Centro de Previsão de Tempo e Estudos Climáticos

CPTEC / INPE

www.cptec.inpe.br

A verification framework for South American sub-seasonal precipitation predictions

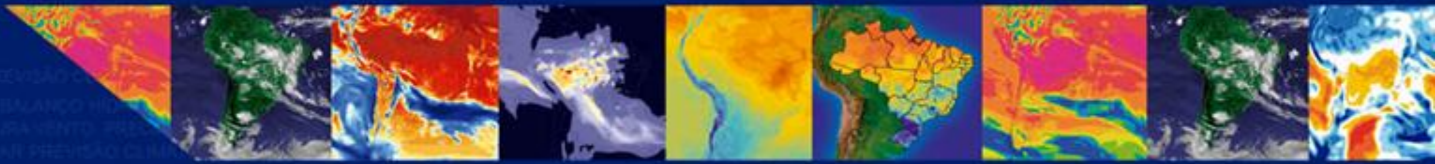
Caio Coelho, CPTEC/INPE

Plan of talk

1. Introduction: Current context, aspects to be considered
2. Elucidation of the sub-seasonal verification problem and associated questions
3. Sampling strategies and information levels for sub-seasonal verification
4. Attribute-based forecast quality assessment
5. Summary

Acknowledgments: Mári Firpo, Felipe Andrade, and ECMWF for making forecast and hindcast data avail. via S2S project data base

Ciência e Tecnologia a serviço da sociedade



The current context

• Recent availability of sub-seasonal predictions produced as part of the WWRP/WCRP Sub-seasonal to Seasonal prediction project (S2S, Vitart et al., 2012; Robertson et al., 2015) allows the investigation of retrospective predictions (hindcast) and real time forecast quality levels of the participating S2S modeling centers.

• Verification strategy is required to document the quality of both deterministic and probabilistic predictions in support of future routine sub-seasonal predictions.

• This strategy is required because verification information detailing past model performance is a key prediction practice component to enhance forecasters´ confidence on the available models predictions and also in support of future model developments. This study proposes a verification framework for these purposes.

Important aspects to be considered

• Large degree of differences in some characteristics of sub-seasonal hindcasts and real time forecasts, directly impacting the verification sample size.

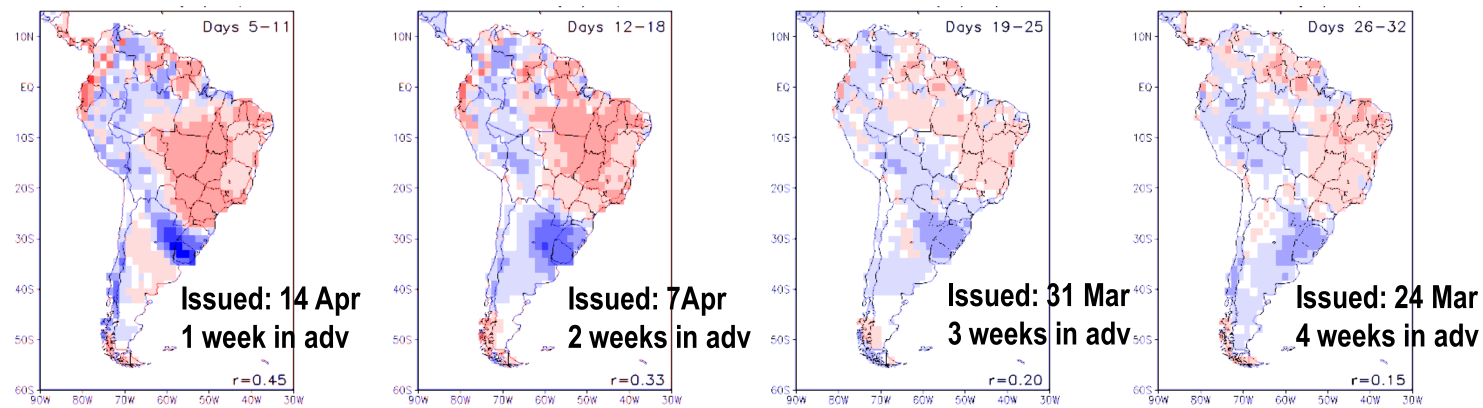
• For example, the number of available sub-seasonal hindcast years (typically 20 years or less) is usually reduced compared to the number of available seasonal hindcast years (typically 30 years).

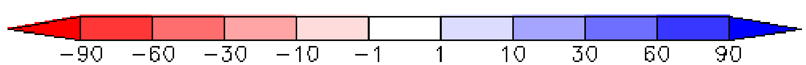
• In the S2S project very few real time subseasonal forecast years are available for verification (about 3-4 years) with a typically much larger ensemble size than usually available for hindcasts.

• These differences in sub-seasonal hindcasts and real time forecasts highlight the need for a strategy for sub-seasonal prediction verification practice.

