Opportunistic mixture model for post-processing S2S temperature and precipitation forecasts using convolutional neural networks

ECMWF Machine Learning Workshop 2022
March 30th, 2022




David Landry
Data Scientist
CRIM

Jordan Gierschendorf
Data Scientist
CRIM

Arlan Dirkson
Post-doctoral researcher
Environment and Climate Change Canada (ECCC)

Bertrand Denis
Former ECCC, UQAM
Now retired

CRIM

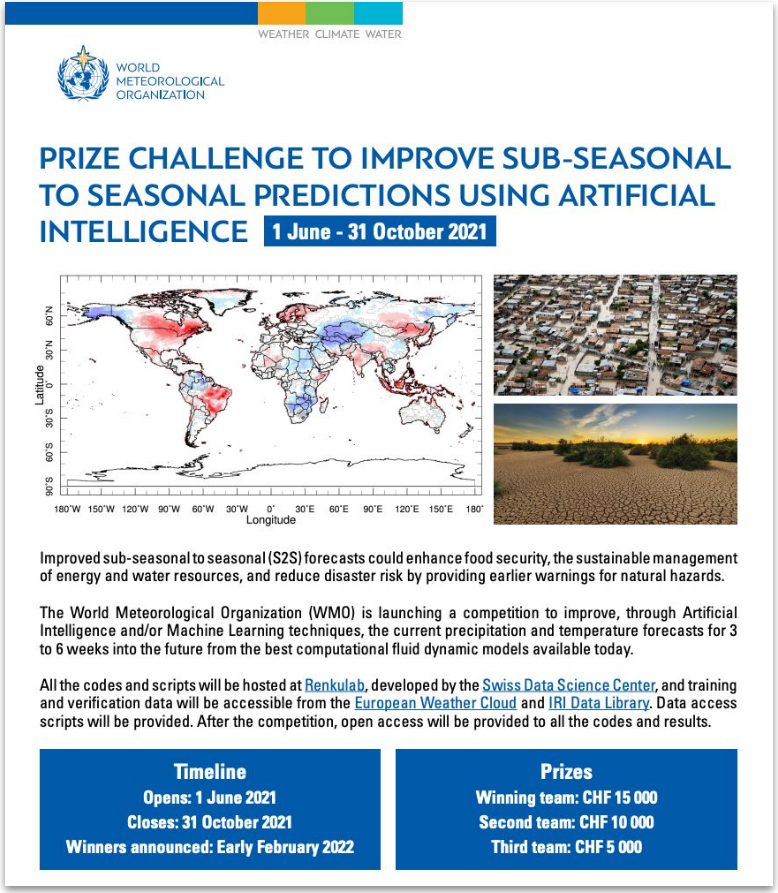
PRINCIPAL PARTERNAIRE FINANCIER
Québec

Overview

- CRIM et al. made a submission to the WMO S2S AI Challenge in 2021
- The object of the challenge was to improve sub-seasonal weather forecasts (3-6 weeks in the future) using AI
- The submitted model produced an opportunistic blend of models using a Convolutional Neural Network
- CRIM and collaborators obtained the first prize

S2S AI Challenge

- Prize money challenge to improve S2S forecasts using AI
- Produce terciled probabilistic forecasts for precipitation and temperature for weeks 3-6
- Use hindcast from ECMWF, ECCC and NCEP from ~2000 to 2019



WEATHER CLIMATE WATER

WORLD METEOROLOGICAL ORGANIZATION

PRIZE CHALLENGE TO IMPROVE SUB-SEASONAL TO SEASONAL PREDICTIONS USING ARTIFICIAL INTELLIGENCE **1 June - 31 October 2021**

Latitude

Longitude

Improved sub-seasonal to seasonal (S2S) forecasts could enhance food security, the sustainable management of energy and water resources, and reduce disaster risk by providing earlier warnings for natural hazards.

The World Meteorological Organization (WMO) is launching a competition to improve, through Artificial Intelligence and/or Machine Learning techniques, the current precipitation and temperature forecasts for 3 to 6 weeks into the future from the best computational fluid dynamic models available today.

