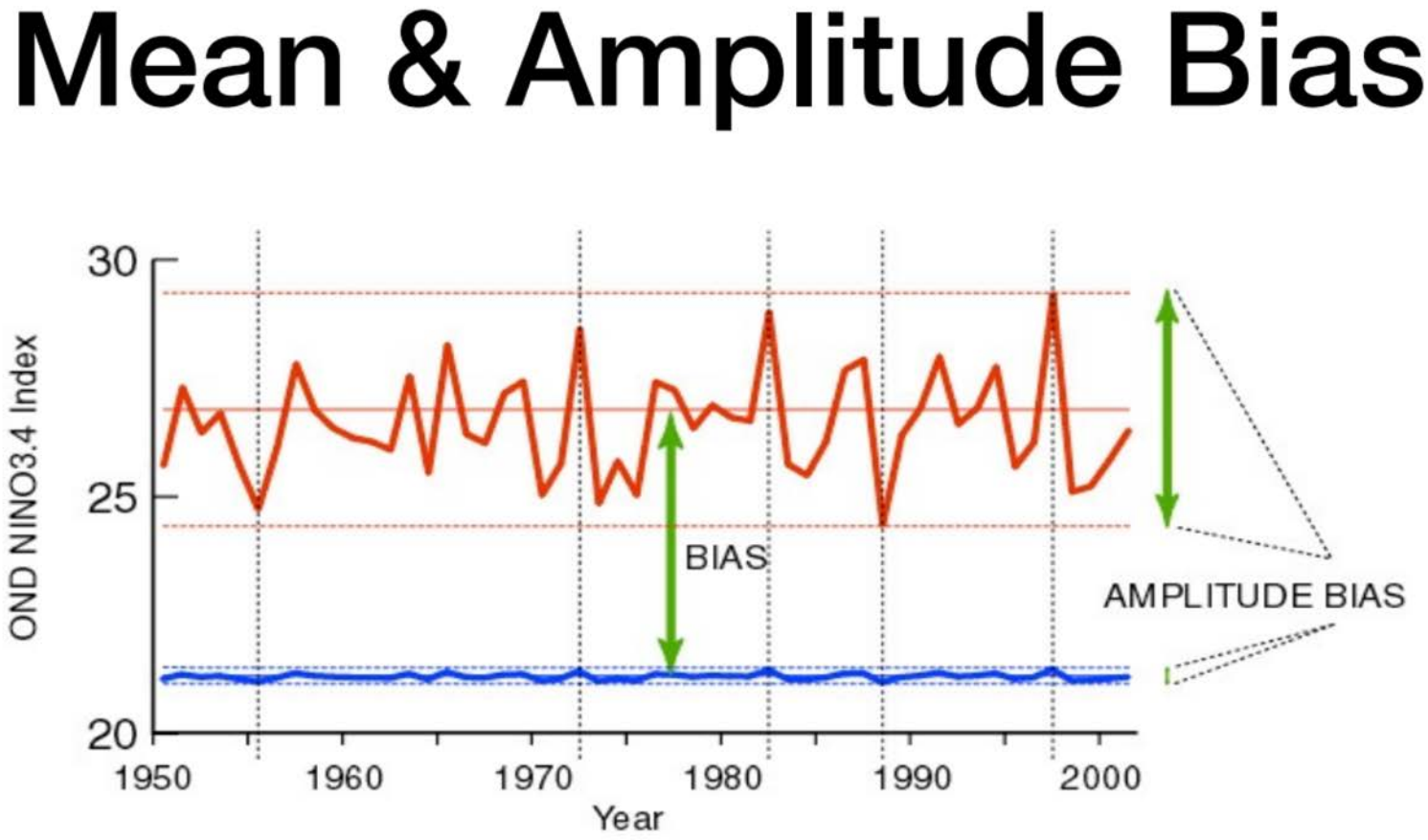


Flexformat! Use entire PDF



Pattern-based calibration (Model Output Statistics), each model independently, then ensemble in probability space

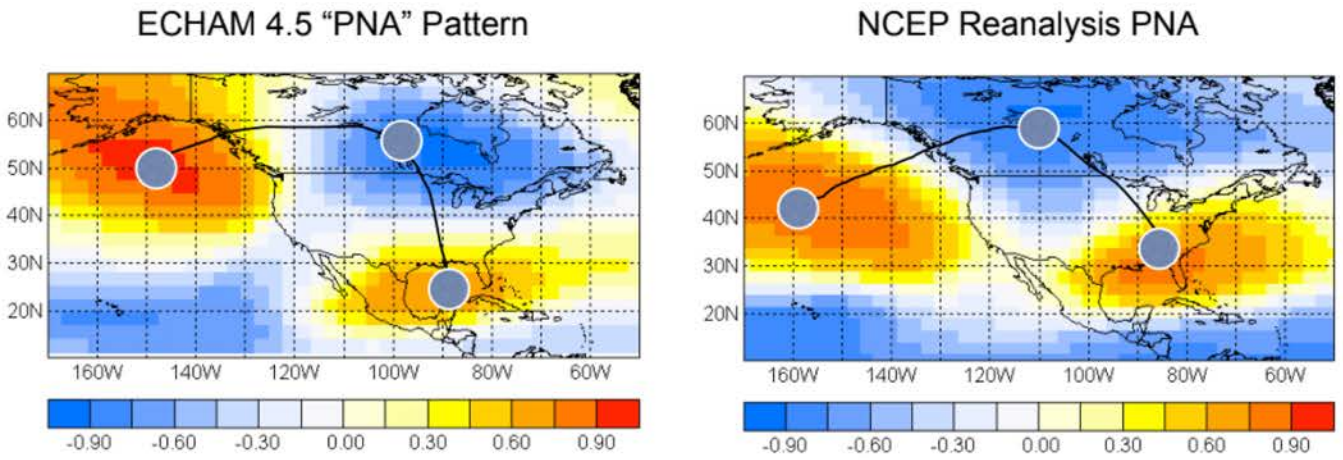
Which biases to take care of?



Courtesy of S. Mason

Which biases to take care of?

Conditional Bias
(errors in patterns of variability)



Important climate features may be displaced in GCMs relative to observations: *Systematic spatial biases*

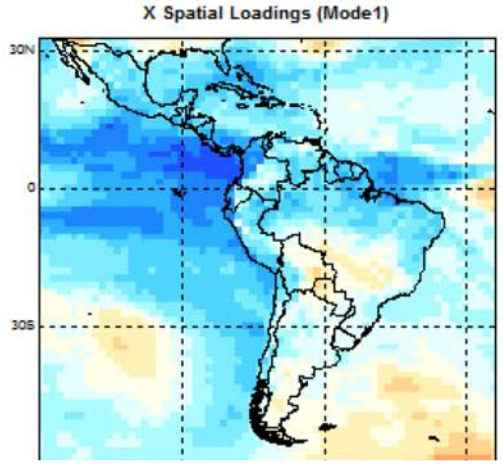
Courtesy B. Lyon (or Simon?)