Elucidation of the sub-seasonal verification problem and associated questions

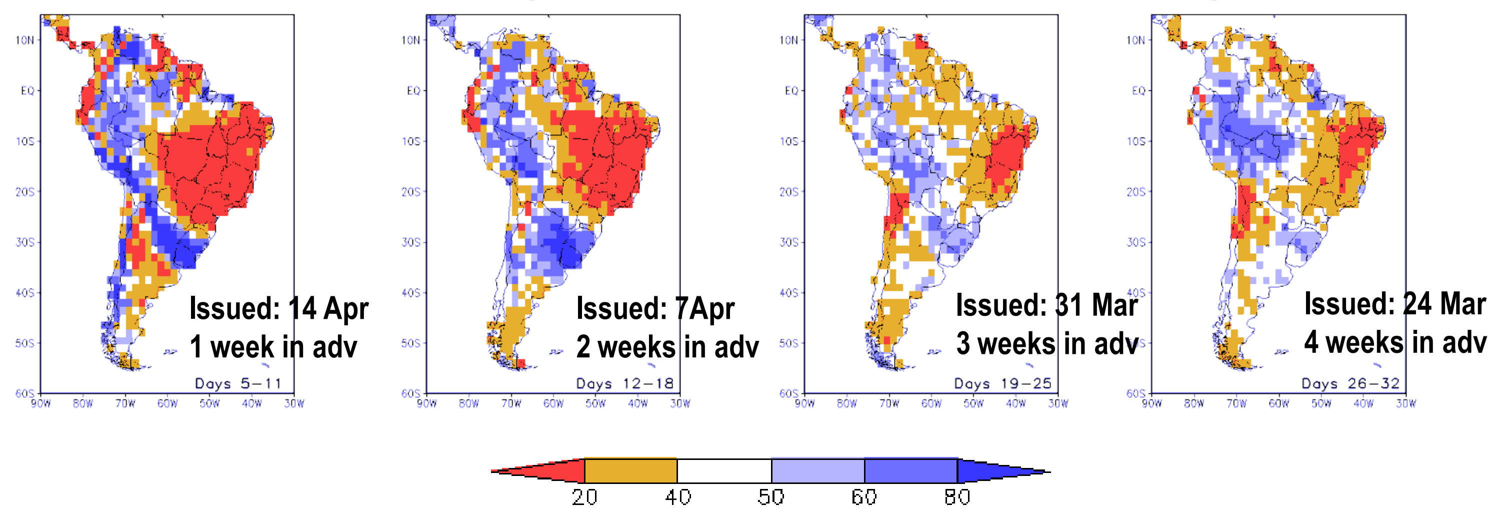
Deterministic (ens. mean anomaly) precip. forecasts for 18-24 April 2016





ECMWF S2S forecasts

Probabilistic forecasts (prob. of pos. precip. anom.) for 18-24 April 2016

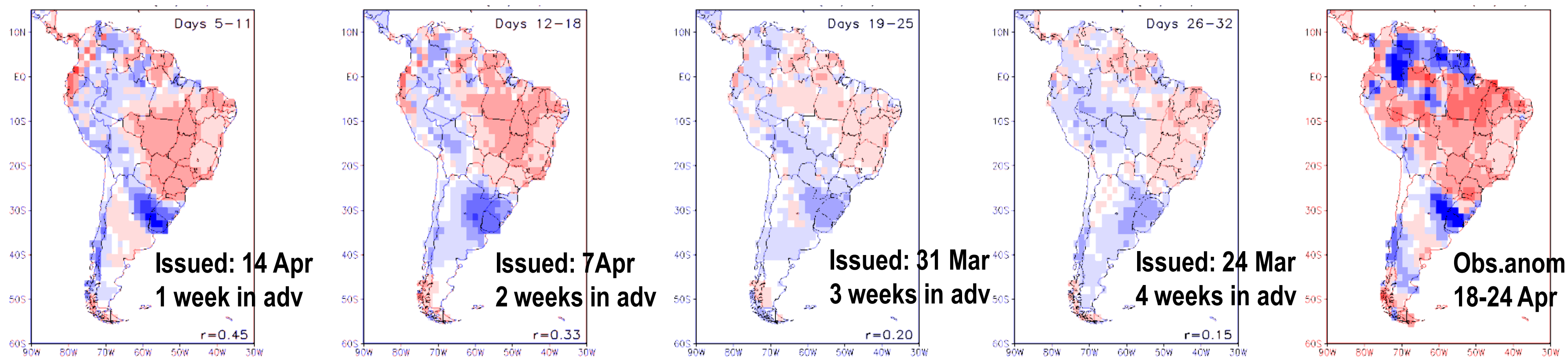


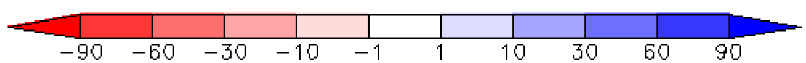
Questions from forecasters having access to forecasts prior to 18-24 April 2016

- How good are these forecasts for the week 18-24 April 2016 produced one to four weeks in advance in terms of correspondence with the observations?
- Where spatially can these forecasts be best trusted?
- How strong is the relationship between the forecast and observed precipitation anomalies?
- How accurate are the forecast precipitation anomalies compared to the accuracy of a reference naïve forecasting strategy of always issuing a constant forecast value (e.g. null anomaly for the climatological forecast)?
- How reliable are the issued forecast probabilities?
- Can the issued forecast probabilities detect the event of interest (i.e. distinguish events from non-events)?

Qualitative assessment and visual identification of regions where forecasts were successful