All the codes and scripts will be hosted at Renkulab, developed by the Swiss Data Science Center, and training and verification data will be accessible from the European Weather Cloud and IRI Data Library. Data access scripts will be provided. After the competition, open access will be provided to all the codes and results.

Timeline
Opens: 1 June 2021
Closes: 31 October 2021
Winners announced: Early February 2022

Prizes
Winning team: CHF 15 000
Second team: CHF 10 000
Third team: CHF 5 000

Our approach

Design principles

We use our expert’s guidance to focus on three “design principles” that guide us through the conception of our submission

- The climatology baseline is strong (i.e. it’s difficult to be better than the uniform distribution)
- Model ensemble often perform better than single models
- Teleconnections can provide predictability (ex. El Niño) - there can be distant connections between weather patterns

Data analysis

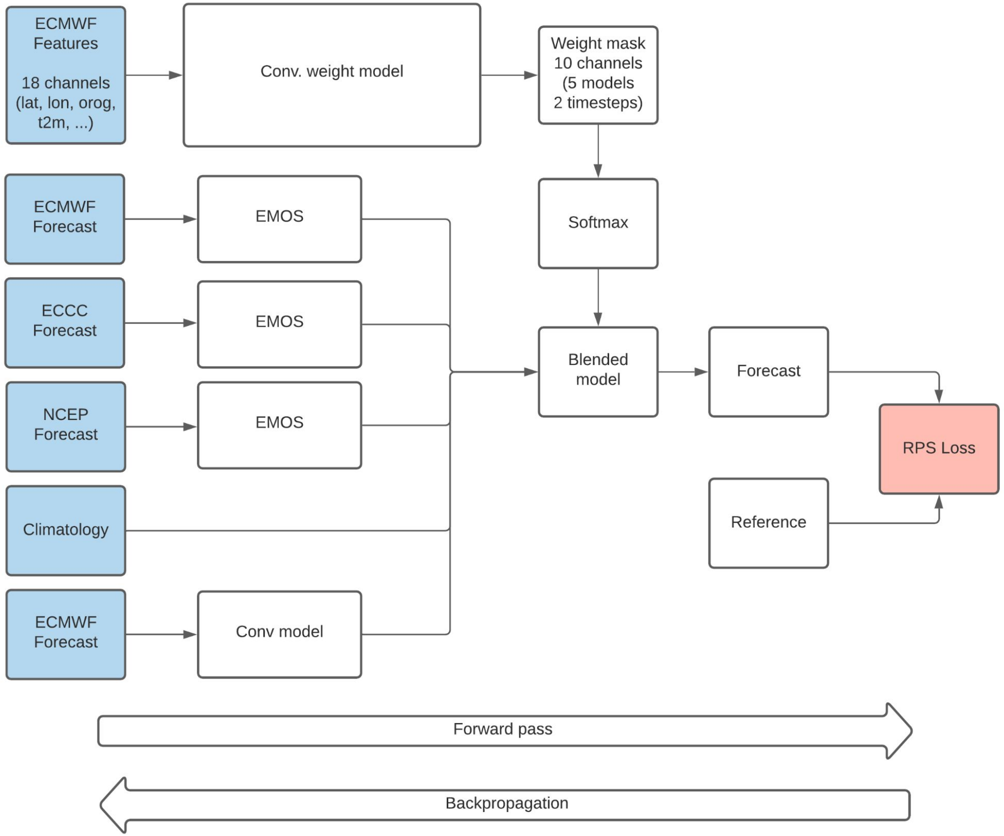
- Probabilistic hindcast from three forecasting centers
    - ECMWF (2000-2020) with 46 lead times
    - NCEP (2000-2010) with 44 lead times
    - ECCC (2000-2017) with 32 lead times (10 missing)
    - For all centers, a set of ~20 variables (slight variations between centers)
    - Temperature, Precipitation, Geopotential height, Soil moisture, Sea ice concentration, …
- **Analysis** grids for temperature and precipitation (2000-2020)