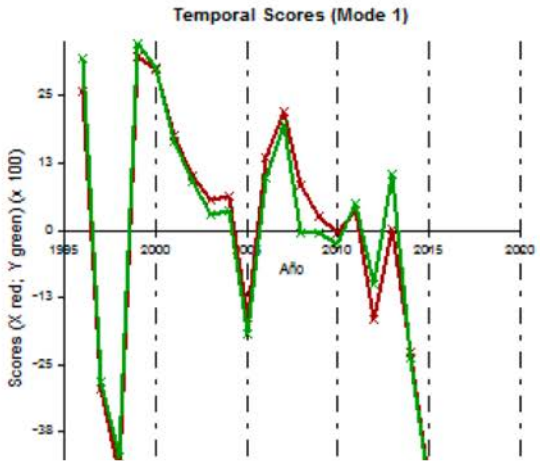
Which biases to take care of?

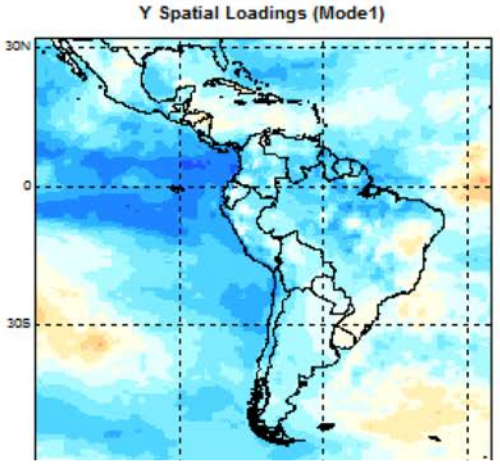
Latin America & Caribbean example

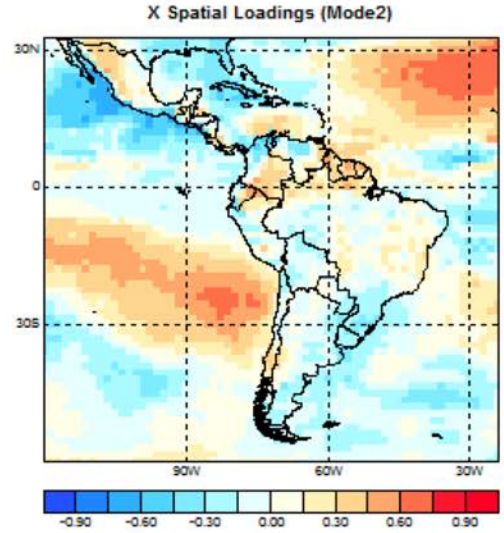
Obs

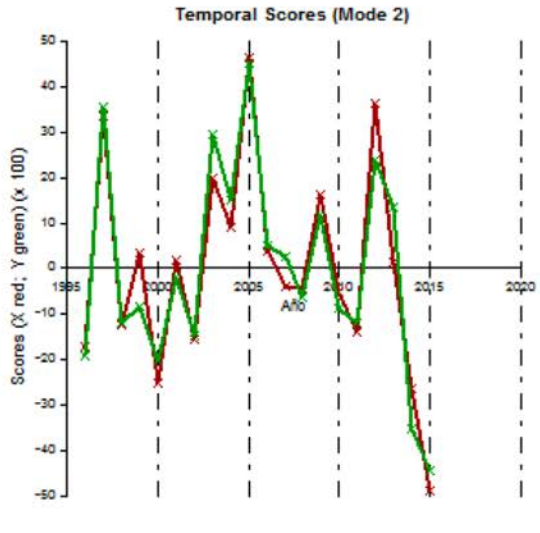
Model

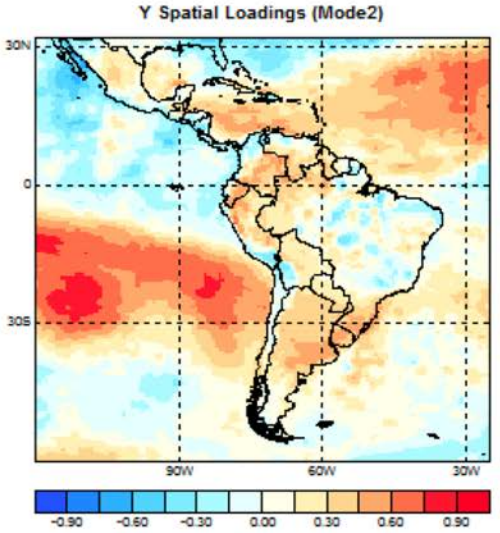












Pattern-based MOS approaches

Summary: both magnitudes (ensemble-mean field and forecast uncertainty) and spatial-patterns are corrected.
Pros: spatial correction; well-known methods; tested by IRI world-wide.
Cons: dependence on domain size due to EOF-prefiltering requires sensitivity analysis.

Steps:

1. GCM set selection via evaluation of physical mechanisms and verification of hindcasts, using
   - Model representation of physical processes
   - No-calibration skill assessment (for reference)
2. Method selection for pattern-based calibration/skill improvement. CPT options:
   - Principal component regression (PCR).
   - Canonical correlation analysis (CCA).
3. Multi-model ensemble (combination) design. Options:
   - Simple average (if forecasts are recalibrated, skill's error bars tend to be too large to justify unequal weighting)
   - Probability/PDF average

A "typical" schedule

	Activity	Month	Outcome
1	Plan development with NMHs	Month 1	Planned schedule + predictand selection
2	Sources of predictability. Skill assessment.	Months 2-4	Physical analysis. Skill maps for each model/season
3	Model selection	Month 5	Set of models ready for calibration
4	Select MOS method and combination strategy	Months 5-6	Local expert prepare scripts to perform MOS
5	Sensitivity analysis and tests	Month 7	Final predictor domain and number of EOFs identified.
6	Produce calibrated hindcasts	Months 8-9	Set of calibrated hindcasts ready for skill assessment
7	Ensemble and evaluation of skill improvement	Months 10-11	NextGen skill maps and first forecasts
8	Implementation (automation) and tests	Month 12	NextGen experimental forecasts in use by NMHs

#NextGen (seasonal): Where?