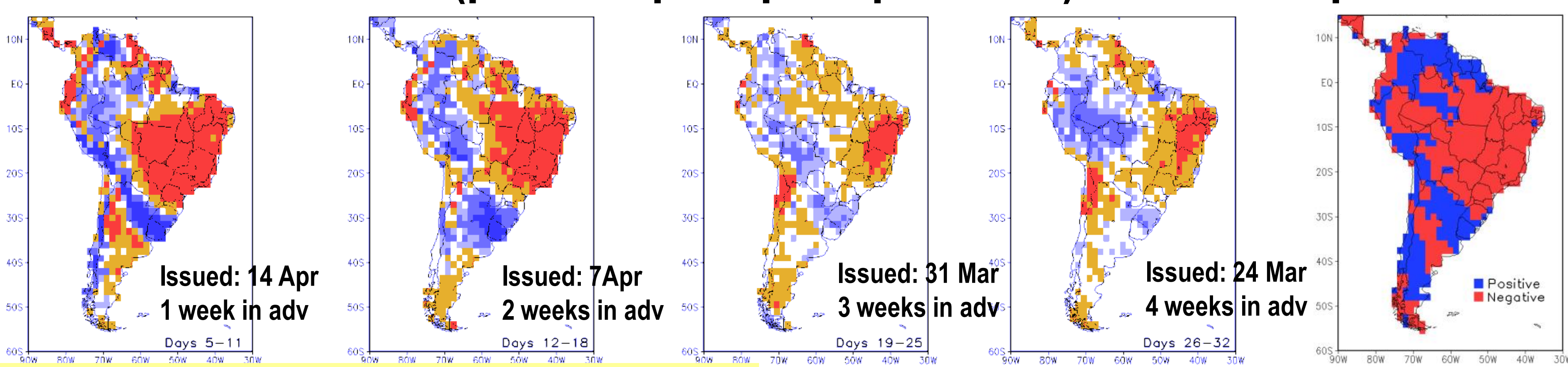
Deterministic (ens. mean anomaly) precip. forecasts for 18-24 April 2016

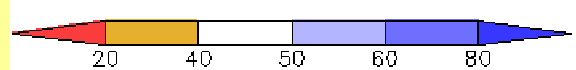




ECMWF S2S forecasts

Probabilistic forecasts (prob. of pos. precip. anom.) for 18-24 April 2016





Binary obs
18-24 Apr

Useful initial assess: Quantitative approach required to document past fcst quality (support to fcst users)

Sampling strategies and information levels for sub-seasonal verification

www.cptec.inpe.br

Proposed framework for quantitative sub-seasonal precip. forecast quality assessment

- **Level 1: Target week hindcast verification (11 ens. members)**

  Similar to traditional seasonal forecast verification (~30 samples)

  Uses ECMWF S2S hindcasts intialized on Thu 14 April, 7 April, 31 March and 24 March of the 2016 calendar for the period 1996-2015 (20 samples)

- **Level 2: All season hindcast verification (11 ens. members)** Increased robustness

  In addition to the hindcasts produced for the four Thu initialization dates previously selected, aggregates hindcasts produced for nine additional initialization dates during the weeks of the previous and following month in order to incorporate in the sample all hindcasts initialized on Thu of March, April and May of the 2016 calendar (260 samples: 13 initialization dates times 20 years)

  MAM: Austral summer season, similar atmospheric features in S. American regions

- **Level 3: All season near real time forecast verification (51 ens. members)**

  Aggregate the real time forecasts produced on Thu during the 13 weeks of March, April and May of each of the past three years (2015, 2016 and 2017).

  This aggregation leads to a verification sample of 39 pairs of near real time forecasts and observations (39 samples: 13 initialization dates times 3 years)

CPTEC
www.cptec.inpe.br

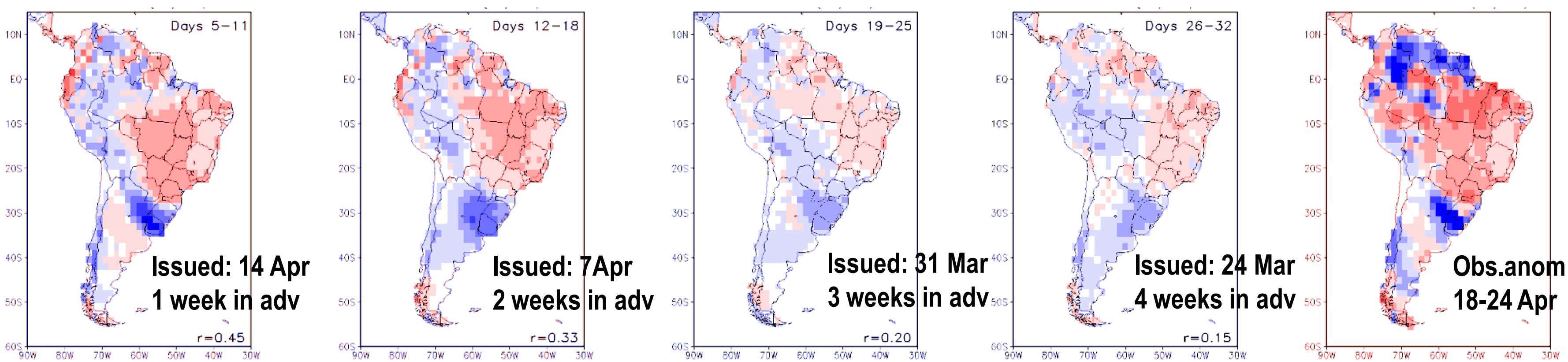
Attribute-based forecast quality assessment

Murphy (1993) defined a number of aspects, so-called attributes, for assessing forecast quality (corresp. btw. fcsts and obs). The proposed procedures for assessing sub-seasonal precipitation predictions is based on a selection of some of the most fundamental attributes: association, accuracy, discrimination, reliability and resolution.

Proposed metrics:

- **Spatial pattern correlation (r) btw forecast and obs anomalies: quantify the degree of correspondence between the deterministic forecasts and the observations**

Deterministic (ens. mean anomaly) precip. forecasts for 18-24 April 2016



r: 0.45 **r: 0.33** **r: 0.20** **r: 0.15**

Association comparative assessment: Corr. btw pred. ens. mean and obs. precip. anom.