Predictors

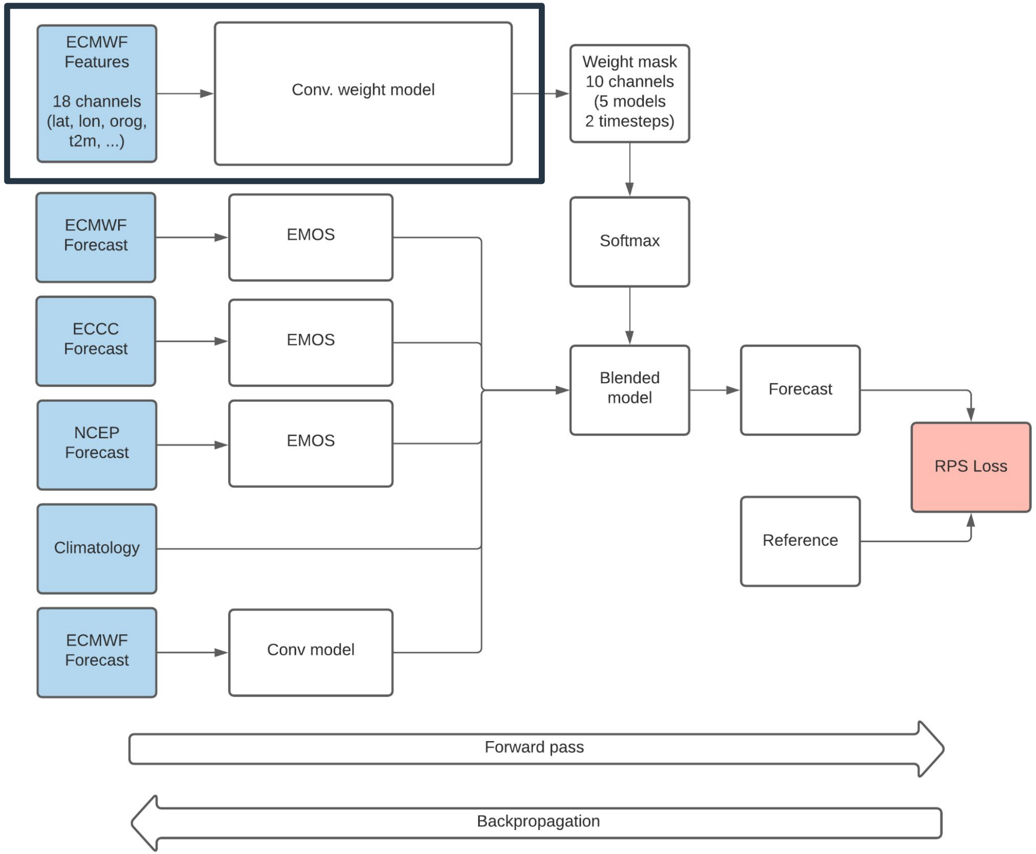
- Sea ice concentration
- Land-sea mask
- Mean sea-level pressure
- Soil moisture
- Sea surface temperature
- **Temperature @ 2m above ground**
- **Total precipitation**
- Wind @ 850, 200 hpa
- Geopotential height @ 1000, 500, 200 hpa

