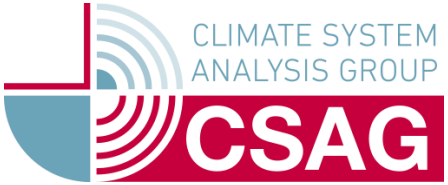
An initiative of the Organisation of African, Caribbean and Pacific States funded by the European Union



Moving SARCOF forecast towards objectivity

Piotr Wolski

Climate System Analysis Group, University of Cape Town

colleagues from ClimSA project (Atkins Realis and SADC CSC):

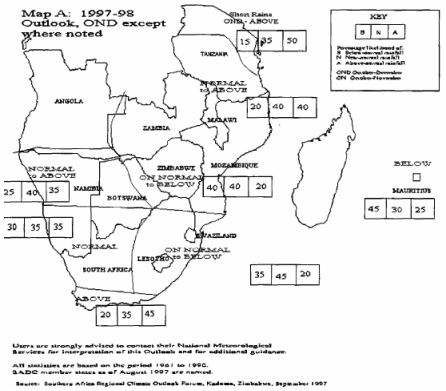
Tiro Nkemelang, Francis College, Sunshine Gamedze

Forecast Verification Workshop, Cape Town, May 2024

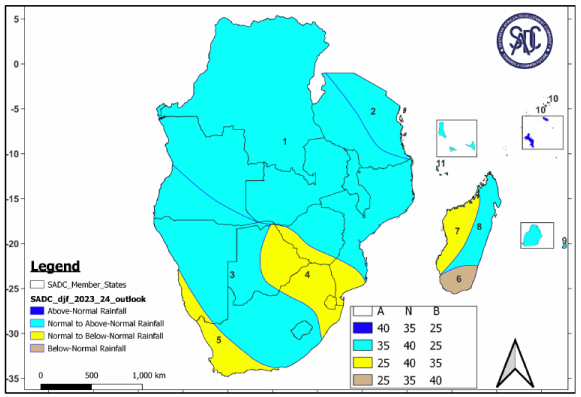
SARCOF forecasts

- SARCOF - Southern African Climate Outlook Forum
- **Regional Climate Outlook Forum** - WMO-mandated process intended to:
  - produce **consensus-based**, user-relevant climate outlook products in real time
  - in order to reduce climate-related risks and support sustainable development for the coming season
  - in sectors of critical socio-economic significance
  - for the region in question.
- Training of forecasters towards the application of forecast methods and development of the forecast is an important element of RCOF
- Creates a platform for interactions with users of climate forecast
- **SARCOF is the first RCOF established in Sept 1997 in Harare, Zimbabwe issuing outlook for SADC.**