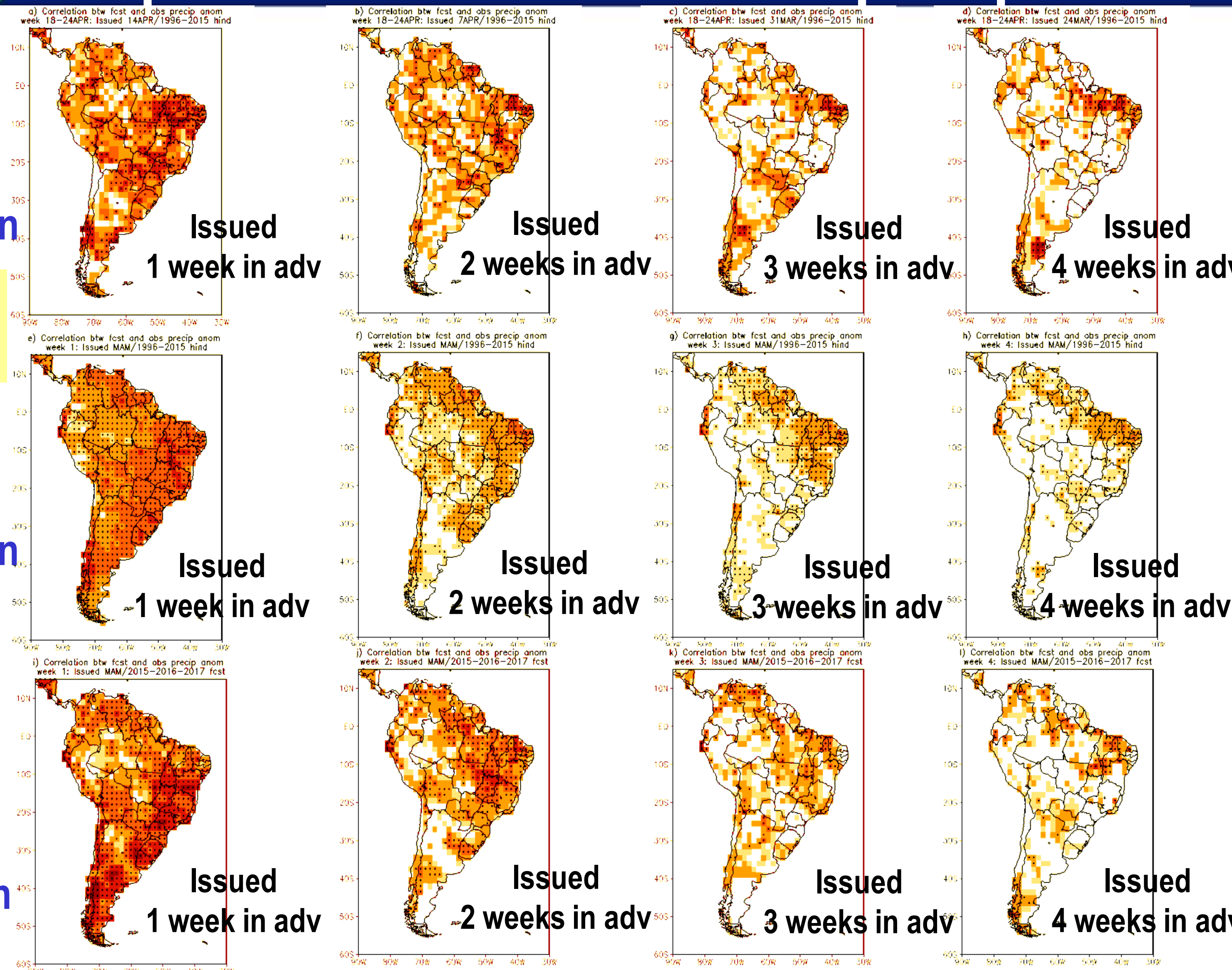
www.cptec.inpe.br

Level 1:
Target week hindcast verification

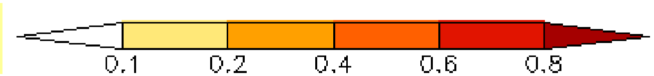
Phase assessment
Consit. across 3 lev.

Level 2:
All season hindcast verification

Level 3:
All season near real time forecast verification



Measure strength of linear association btw pred ens mean and obs anom



Accuracy comparative assessment: MSSS for ens. mean precip.anom pred. wrt clim.