Non-model predictors

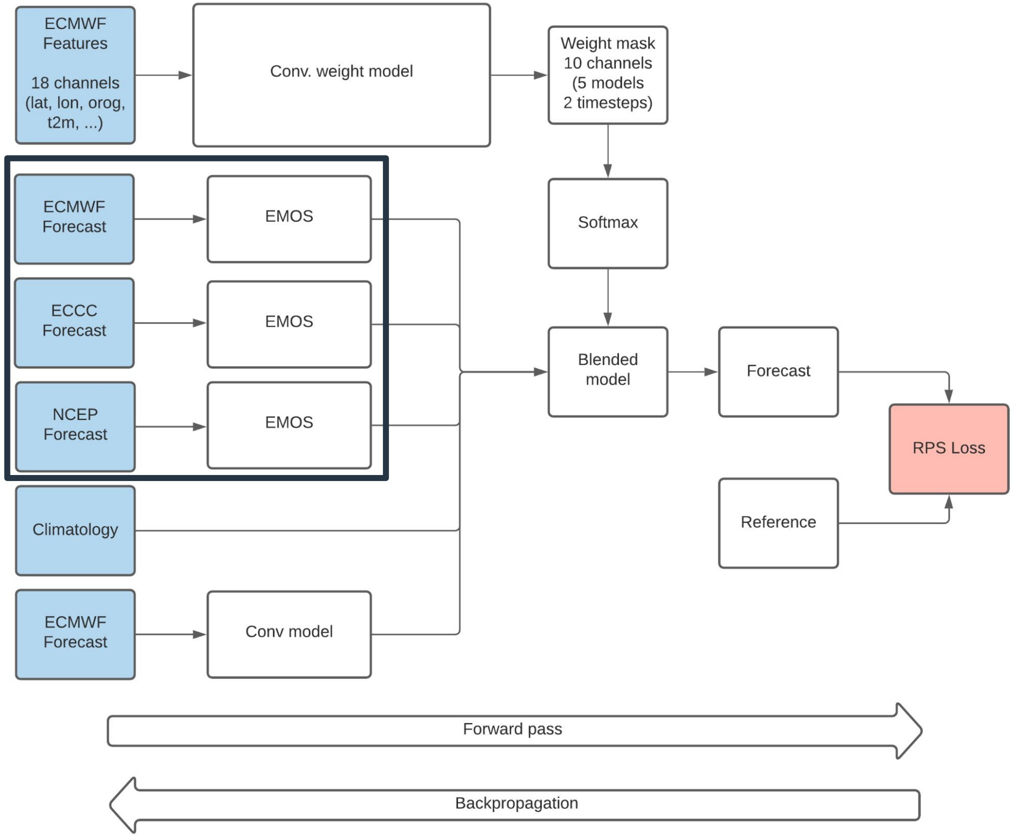
- Time of year
- Latitude and longitude
- Orography