1997

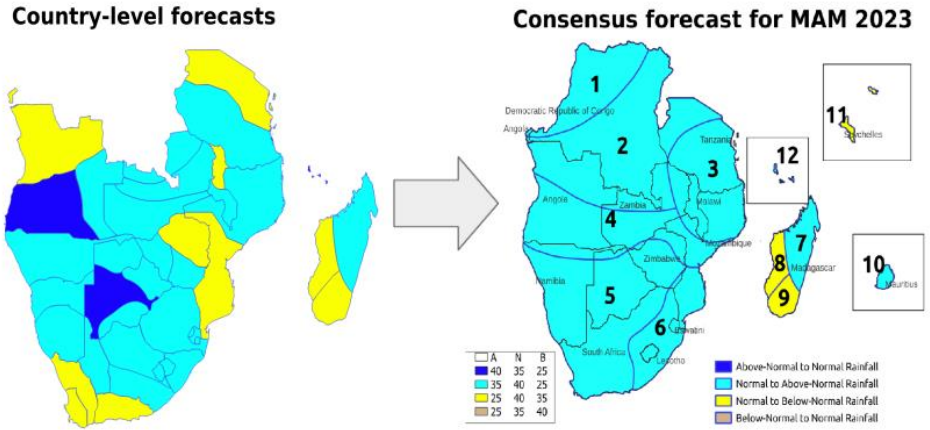


2023



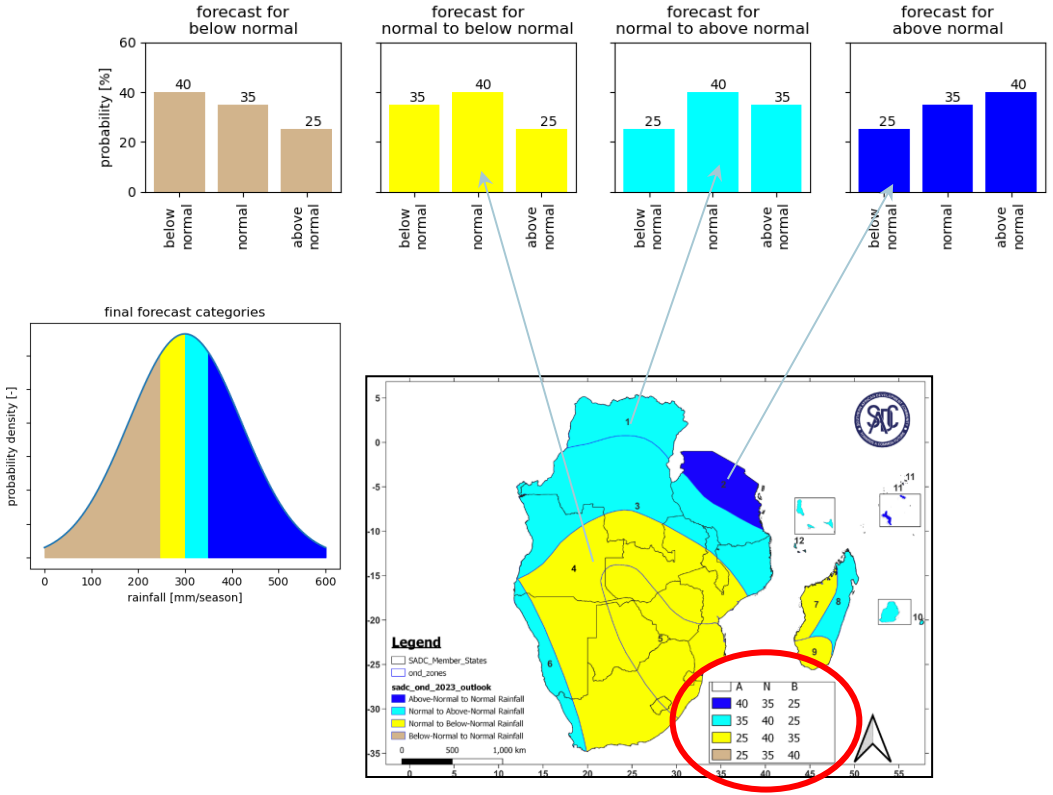
Forecast process

- forecast at country level derived for country zones, using statistical approaches, using multiple forecasts and multiple predictors
- integration of multiple sources into final country forecast based on expert opinion, and thus not transparent
- process of “negotiation” between countries to derive forecast for regional zones from zonal forecast at country level
- in practice - skill of individual sources of country level forecast is rarely taken into account
- process is **not transparent, does not follow any pre-defined rules**
- strongly dependent on “who is in the room” and who speaks the loudest
- the need to fit large, pre-defined “homogeneous” zones make the process more difficult



What does SARCOF forecast map mean?

- Four (4) classes are used. In “deterministic” interpretation, the “normal” class is split into two:
  - normal to below normal
  - normal to above normal
- this is intended to avoid forecasting normal and indicating most likely direction of the forecast
- The probabilities of the **tercile** forecast associated with each of the **four final classes are “fixed” or pre-determined**, and are not dependent on the actual outcome of the empirical model.



In summary…

- SARCOF forecast is essentially **a subjective expert interpretation** of results of a statistical forecast at the level of individual stations
- SARCOF forecast is generated in a process that is **not transparent and most likely not reproducible**
- The use of large regional pre-determined "homogeneous" zones causes likely ignoring of "sub-zonal" level signal (there is no assurance that signal within the regional zones in a particular season is identical)
- The large zones might make sense, if the forecast is used at regional level, but from our experience - it is **interpreted locally, even by regional organizations.**
- The form of the forecast (4 categories that are linked to probabilistic tercile forecast) appears very confusing to users (we suspect, but we don't know, because there is no data/research on this).
- SARCOF forecast has no skill or confidence information attached

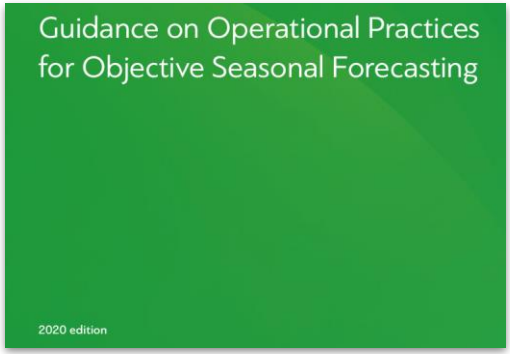
but:

- **the process appears to have developed a tightly-knit community of regional forecasters**
- **it facilitates peer-learning and lateral transfer of knowledge**
- **it empowers the forecasters and capitalizes on their expertise, thus creating a sense of ownership of the product**
- **SARCOF forecast has a very strong mandate in the region**

Objective Seasonal Forecast

OSF principles favour dynamical forecasting models

1. Follow a traceable, reproducible, and well-documented procedure (including model selection, bias correction, calibration and statistical downscaling) that is amenable to assessments of forecast quality (verification);



but there is still room for statistical approaches:

At present, for some regions and/or seasons, dynamical models may not be as skilful as statistical (or empirical) models. In such cases, in order to provide the best seasonal forecast information, an objective blend of empirical (statistical) and dynamical models may be used. However, statistical models should be properly trained to avoid overfitting. They should also be cross-validated to adequately assess their past performance (and to ensure that they adhere to other components in the set of principles, for example, the use of a well-documented forecast procedure). The use of statistical models should only be considered an interim measure given the routine advances in dynamical models.