CPTEC
www.cptec.inpe.br

Level 1: Target week hindcast verification

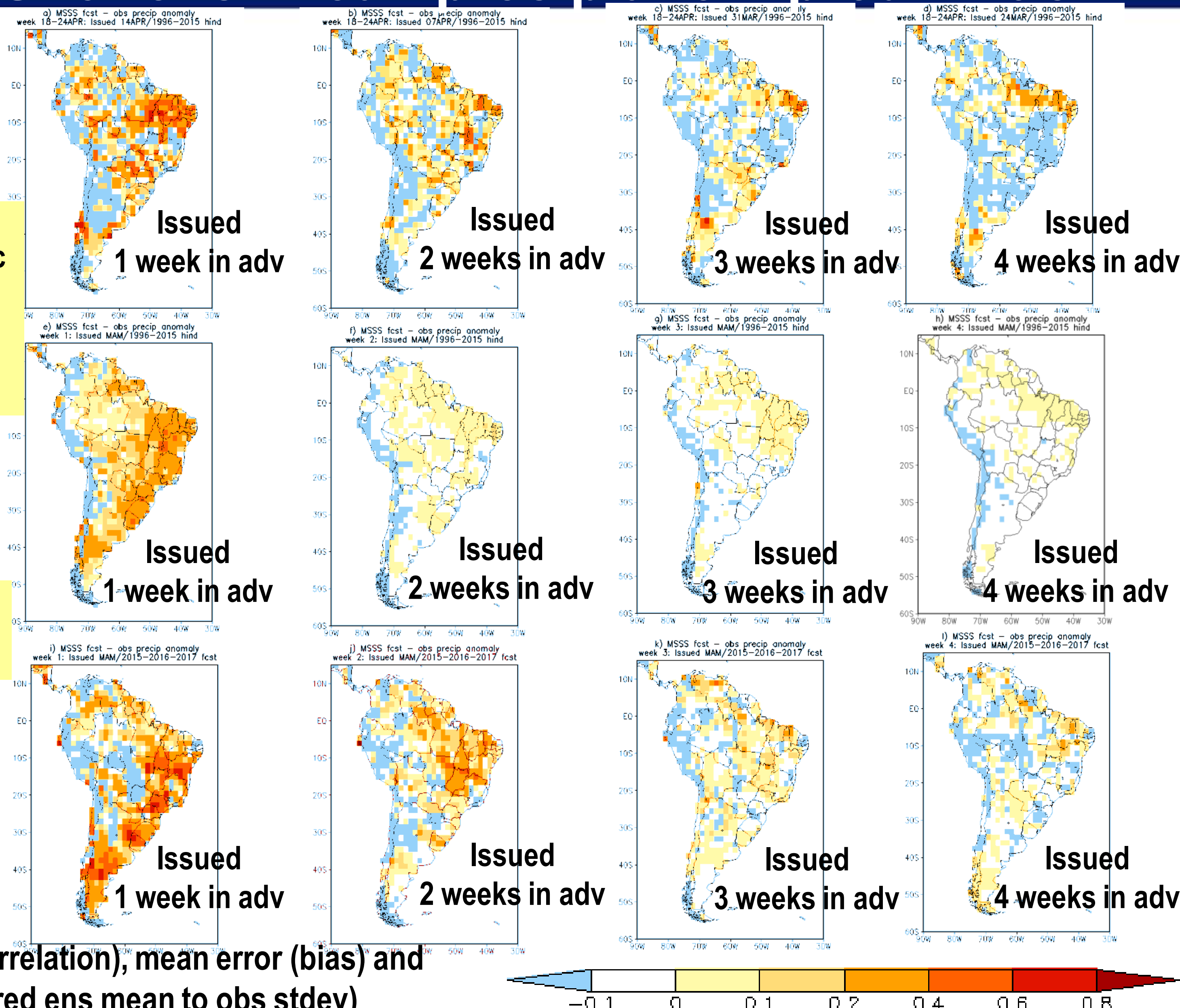
$MSSS=1-MSE/MSE_c$

MSSS>0 greater accuracy than ref (climat.) prediction

Level 2: All season hindcast verification

MSE: det. meas. (ave square diff btw pred ,obs)

Level 3: All season near real time forecast verification



Murphy (1988): MSSS incl phase error (correlation), mean error (bias) and amplitude error (ratio of pred ens mean to obs stdev)

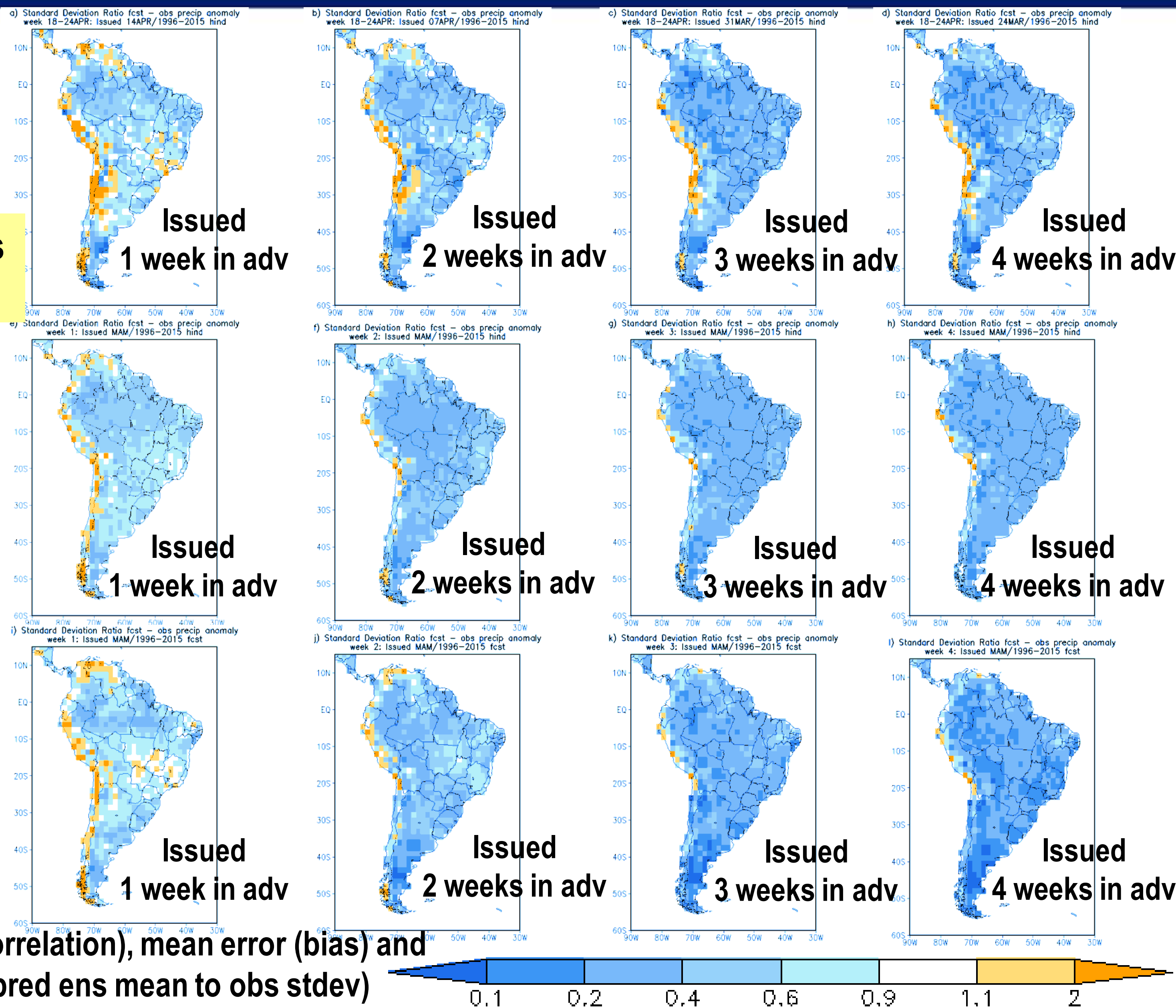
Amplitude error comparative assessment: Ratio btw. pred. precip. ens. mean anom. stdev and obs. anom. stdev