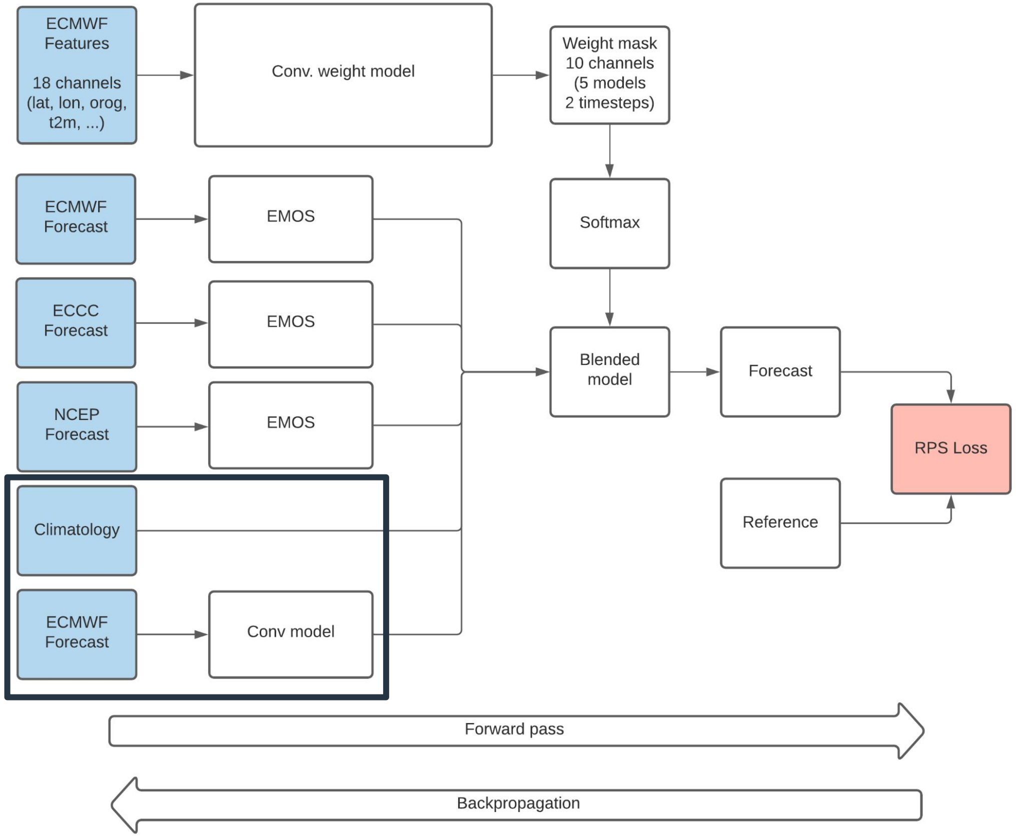
Model architecture



Model architecture



Model architecture



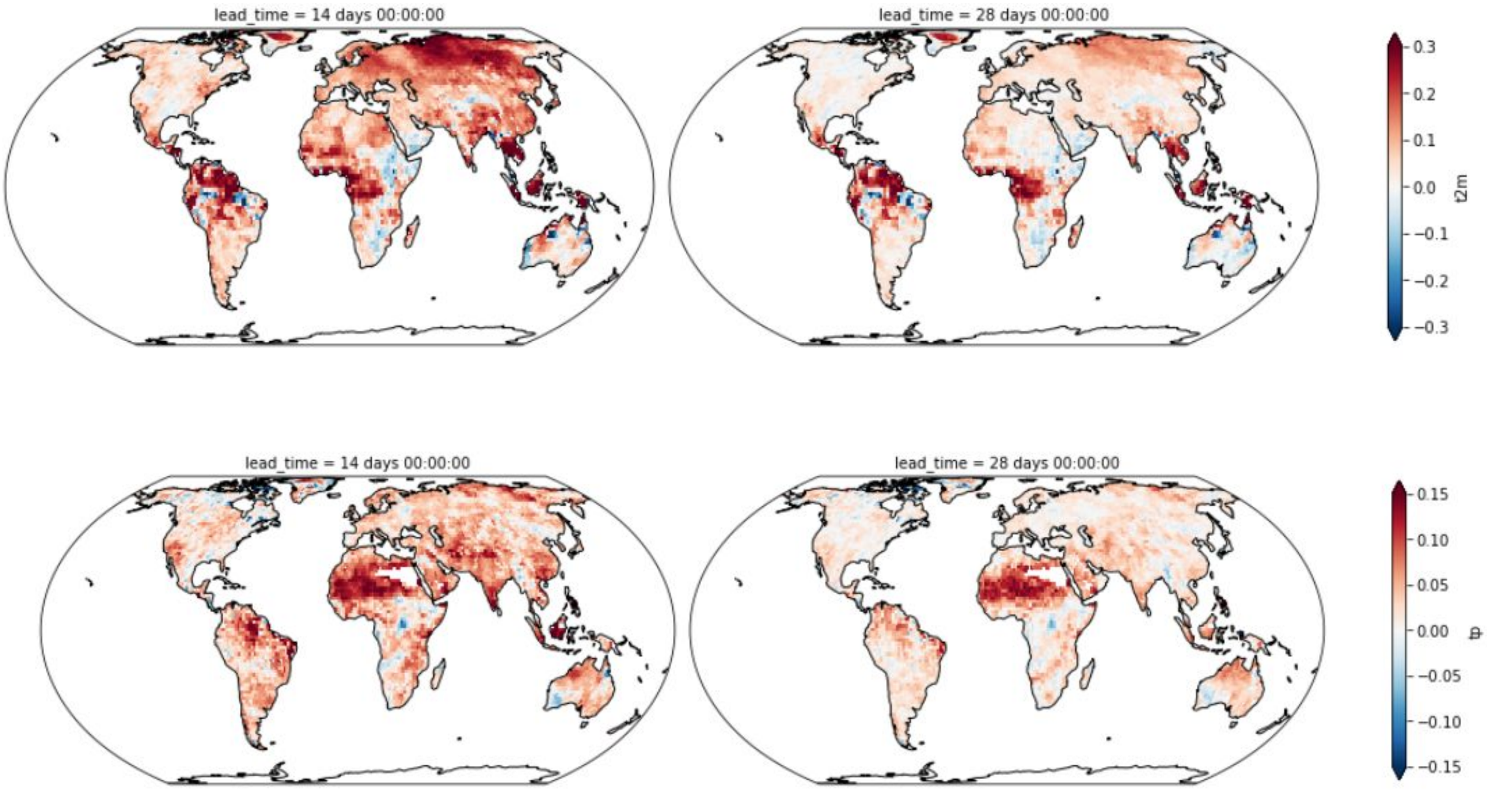
Model architecture

Results

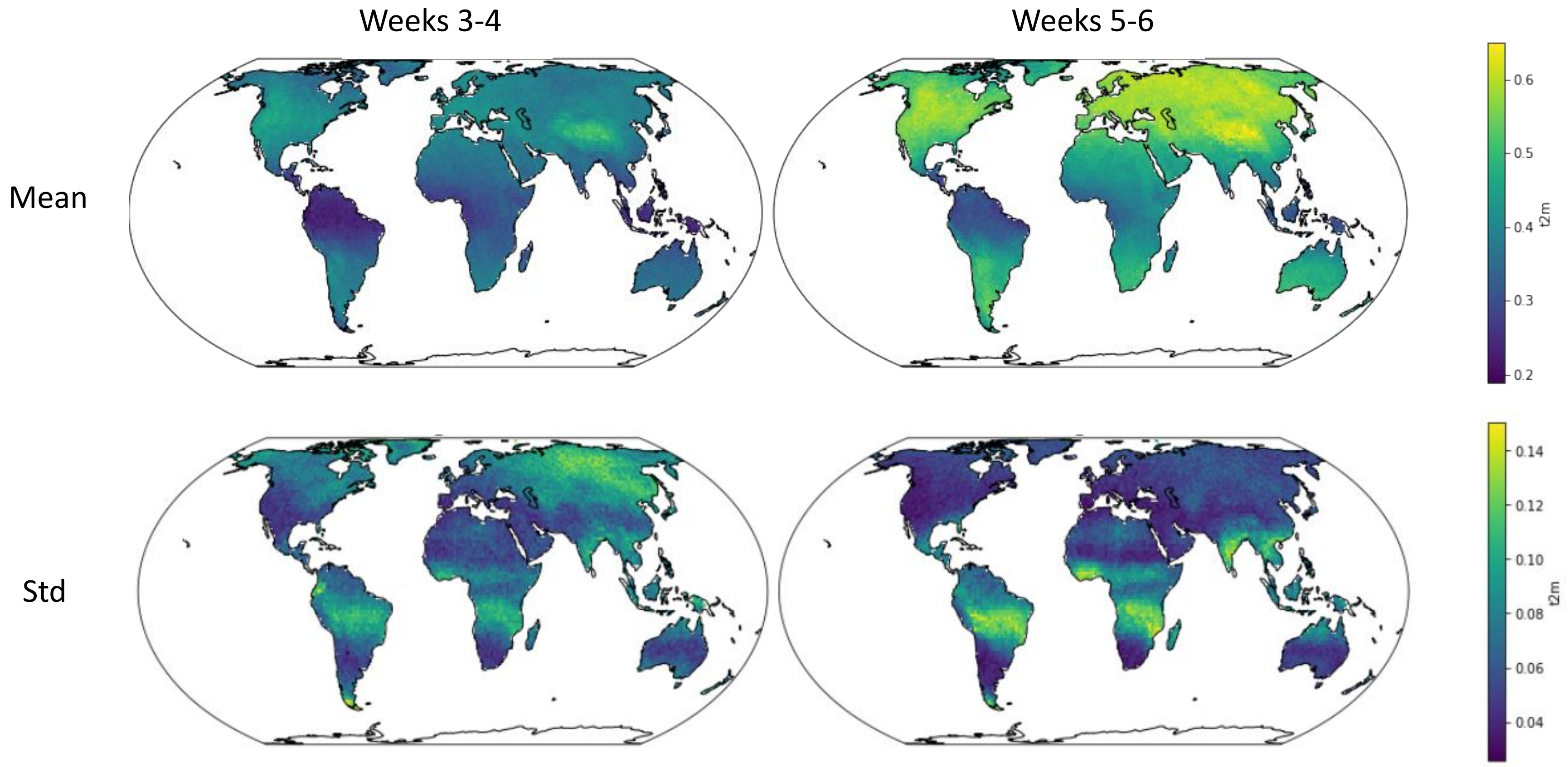
Leaderboard

Position	Team	Affiliation	RPSS
1	**CRIM S2S**	**CRIM**	**0.045**
2	BSC	Barcelona Supercomputing Center	0.028
3	UConn	University of Connecticut	0.006

+2 more submissions with RPSS > 0



RPSS for temperature and precipitation forecast



Mean weight of climatology - Surface temperature