Completely implemented -via ACToday:

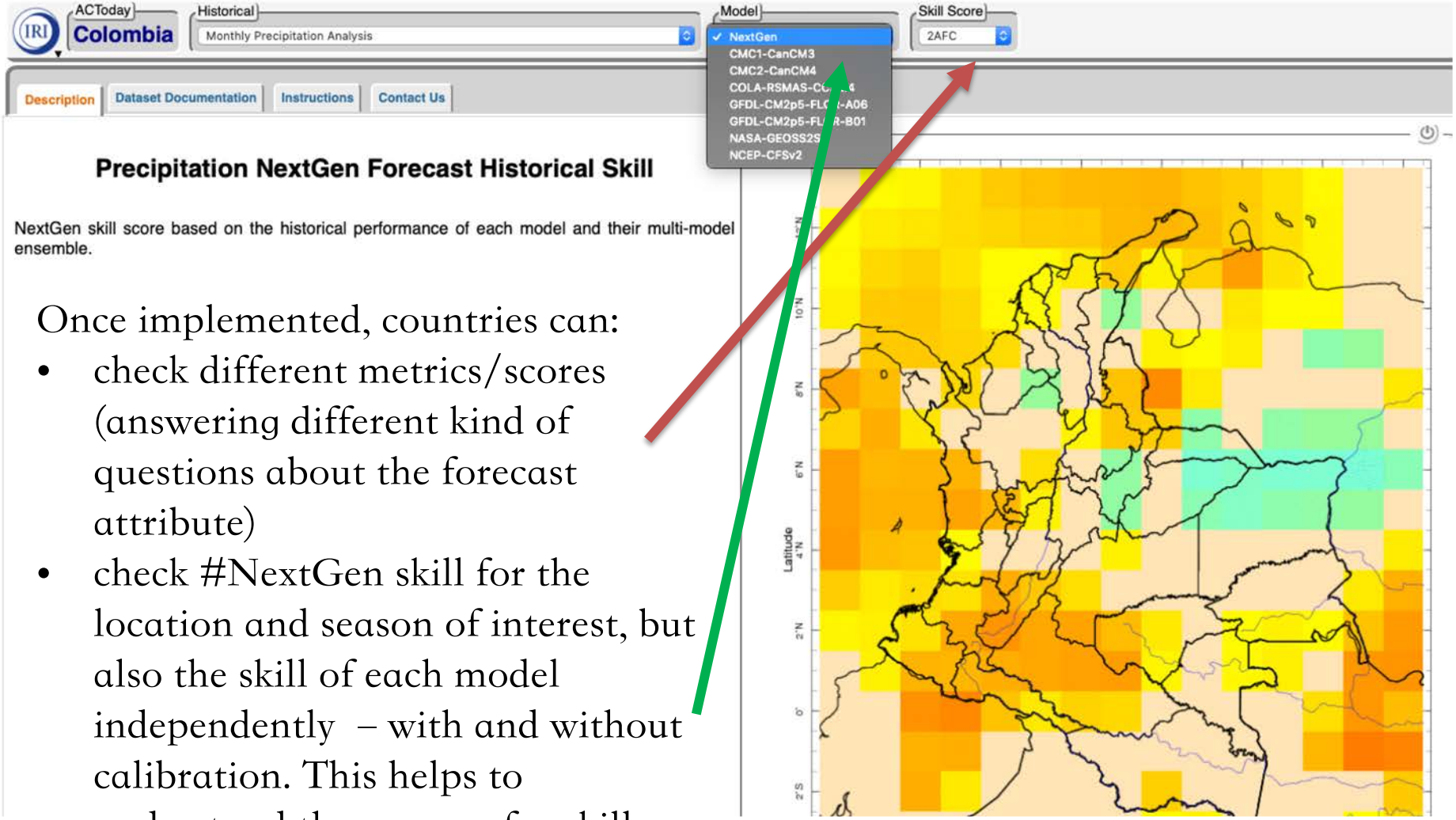
+ Colombia (IDEAM, March 2019)
+ Guatemala (INSIVUMEH, March 2019)

Potential next countries:

+ Chile (ENANDES + AEMET)
+ Peru (ENANDES)
+ Mexico
+ CRRH (Central America is very interested)
+ Other ACToday countries

How does it look like?

#NextGen: skill

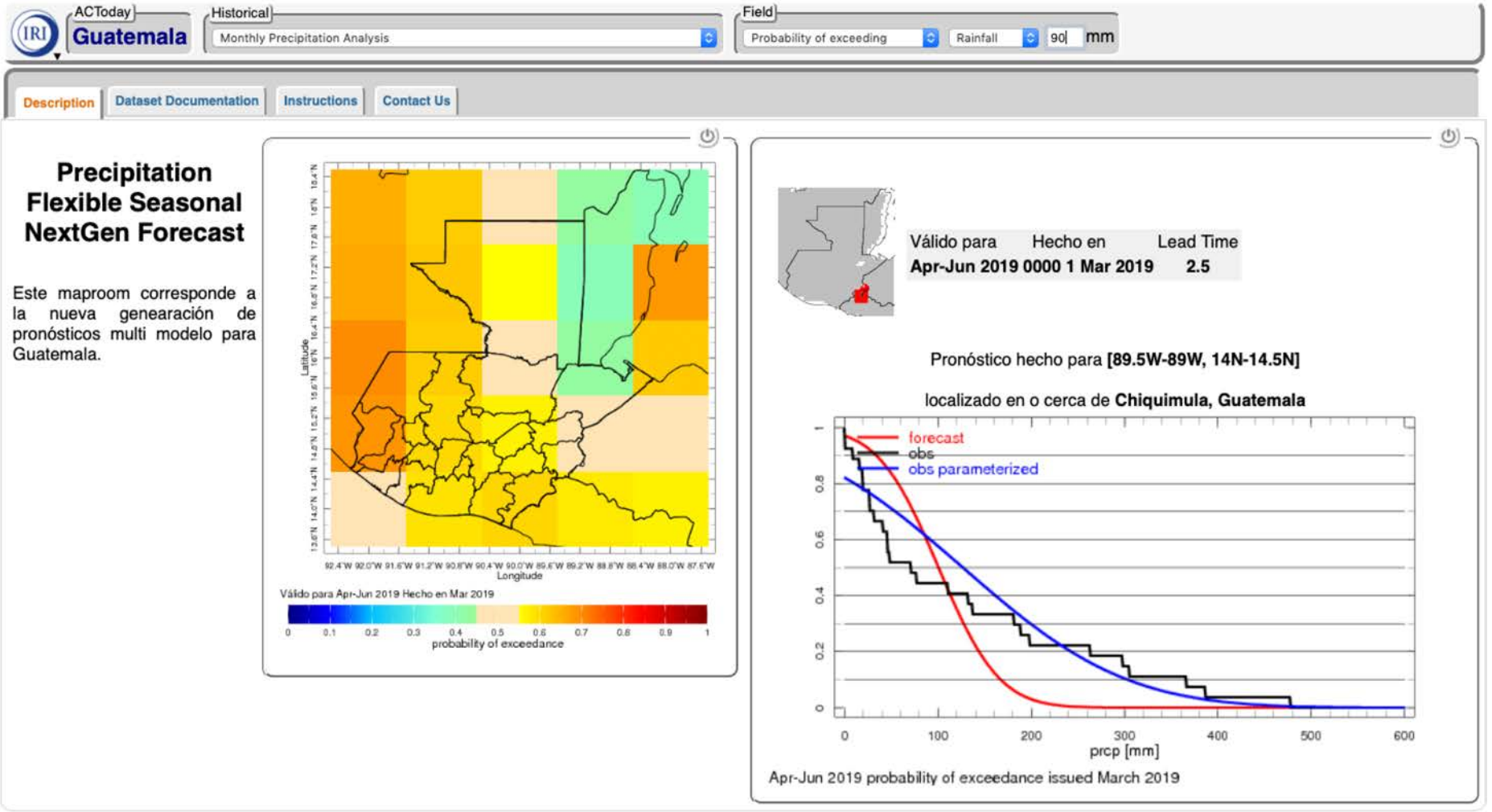


Once implemented, countries can:

- check different metrics/scores (answering different kind of questions about the forecast attribute)
- check #NextGen skill for the location and season of interest, but also the skill of each model independently – with and without calibration. This helps to understand the reasons for skill improvement brought by #NextGen

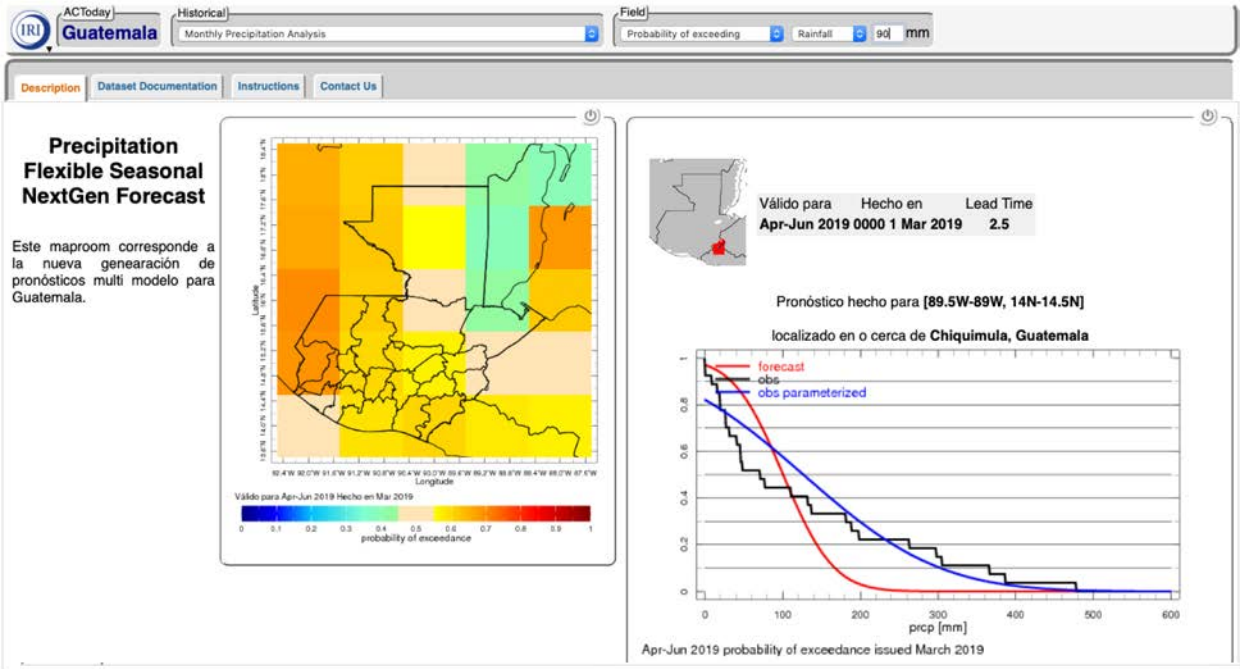
#NextGen: flexformat

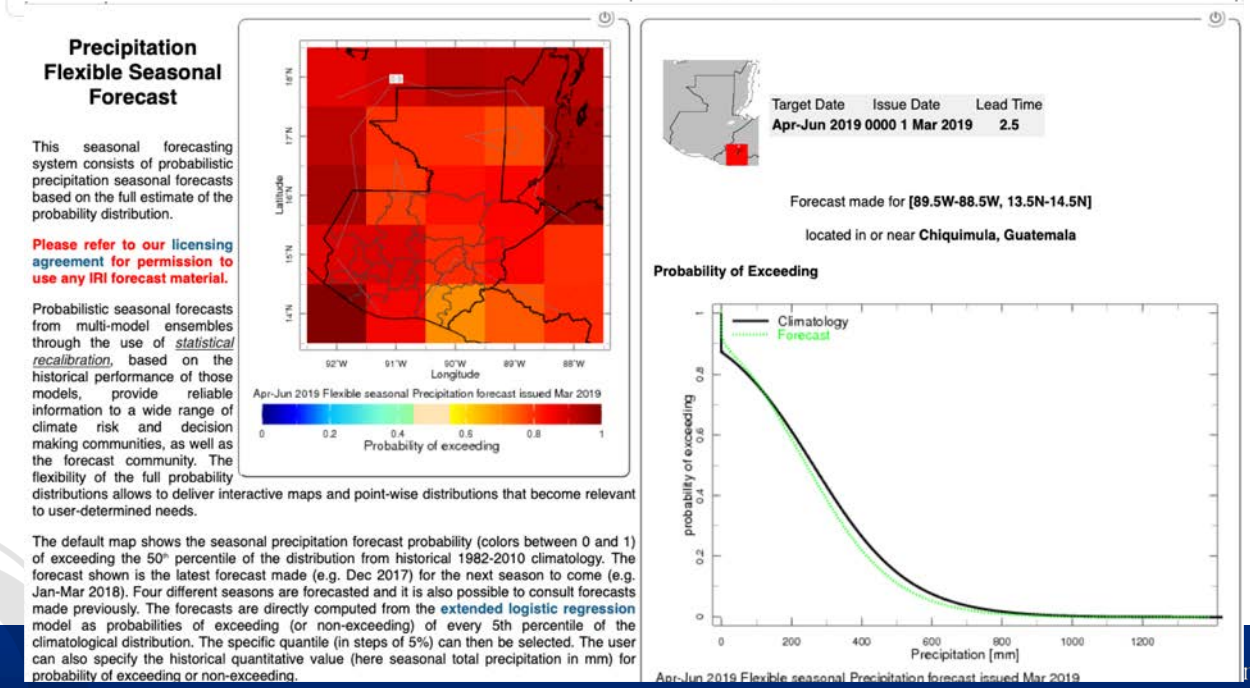
The system provides both spatial maps with deterministic values and probabilities of exceedance (or not) of particular thresholds, and the entire PDF once a location is selected.



NextGen (CCA)

IRI Forecast (ELR)





And what about sub-seasonal timescales?