Level 1:
Target week hindcast verification

Ratio < 1 pred with less variability than obs

Level 2:
All season hindcast verification

Level 3:
All season near real time forecast verification



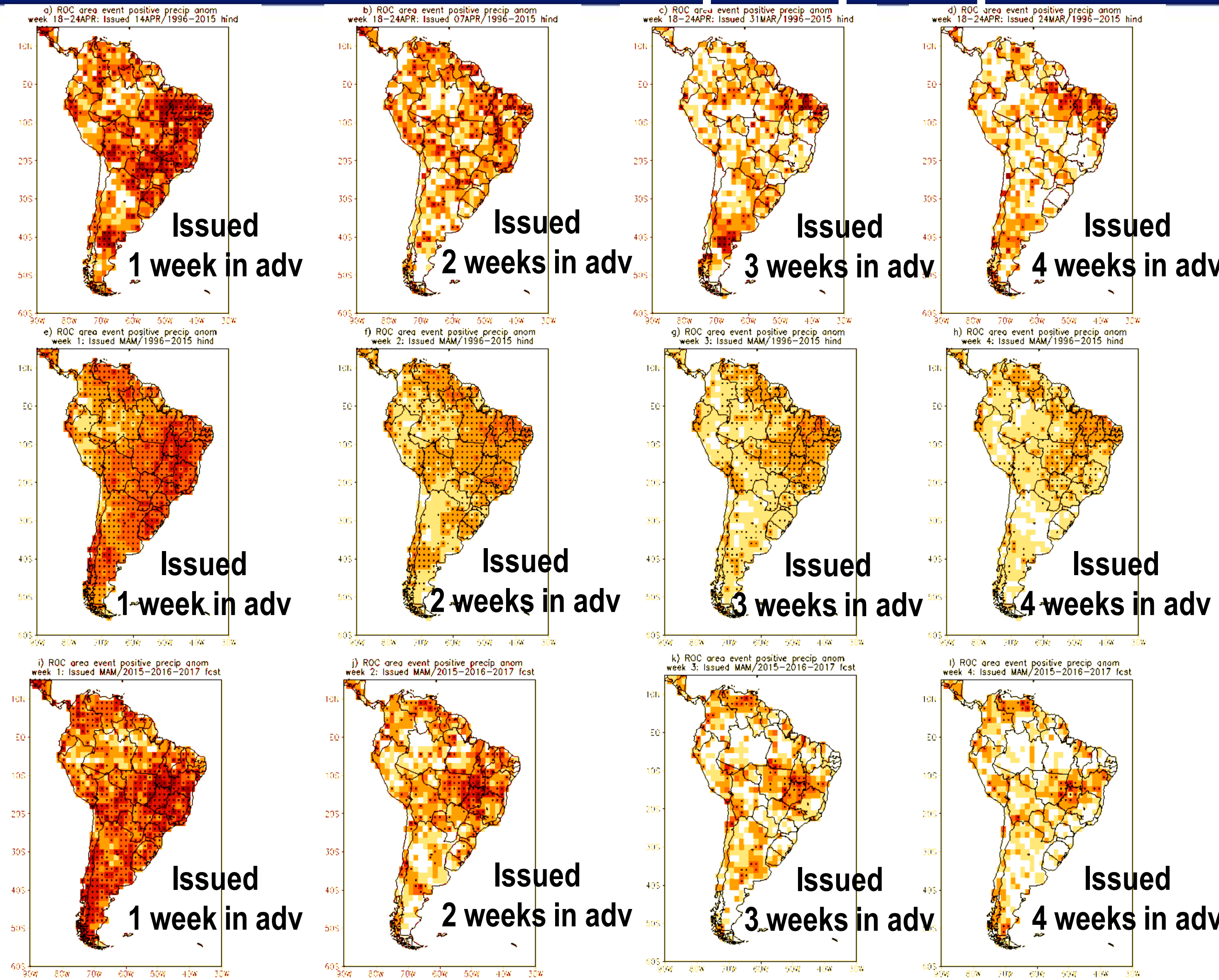
Murphy (1988):
MSSS incl phase error (correlation), mean error (bias) and amplitude error (ratio of pred ens mean to obs stdev)

Discrimination comparative assessment: Area under the ROC curve for event pos. precip. anom.