Verification is a necessary prerequisite towards OSF:

5. Ensure that forecasts are verified according to established standards, keep archives of past forecasts, and conduct post-season assessments;

6. Provide forecast information together with historical performance (for example, skill and reliability);

Constraints in the process of introducing OSF at SARCOF

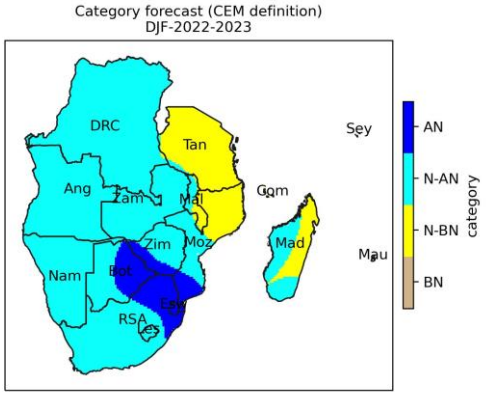
- country experts express reluctance to abandon consensus process and adopt dynamical models, which they feel would reduce their role to post-processing of data generated elsewhere
- forecast exists in a political context, there is a reluctance to:
  - forecast normal (perceived as non-informative, thus the “unusual” four categories)
  - forecast climatological odds (“if we say we don’t know - we will lose our jobs”)
  - admit low skill of forecast (as above)

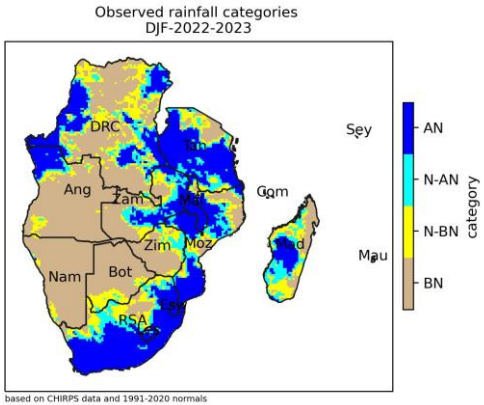
Moving SARCOF towards OSF

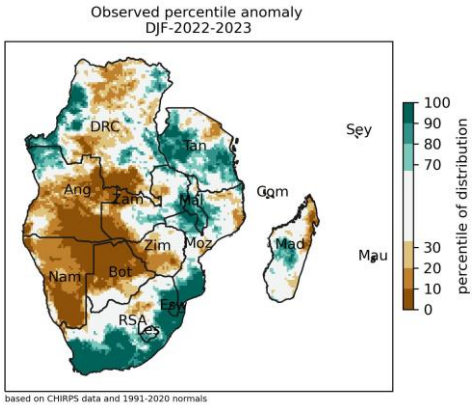
Under ClimSA project the following were introduced:

- verification of a single forecast as per WMO recommendations
- evaluation of skill of previous SARCOF forecasts
- a transparent process of synthesis of multiple forecasts at country level that includes articulation of confidence in the forecast (although no explicit skill)
- regional consensus process that:
  - moves away from “homogenous” zones and adopts zones that follow “signal”
  - transparently includes confidence in the “consensus” forecast at the country level

SADC CFT verification tool - visual verification







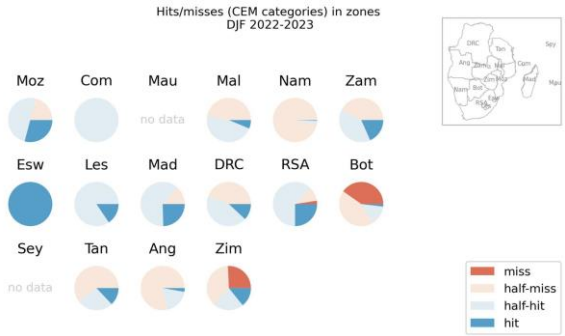
Visual interpretation of observations vs. forecast:

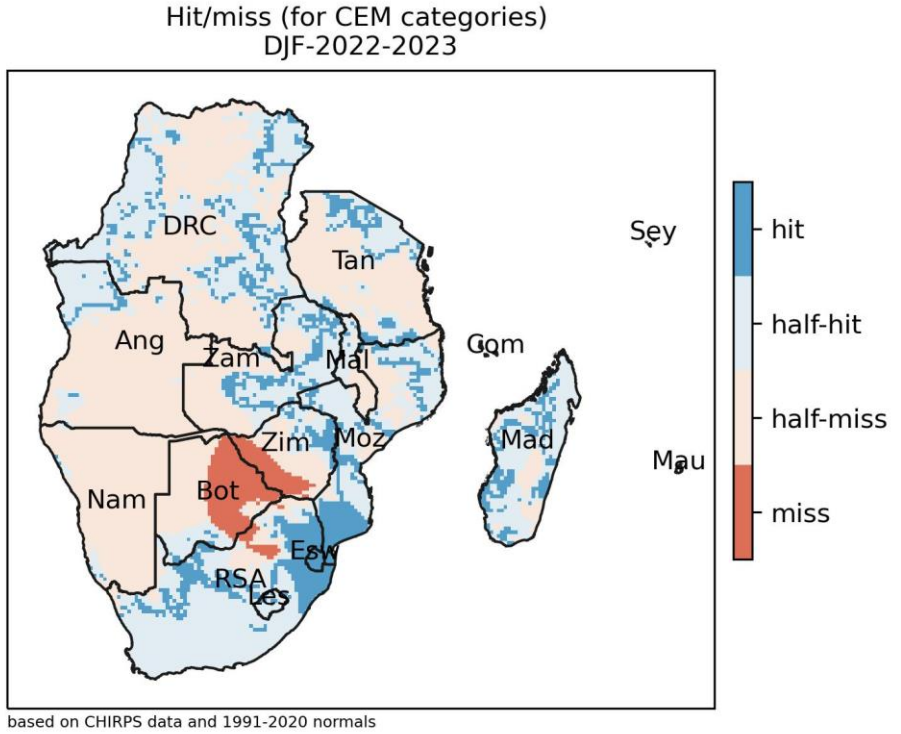
- observations presented in identical categories/colours to those used in the forecast
- percentile anomaly, as recommended by WMO

SADC CFT verification tool - verification of four categories

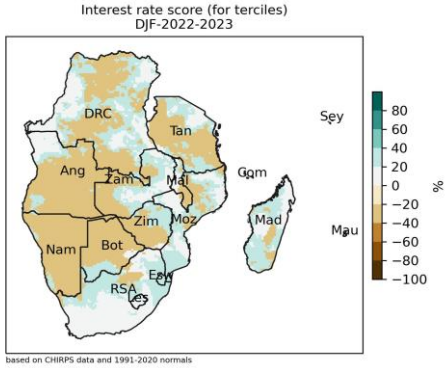
Evaluation of the four category forecast interpreted in deterministic way:

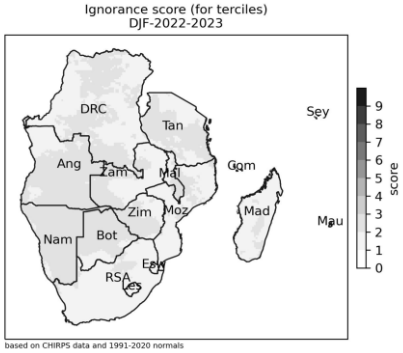
- hits/half hits/misses against observed categories
- presented as a map, but also as spatial percentage, which corresponds to interpretation of the forecast “locally”

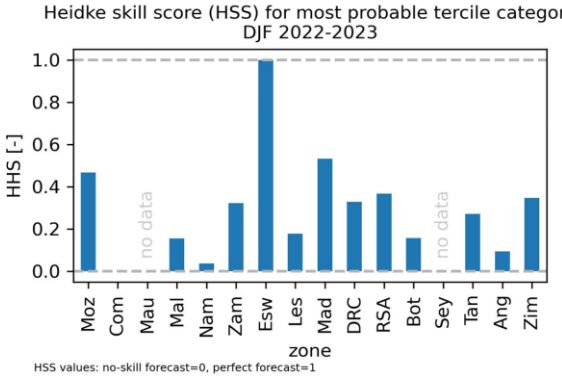


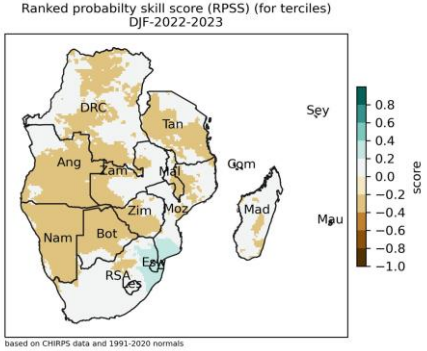


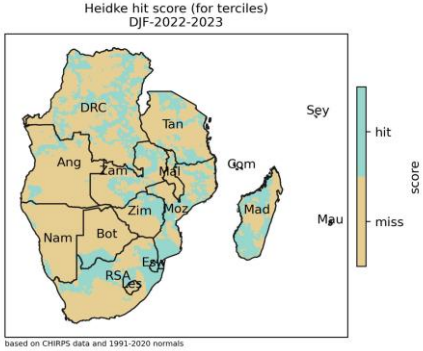
SADC CPT verification tool - probabilistic tercile forecast