#NextGen: sub-seasonal

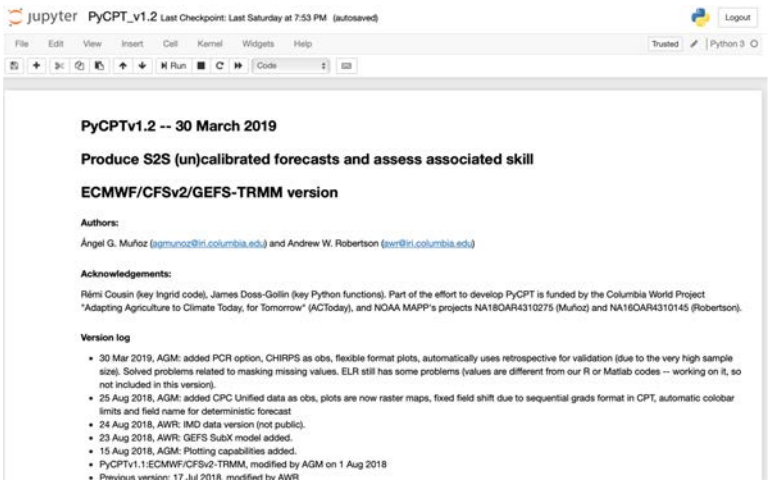
+ Forecasting subseasonal characteristics, e.g., rainfall frequency, rainy/monsoon/MSD season onset, demise, duration.

+ Exactly the same approach used for seasonal skill and predictions is used for the sub-seasonal case (PyCPT for sub-seasonal, + CPTv16!)

+ The idea is to first take advantage of the seasonal-scale experience in place -in the countries and the scientific community more broadly- before implementing #NextGen: sub-seasonal.

PyCPT: just a Python interface for CPT

https://github.com/agmunozs/PyCPT


jupyter PyCPT_v1.2 Last Checkpoint: Last Saturday at 7:53 PM (autosaved)

PyCPTv1.2 -- 30 March 2019

Produce S2S (un)calibrated forecasts and assess associated skill

ECMWF/CFSv2/GEFS-TRMM version

Authors:

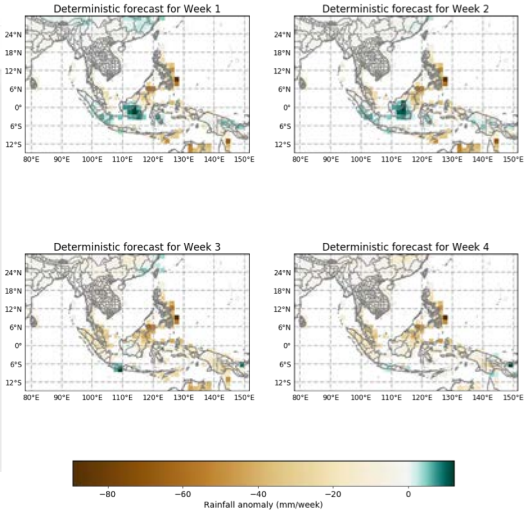
Ángel G. Muñoz (agmunoz@iri.columbia.edu) and Andrew W. Robertson (awr@iri.columbia.edu)

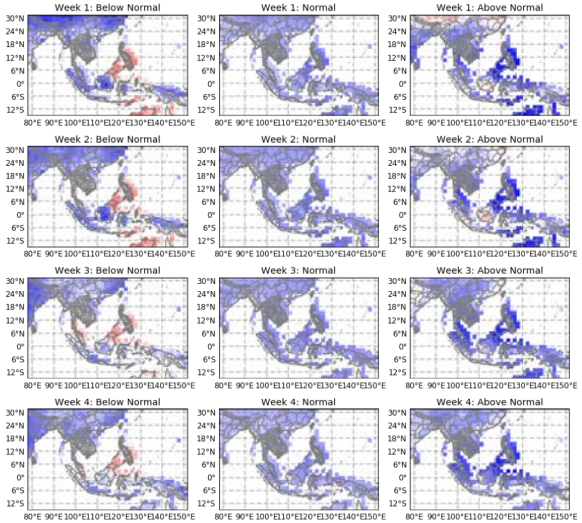
Acknowledgements:

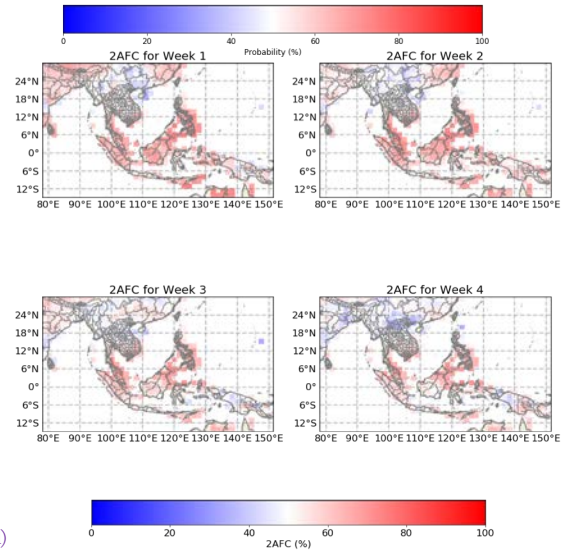
Rémi Cousin (key Ingrid code), James Doss-Gollin (key Python functions). Part of the effort to develop PyCPT is funded by the Columbia World Project "Adapting Agriculture to Climate Today, for Tomorrow" (ACToday), and NOAA MAPP's projects NA18OAR4310275 (Muñoz) and NA16OAR4310145 (Robertson).

Version log

- 30 Mar 2019, AGM: added PCR option, CHIRPS as obs, flexible format plots, automatically uses retrospective for validation (due to the very high sample size). Solved problems related to masking missing values. ELR still has some problems (values are different from our R or Matlab codes -- working on it, so not included in this version).
- 25 Aug 2018, AGM: added CPC Unified data as obs, plots are now raster maps, fixed field shift due to sequential grads format in CPT, automatic colobar limits and field name for deterministic forecast
- 24 Aug 2018, AWR: IMD data version (not public).
- 23 Aug 2018, AWR: GEFS SubX model added.
- 15 Aug 2018, AGM: Plotting capabilities added.
- PyCPTv1.1:ECMWF/CFSv2-TRMM, modified by AGM on 1 Aug 2018