Model	Mixture model	Forecasting Center	RPSS (2020)
EMOS	N/A	ECMWF	0.032
		NCEP	-0.021
		ECCC	-0.021
Conv	N/A	ECMWF	0.020
EMOS + Clim	Mean	ECMWF	0.042
		All*	**0.049**
	Conv	ECMWF	0.040
		All*	0.048

*Two versions of ECMWF: One corrected via EMOS, the other corrected by a Conv. Network

Regression study

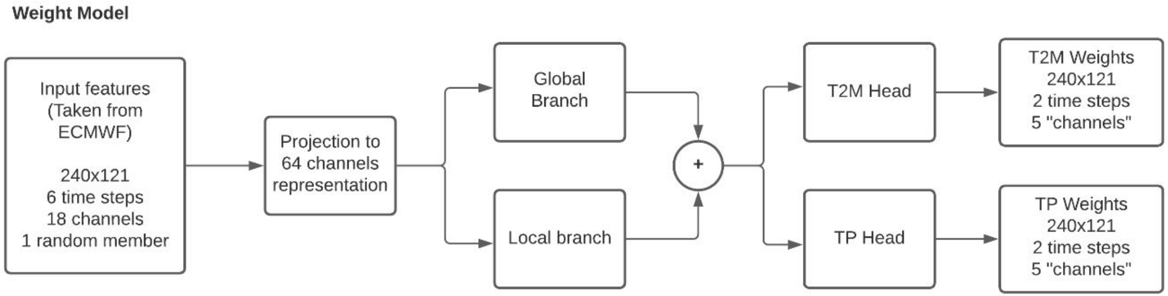
Open questions

- Can we demonstrate that CNNs are necessary for this task?
- Did our global branch capture teleconnections?
    - Could we infer important known indices such as the Madden Julian Oscillation from the embedding?