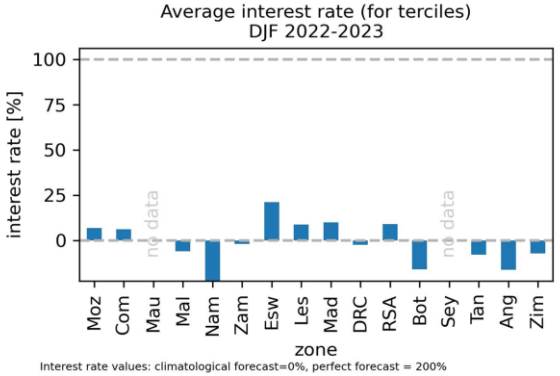






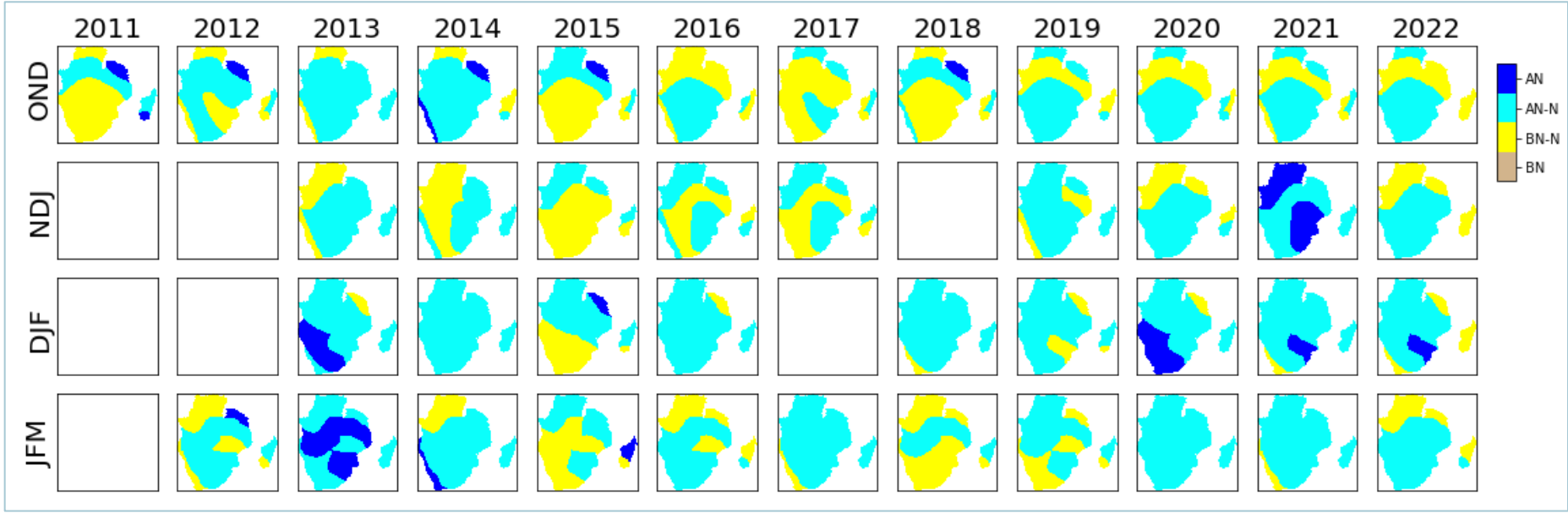






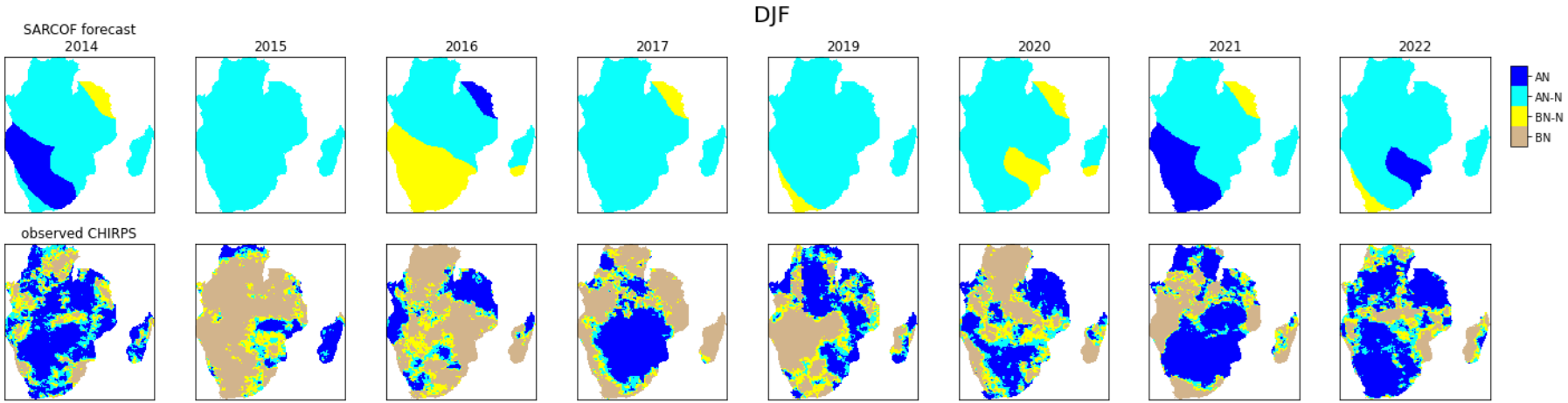
Previous SARCOF forecasts

note that Below Normal was never forecasted in the 2011-2022 period!