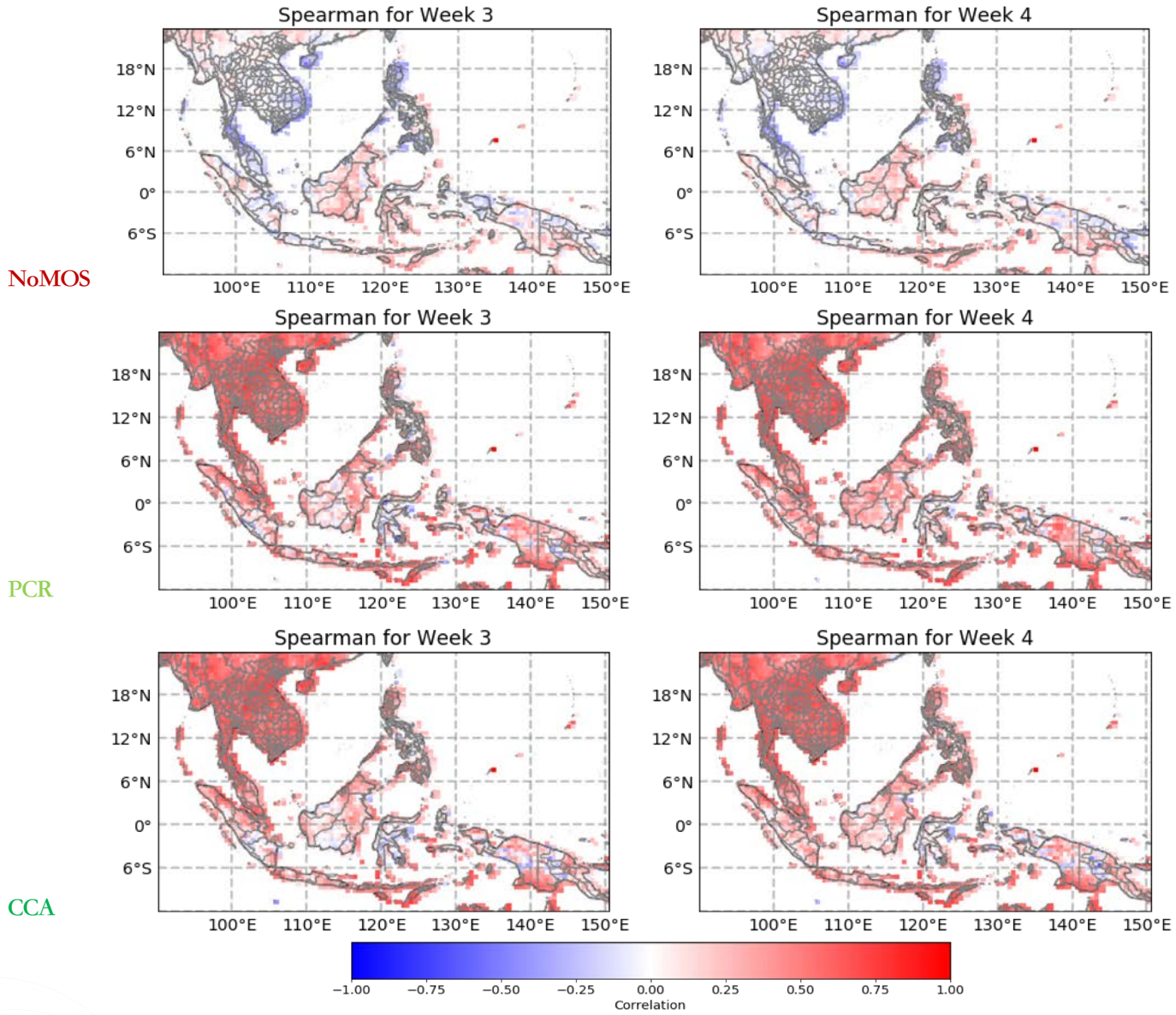


- Python interface for IRI's Climate Predictability Tool (CPT), a widely used research and application Model Output Statistics/Prediction toolbox.
- Publicly available: GitHub.
- Automatically downloads required observations (CHIRPS, TRMM, CPC Unified) and S2S model data from the IRI Data Library (**S2S Database** and **SubX** – ECMWF, CFSv2, GEFS, others are being included).
- Computes climatologies, anomalies, a variety of skill metrics (uncalibrated and PCR/CCA-calibrated hindcasts) and probabilistic sub-seasonal forecasts.

Funded by:

NOAA NA16OAR4310145 (Robertson)
NOAA NA18OAR4310275 (Muñoz)
Columbia World Project "ACToday" (Goddard)