- Could a transformer architecture do a better job at capturing teleconnections?

Ressources

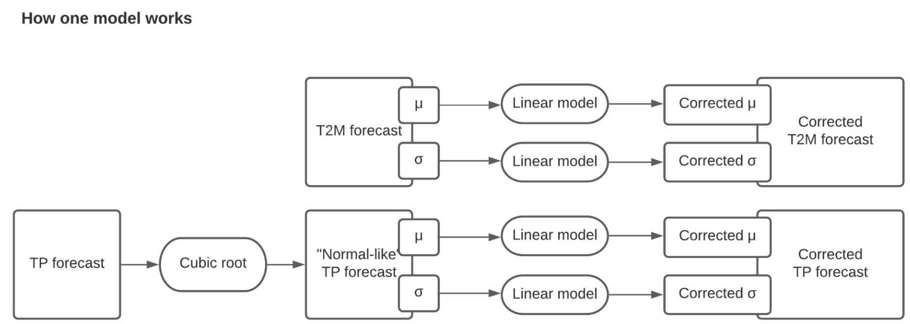
- Challenge website
  - https://s2s-ai-challenge.github.io/
- Github to our contribution
  - https://github.com/crim-ca/crims2s
- Description of our contribution destined to the challenge judges
  - https://renkulab.io/gitlab/jordan.gierschendorf1/s2s-ai-challenge-template/-/issues/2
- Review paper to be submitted to BAMS (Frederic Vitart et al.)