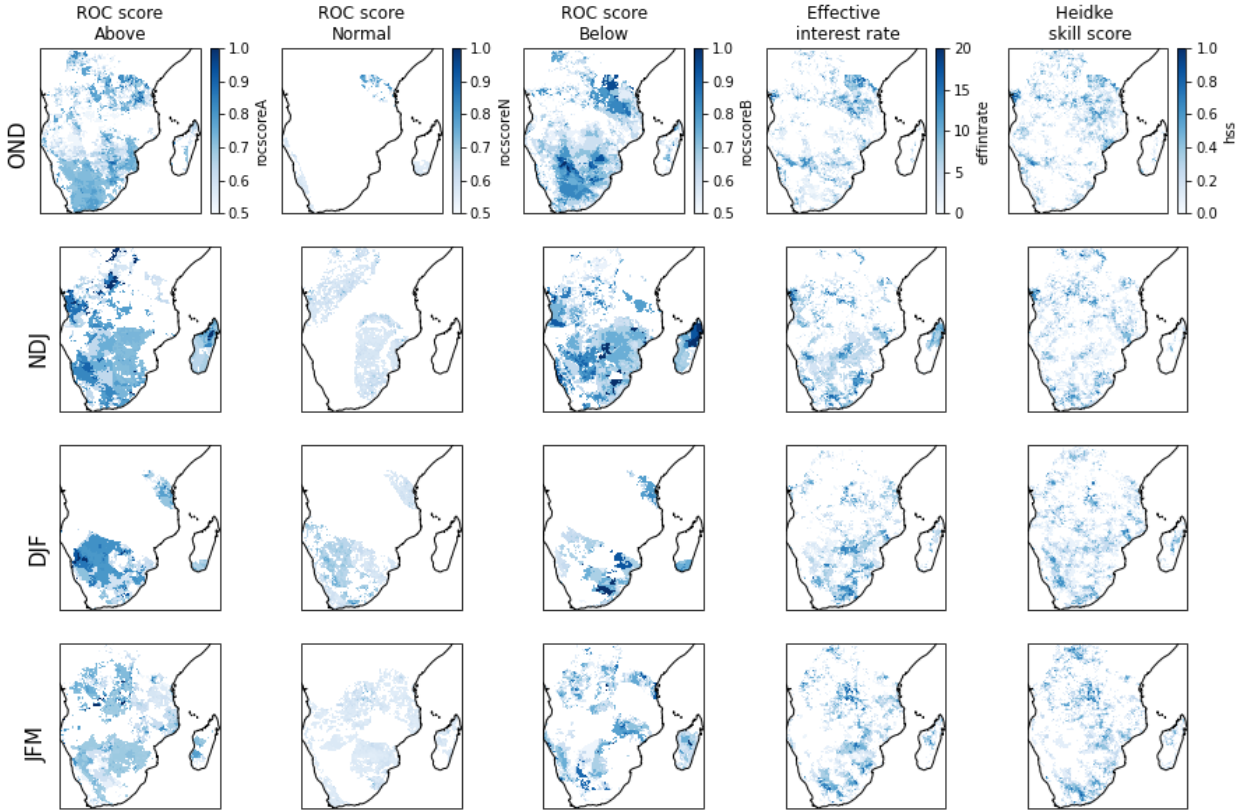
Forecast vs. observed - 4 categories

Note that the two central categories appear to be “rare” in observations, but dominant in the forecast



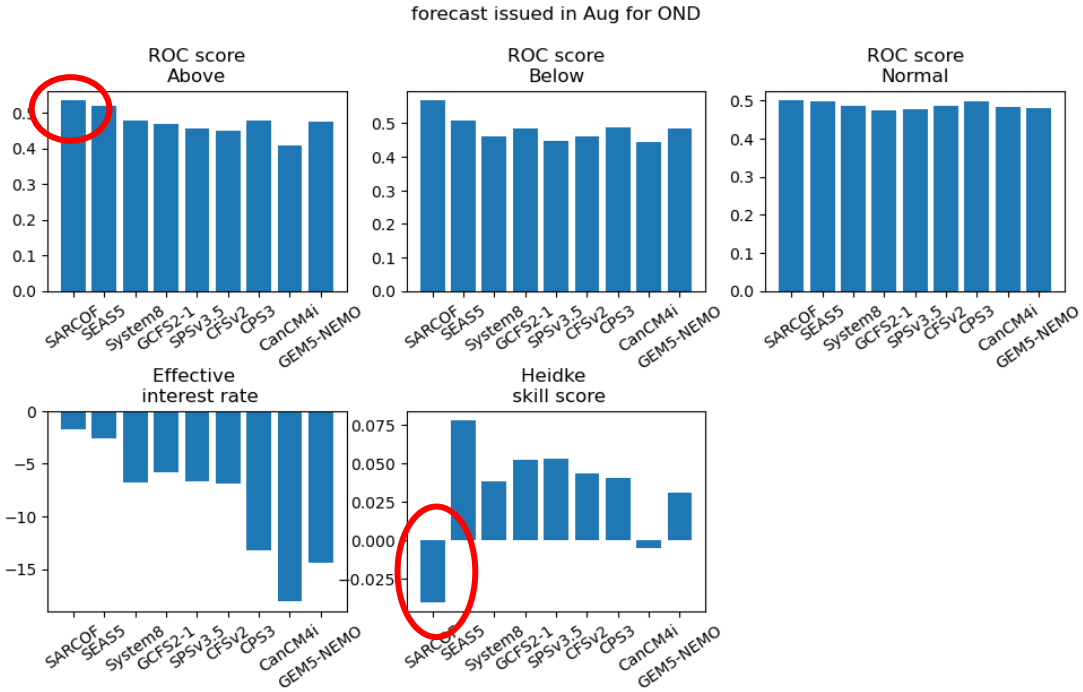
Skill of 2010-2022 SARCOF forecast (teciles)