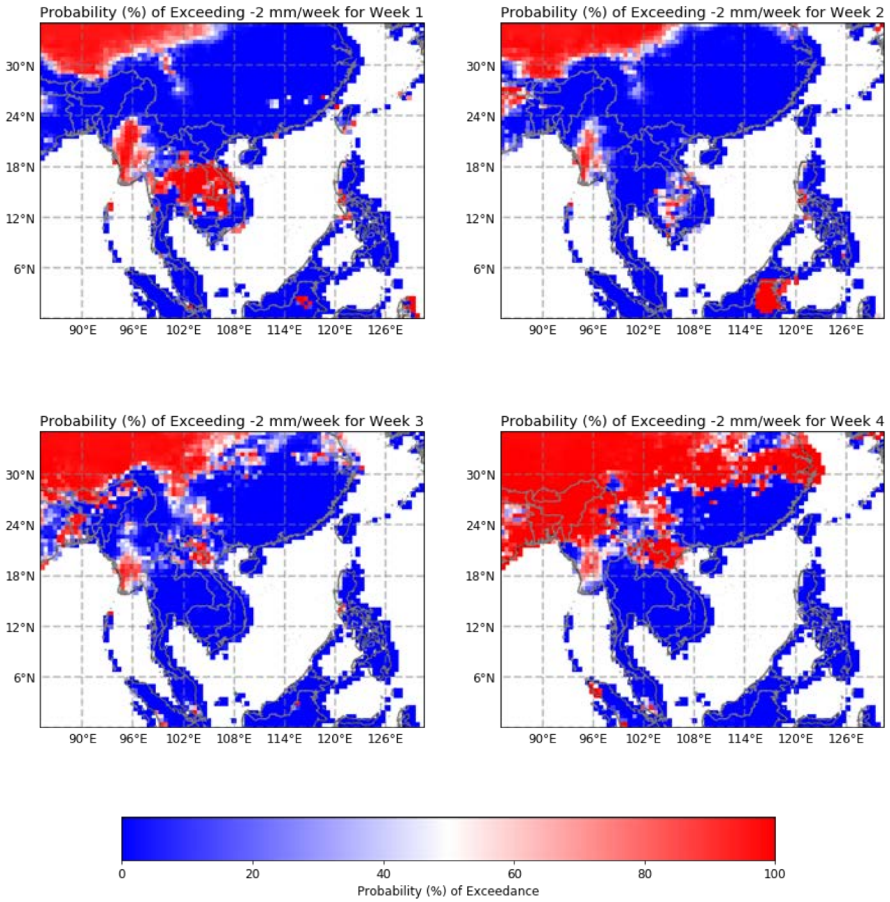
NextGen: sub-seasonal, in action

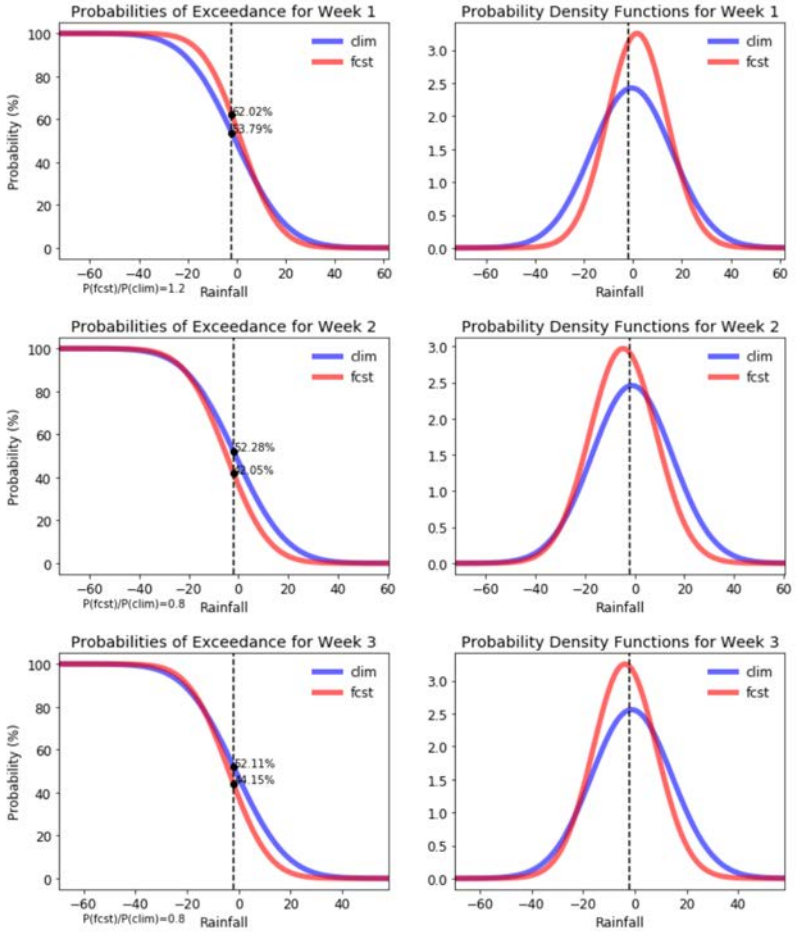


CFSv2 Rainfall. Init: ~14 March 2019

Turkington *et al.* (in prep)

NextGen (subseasonal) in action



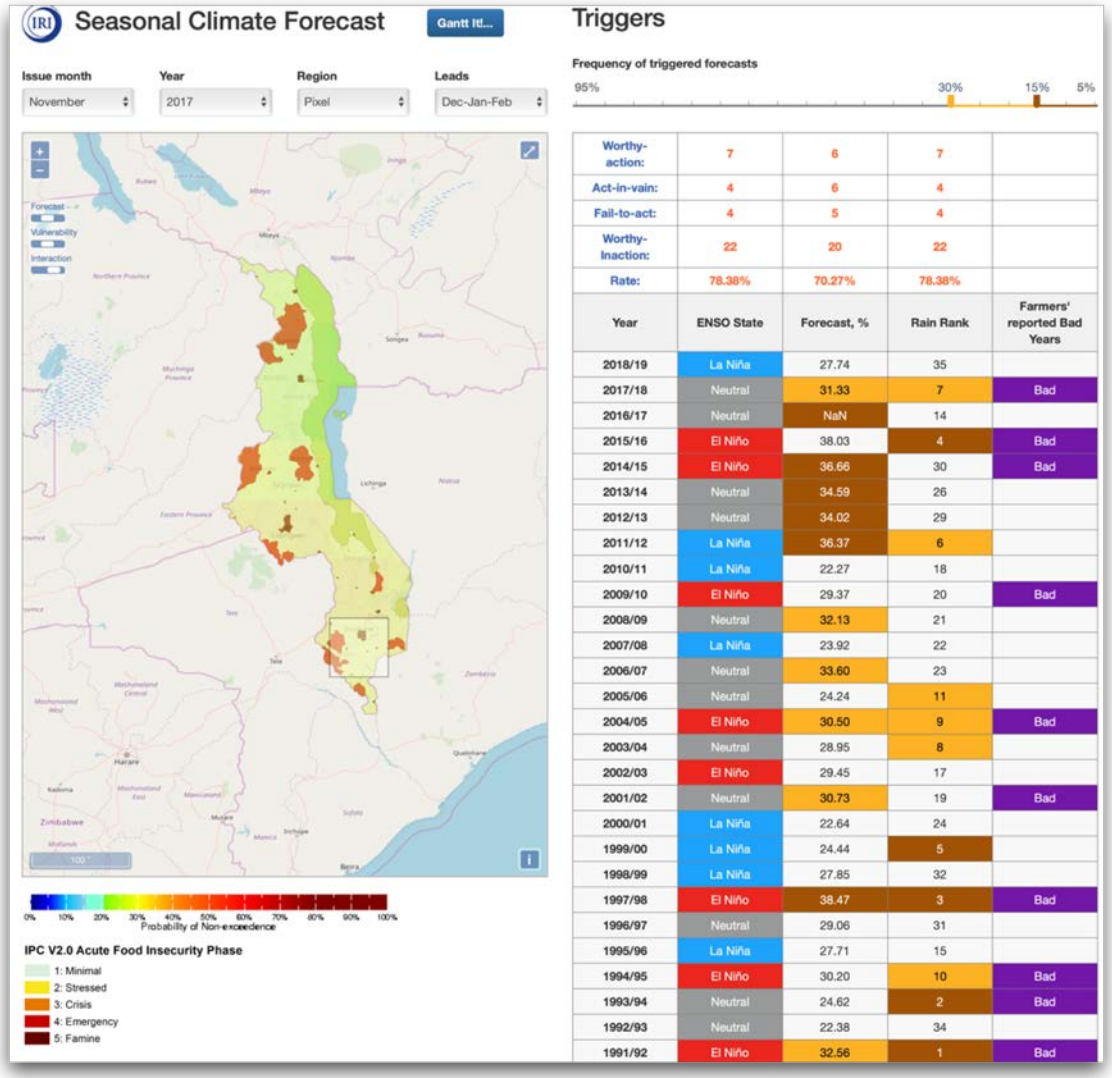


So what?

NextGen: so what? #1

ACToday is working with WFP, INSIVUMEH (NMS) and other partners in Guatemala to use the recently implemented #NextGen for a drought FbF in the Dry Corridor.

Figure shows a Python-based mapsoom example for Malawi.