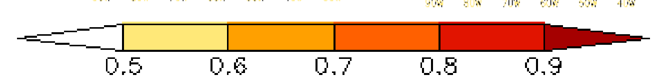
Level 1:
Target week
hindcast verification

Level 2:
All season
hindcast verification

Level 3:
All season
near real time
forecast verification



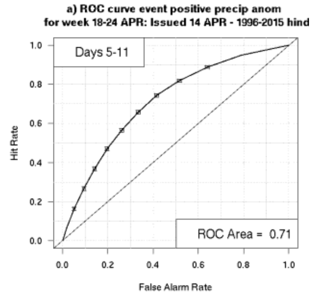
Discrimination: ability to distinguish events from non-events



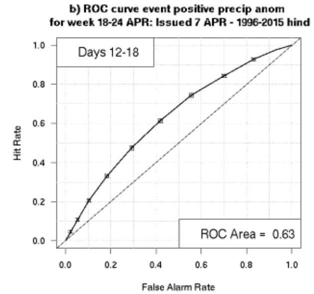
Discrimination comparative assessment: ROC curve for event pos. precip. anom.

www.cptec.inpe.br

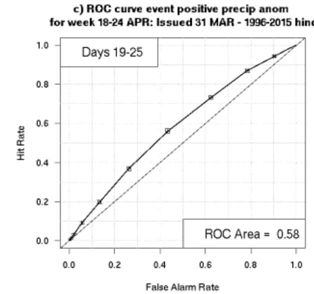
Level 1: Target week hindcast verification



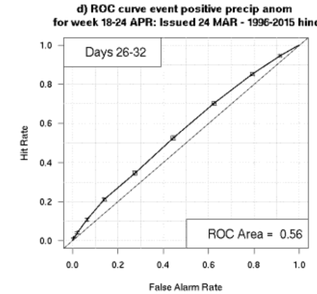
Issued 1 week in adv



Issued 2 weeks in adv

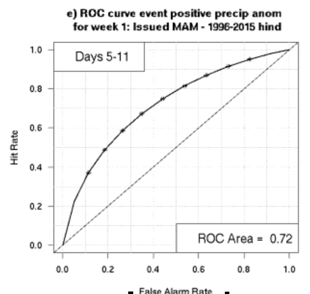


Issued 3 weeks in adv

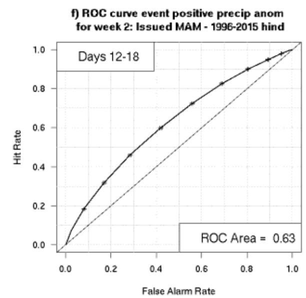


Issued 4 weeks in adv

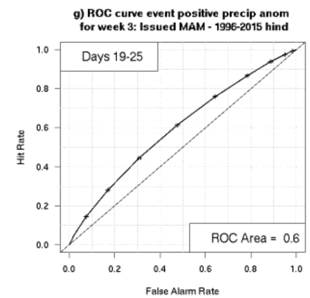
Level 2: All season hindcast verification



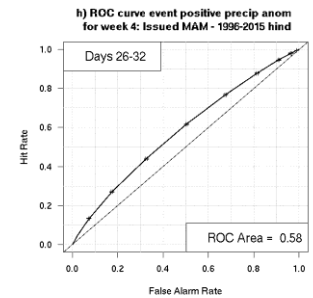
Issued 1 week in adv



Issued 2 weeks in adv

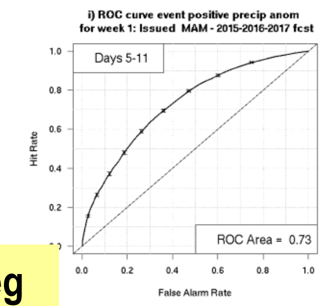


Issued 3 weeks in adv

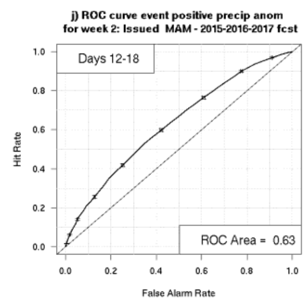


Issued 4 weeks in adv

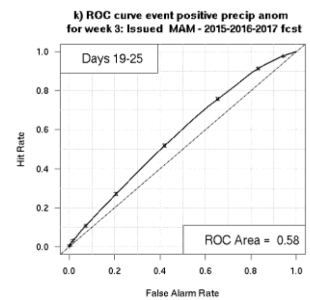
Level 3: All season near real time forecast verification



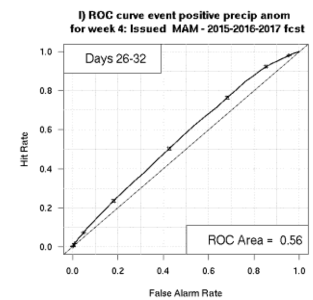
Issued 1 week in adv



Issued 2 weeks in adv



Issued 3 weeks in adv



Issued 4 weeks in adv

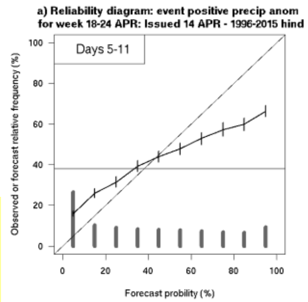
Overall discrimination: aggreg all hindcasts/forecasts in space and time

www.cptec.inpe.br

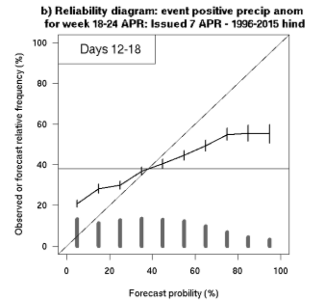
Reliability/Resolution comparative assessment: Reliability diagram for event pos. precip. anom.