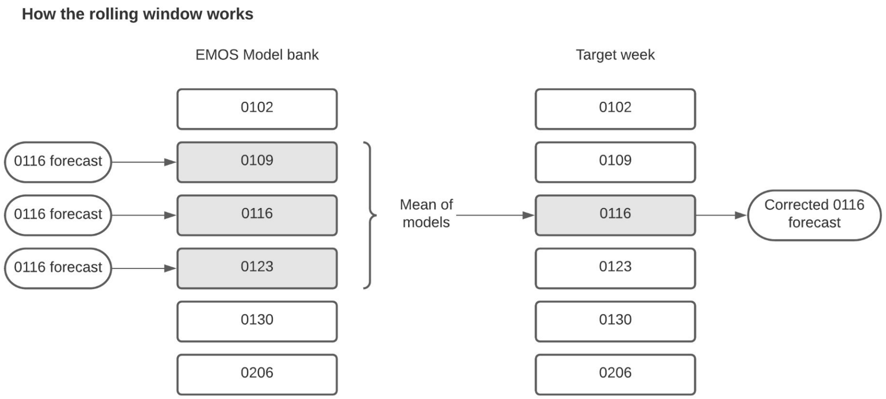
Any questions?

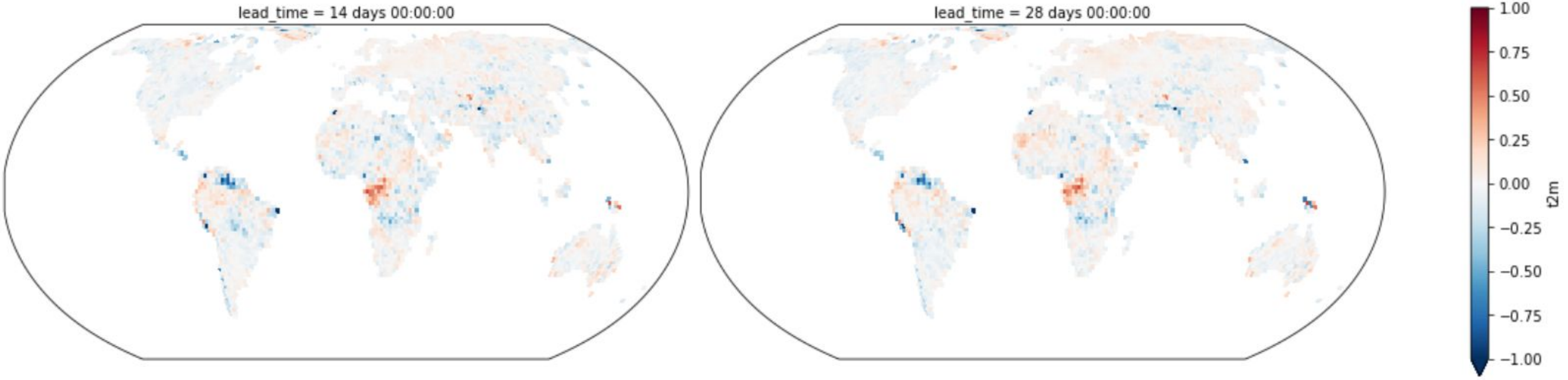


Blending model architecture

EMOS Model Architecture







Comparison of Conv weight model skill vs mean model. T2M, 2020.