- OND and NDJ appear to be skillful over Tanzania and central southern Africa for the Above and Below terciles, South Africa also in DJF for Above
- Elsewhere - and in terms of other skill indices, skill is “scattered”



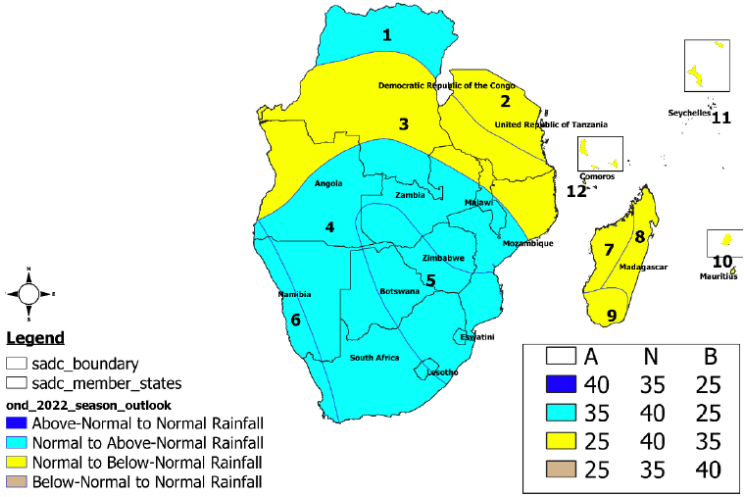
Region-average skill measures

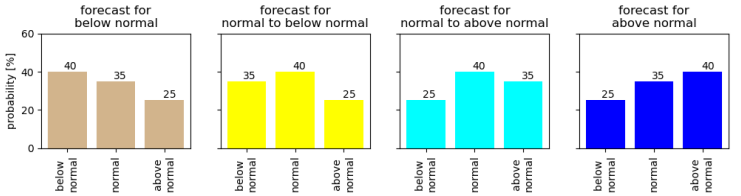
- SARCOF forecast has a better ROC score for above and below, as well as effective interest rate than (uncalibrated) dynamical forecasts.
- But it has a **considerably worse** Heidke skill score than those forecasts



“Appears better” but is it?

- **Do users use probabilistic information or do they interpret categories of forecast in deterministic way?**
- It appears that the probabilities are often ignored by users, with two implications:
  - on the one hand, probabilities are actually skillful, so ignoring them reduces utility of the forecast
  - on the other - probabilities are only skillful when considered “in full”, i.e. all three have to be considered. Interpreting SARCOF forecast as deterministic through the tercile with the maximum probability is basically skill-less.
- The simplistic interpretation of the forecast through its “face value” (“colours” with naive interpretation of category names) actually works!