Worthy-action:	7	6	7	
Act-in-vain:	4	6	4	
Fail-to-act:	4	5	4	
Worthy-Inaction:	22	20	22	
Rate:	78.38%	70.27%	78.38%	
Year	ENSO State	Forecast, %	Rain Rank	Farmers' reported Bad Years
2018/19	La Niña	27.74	35	
2017/18	Neutral	31.33	7	Bad
2016/17	Neutral	NaN	14	
2015/16	El Niño	38.03	4	Bad
2014/15	El Niño	36.66	30	Bad
2013/14	Neutral	34.59	26	
2012/13	Neutral	34.02	29	
2011/12	La Niña	36.37	6	
2010/11	La Niña	22.27	18	
2009/10	El Niño	29.37	20	Bad
2008/09	Neutral	32.13	21	
2007/08	La Niña	23.92	22	
2006/07	Neutral	33.60	23	
2005/06	Neutral	24.24	11	
2004/05	El Niño	30.50	9	Bad
2003/04	Neutral	28.95	8	
2002/03	El Niño	29.45	17	
2001/02	Neutral	30.73	19	Bad
2000/01	La Niña	22.64	24	
1999/00	La Niña	24.44	5	
1998/99	La Niña	27.85	32	
1997/98	El Niño	38.47	3	Bad
1996/97	Neutral	29.06	31	
1995/96	La Niña	27.71	15	
1994/95	El Niño	30.20	10	Bad
1993/94	Neutral	24.62	2	Bad
1992/93	Neutral	22.38	34	
1991/92	El Niño	32.56	1	Bad

NextGen: so what? #2

ACToday is working with FAO to assimilate #NextGen into their Agriculture Stress Index System (ASIS), which so far has been used only for monitoring purposes.

Transforming ASIS into a monitoring & forecast system is expected to help FAO's work in several countries.

Both Guatemala and Colombia are being explored as case studies right now, and Ethiopia might follow soon.

NextGen: so what? #3

MTAs in Guatemala and Colombia will start using NextGen's flexible format predictions soon. The system is already being used by INSIVUMEH and IDEAM to make decisions regarding the present below-normal rainfall conditions in those countries.

ENANDES, an USD$7.3M WMO-Adaptation Fund's project, is planning to use NextGen in Colombia, Peru and Chile.

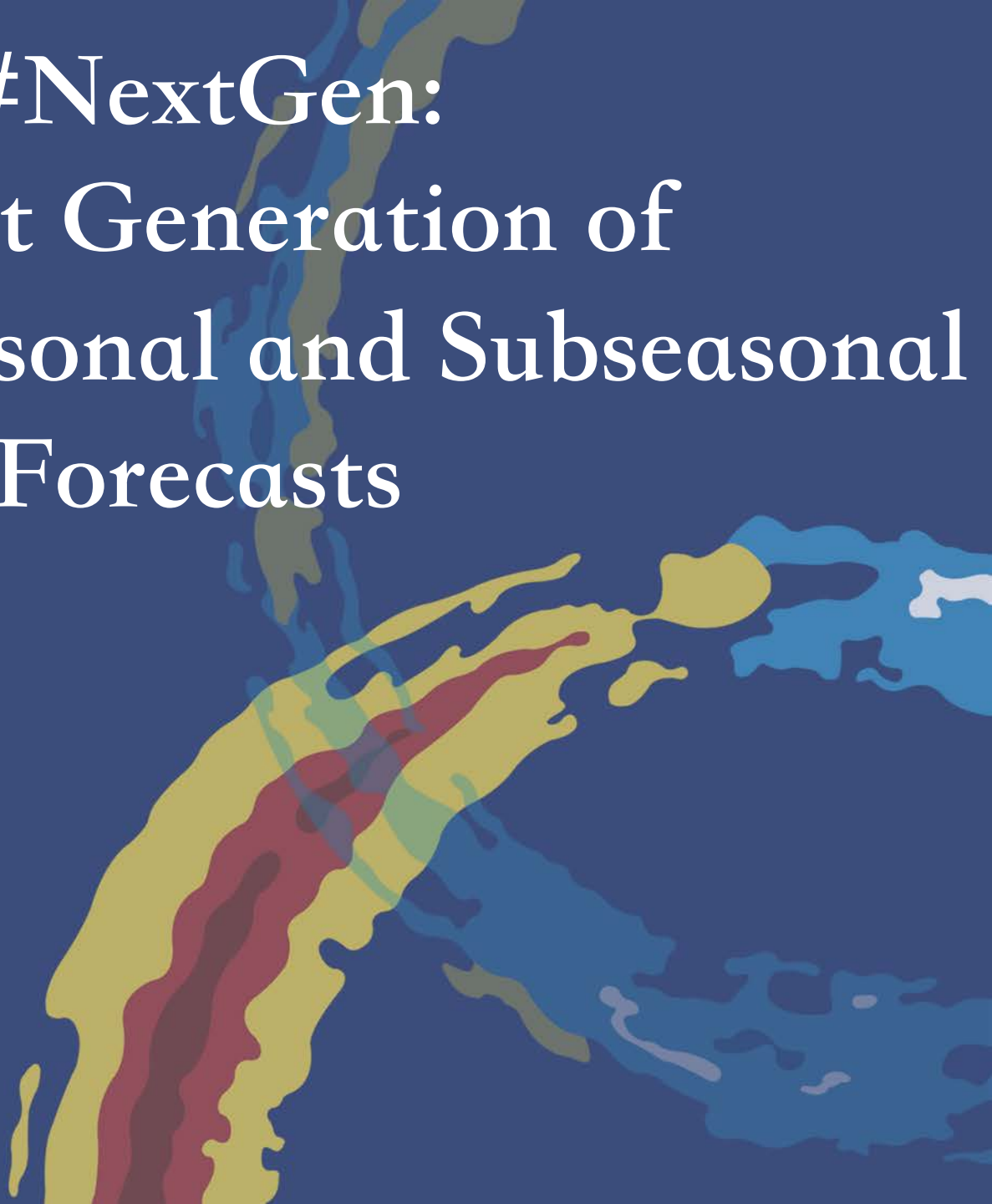
WMO has indicated interest in using NextGen as part of the RCOF's standard methodologies.

Discussion time!

#NextGen: A Next Generation of Regional Seasonal and Subseasonal Forecasts

Á.G. Muñoz
(agmunoz@iri.columbia.edu),
A.W. Robertson,
S.J. Mason,
L. Goddard
D. Pons,
X. Chourio

International Research Institute
for Climate and Society
Earth Institute | Columbia University



Extra slide(s)

No pattern-based MOS approaches

Summary: only magnitudes of ensemble-mean field and forecast uncertainty are corrected, gridbox by gridbox.
Pros: "simple" and fast set of methods
Cons: models tend to have biases in spatial patterns, that are not corrected in this approach; several options need to be explored (see below).

Steps:

1. Evaluation and selection of GCMs via verification of hindcasts, using
   - No correction (for reference)
   - Correct mean and variance biases
   - Correct based on skill of ensemble mean
2. Multi-model ensemble (combination) design. Options:
   - Simple average (calibrated: mean and variances are corrected)
   - Simple average (recalibrated: model average corrected for skill)
   - Skill-weighted average (MLR is used to combine models)
   - Best model by location

Not explored yet
To be explored later, if necessary