Short-term climate extremes: prediction skill and predictability

Emily J. Becker¹, Huug van den Dool¹, and Malaquias Peña²
1: Climate Prediction Center (NOAA/NWS/NCEP), 2: IMSG at Environmental Modeling Center (NOAA/NWS/NCEP)

CDPW - 23 October 2012 - Fort Collins, CO
How well can we currently predict short-term climate extremes?

- Short-term climate extremes (STCE): monthly or seasonal means well above or below the mean, at leads of 1 – 8 months.
- Present-day prediction skill and predictability
- 2-meter surface temperature and precip. rate over North and South America; sea-surface temperature in Niño 3.4 region and Atlantic hurricane main development region (MDR)
Forecasts and Observations

• Climate Forecast System version 2 ensemble mean reforecasts
  – 9-month leads, 1982-2010
  – all 12 initial months
  – regridded to resolution 1° x 1° (Saha et al. 2012)
• Tmp2m: GHCN+CAMS, regridded to 1° x 1° (Fan and van den Dool 2008)
• Precipitation rate: CPC global Unified Rain-Gauge Database, regridded to 1° x 1° (P. Xie et al. 2010).
• Sea-surface temperature: OI-2 (Reynolds et al. 2002), native resolution is 1° x 1°.
Methods

• Systematic error correction to remove model bias: model climatology is removed, and replaced with climatology from observations
• Cross-validation: “CV3RE”
• Verification measures:
  – Anomaly correlation (AC)
  – Root-mean-square error (RMSE)
• AC and RMSE are area-averaged over North America, South America (tmp2m and prate) and Niño3.4 and MDR regions (SST).
The definition of extreme
The definition of extreme

• Gridpoint monthly mean anomaly above/below +/- 1.645 standard deviations
  – Approximately equivalent to 5\textsuperscript{th}/95\textsuperscript{th} percentiles.
  – Other definitions were tested, with similar skill findings.

• Two scenarios:
  – Extreme is forecast: did it come true?
  – Extreme was observed: was it forecast?
Tmp2m: North Amer.

- Sliding scale ranging from “all cases” included to 3.0xSD.
- “Mod” = skill assessed for cases where an extreme was forecast
- “Obs” = skill when an extreme was observed
- Permutation tests of 1,000 random subsets
Tmp2m: South Amer.

T2m skill is higher in South America than North America.
Tmp2m: North Amer.

- Skill is often a function more of target than lead.
- STCE defined here by +/- 1.645 std. dev. threshold.
Tmp2m: South Amer.

![Diagram of Tmp2m South America showing different cases and extractions.](image)
Prate: NA

- Precipitation is very difficult to forecast!
- Cross-validation affects lower scores the most, and leaves us with near-zero scores for precipitation rate.
Precipitation rate: South Amer.
SST: Niño3.4 Region

- Higher skill scores are not noticeably reduced by CV
- RMSE is essentially flat with increase in threshold defining extreme
SST: Niño3.4 Region

SST Nino3.4 Reg.

a) all cases

b) model extr.

c) all cases CV3RE

d) model extr. CV3RE

[Graph showing SST patterns over time, with color-coded lead time and target month axes.]
SST: Atlantic hurricane dev. region

SSTs in this region are important for hurricane forecasting.
SST: Atlantic hurricane dev. region
Potential predictability

• Test how well the model predicts itself, under “perfect model” assumption.
• Take one member from the ensemble of N members, verify “prediction” of ensemble mean of N-1 members against this member.
• Employ both CFSv1 and CFSv2, as (ideally) the predictability should not depend on the model.
Tmp2m predictability, North Amer.
Tmp2m predictability, South Amer.

g) CFSv2

h) CFSv2 extremes

e) CFSv1

f) CFSv1 extremes
Potential predictability: precip rate
Potential predictability: SST in Niño3.4

SST predictability

a) CFSv1

b) CFSv1 extremes

c) CFSv2

d) CFSv2 extremes
Potential predictability: SST in MDR

- SST predictability MDR reg.
  - e) CFSv1
  - f) CFSv1 extremes
- g) CFSv2
- h) CFSv2 extremes
Summary (1/2)

• Assessed model forecast skill of short-term climate extremes (STCE) in 2 m temperature, precipitation rate, and sea surface temperature using CFSv2, 1982-2010.

• Anomaly correlations for STCEs are routinely higher than for “all cases”.

• RMSE, which can be considered as noise in a signal-to-noise ratio such as the AC, grows more slowly than the threshold: the signal grows despite increased noise.
Summary (2/2)

• Cross-validation leads to lower ACs. ACs that are already low are affected the most: precipitation ACs are reduced to near zero, while SST ACs are largely unaffected.

• Highest potential predictability for T2m in the first half of the year (both So. and No. Amer.)

• Target months with higher predictability tend to have higher forecast skill, but forecast skill is generally lower than its potential.