Level 1:
Target week hindcast verification

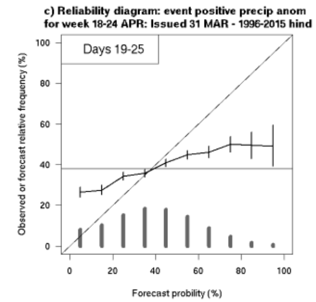
Reliab: how well calibr. issued probs. are
Resol.: how the freq of occurr of the event differs as issued prob changes



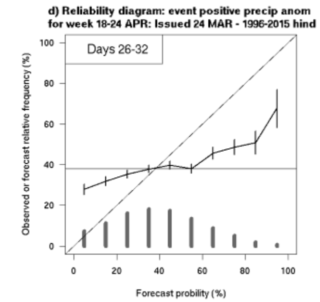
Issued
1 week in adv



Issued
2 weeks in adv



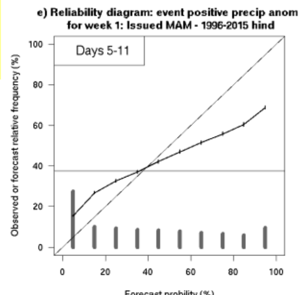
Issued
3 weeks in adv



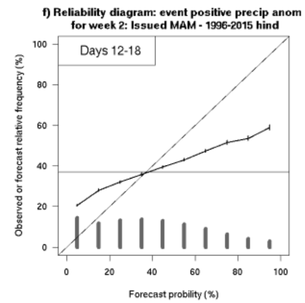
Issued
4 weeks in adv

Level 2:
All season hindcast verification

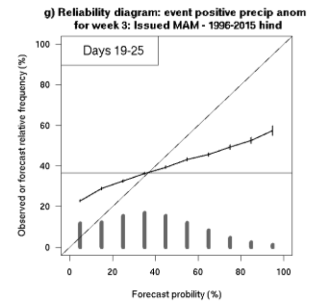
Assessment reveals need for calibration



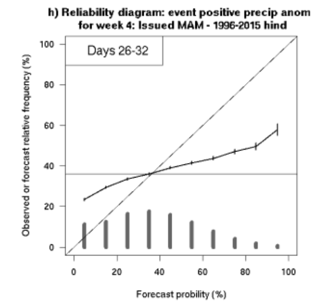
Issued
1 week in adv



Issued
2 weeks in adv



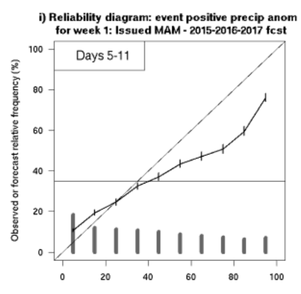
Issued
3 weeks in adv



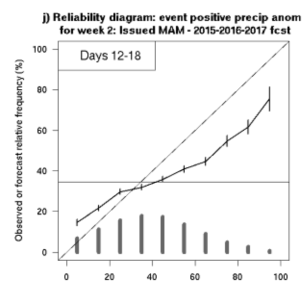
Issued
4 weeks in adv

Level 3:
All season near real time forecast verification

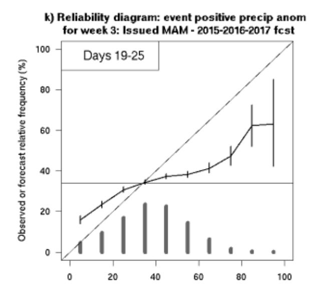
Aggreg all hindcasts/forecasts in space and time



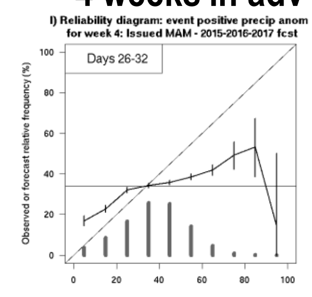
Issued
1 week in adv



Issued
2 weeks in adv



Issued
3 weeks in adv



Issued
4 weeks in adv

Summary

• Proposed a necessary verification framework for sub-seasonal precip. predictions based on a three level strategy according to the available hindcasts and near real time forecasts of recent past years characterized by a large degree of complexity/differences

• Most fundamental prediction quality attributes were assessed: association, accuracy, discrimination, reliability and resolution

• Verification information provided in the three levels found to be consistent and complementary and when used together with forecasts help a) forecasters building confidence in the model forecast guidance information when issuing sub-seasonal forecasts (by addressing various questions), and b) model developers/forecasters indentifying forecast aspects in need of improvement

• Probabilistic assessment aggregating all predictions over South America revealed modest discrimination ability, with predictions clearly requiring calibration for improving reliability and possibly combination with other model predictions for improving resolution

• C.A.S. Coelho, M.A.F. Firpo, F.M. de Andrade , 2018: A verification framework for South American sub-seasonal precipitation predictions. Meteorologische Zeitschrift
DOI: 10.1127/metz/2018/0898

CPTEC Centro de Previsão de Tempo e Estudos Climáticos

CPTEC / INPE

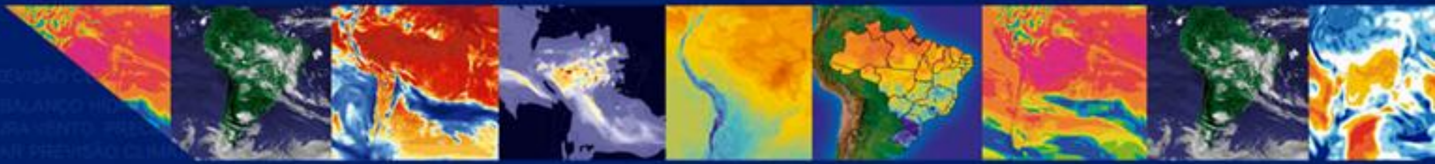
www.cptec.inpe.br

Thank you for your attention!

Caio A. S. Coelho

caio.coelho@inpe.br

Ciência e Tecnologia a serviço da sociedade



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Ministério da Ciência, Tecnologia e Inovação

BRASIL GOVERNO FEDERAL