Approach to synthesizing country level forecast

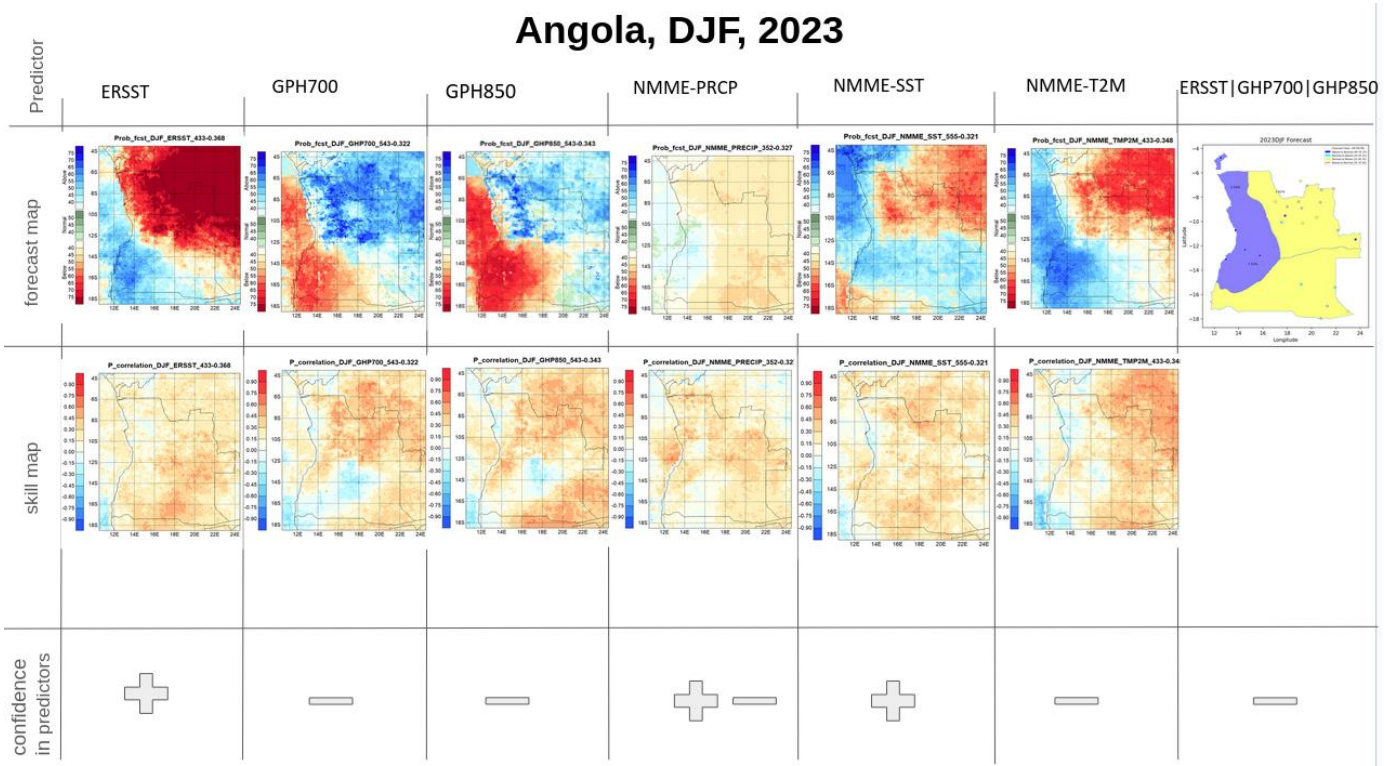
- set of rules towards determination of forecast signal and forecast confidence (proxy of skill) based on:
  - agreement between various sources of forecast/predictors
  - average skill of various forecasts
  - weighted to discount some predictors or approaches/tools

Proposed rule-base for country-level forecast from various sources

	skill		
agreement	all low	mixed	all high
low (categories diverge)	no signal/expert decision low confidence	where most skill low confidence	expert decision moderate confidence
moderate (there is some agreement)	expert decision low confidence	where most skill moderate confidence	expert decision moderate?/high confidence
high (all source in the same category)	as per agreement moderate confidence	as per agreement high confidence	as per agreement very high confidence

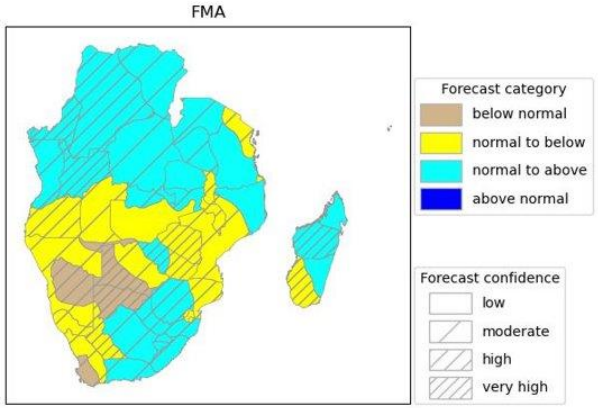
signal in red
confidence in blue

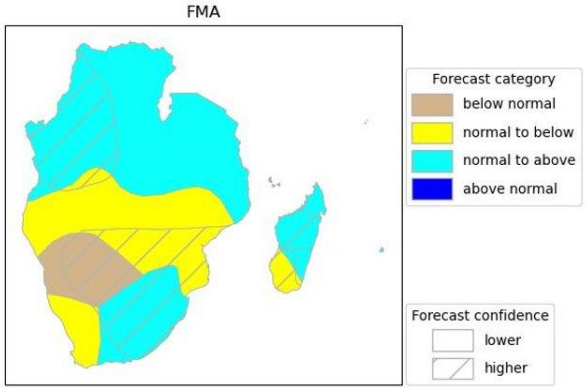
Country forecasts from a number of sources



Combining country forecasts into regional one

- four categories remain
- forecast zones follow signal, which creates fewer situations of “conflict” between countries that in the past had to be resolved in a process that was not transparent
- confidence (proxy for skill) is “inherited” from the country-level forecasts, but is expressed in two categories only





Summary and conclusions

- SARCOF forecast has been generated using methods basically developed in 1990s, with some alterations on the way
- it does not have the best “press”, for a number of reasons
- it has skill, but only when interpreted in a particular, deterministic “naive” way (which is most likely how it is interpreted by users)
- verification tool introduced that allows for defensible evaluation of a single forecast
- system of rules introduced to the country level forecast process that provides an expression of confidence
- the expression of confidence is included in the final product
- forecast confidence is a vague concept, but the best that could be achieved within the current framing of SARCOF forecast in order to demonstrate to the users some information about skill of the forecast