Identifying relevant large-scale predictors of sub-seasonal precipitation using explainable AI

September 2022 | EMS22 | Bonn, Germany
Niclas Rieger, Laura Ferranti, Linus Magnuson, Álvaro Corral, Estrella Olmedo, Antonio Turiel
Mathematical Research Centre (CRM), Barcelona, Spain




Where to look for S2S predictability of precipitation?

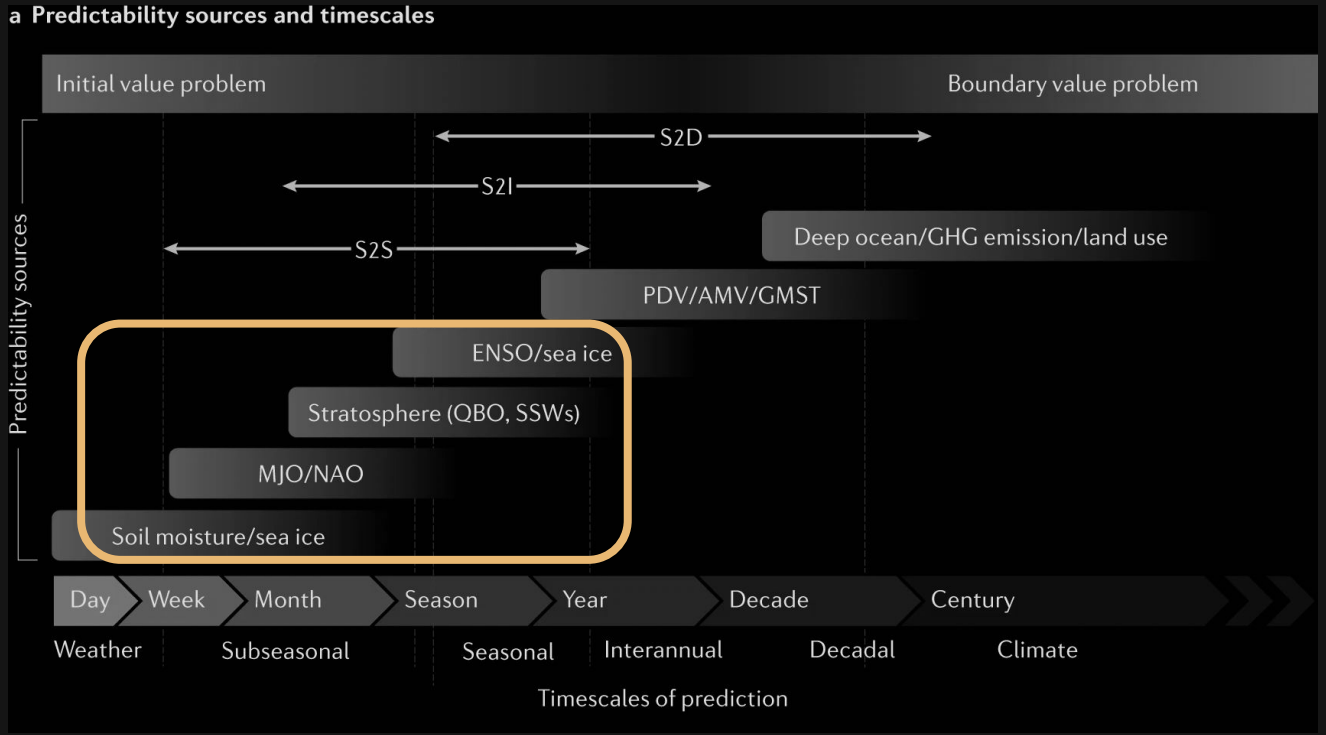
Explainable AI has successfully been used to identify relevant predictors on sub-seasonal time scales

For example:

Summer temperatures (Straaten et al. 2022)

Midwest precipitation (Carter & Herrera 2022)

Stratospheric polar vortex (Wu et al. 2022)



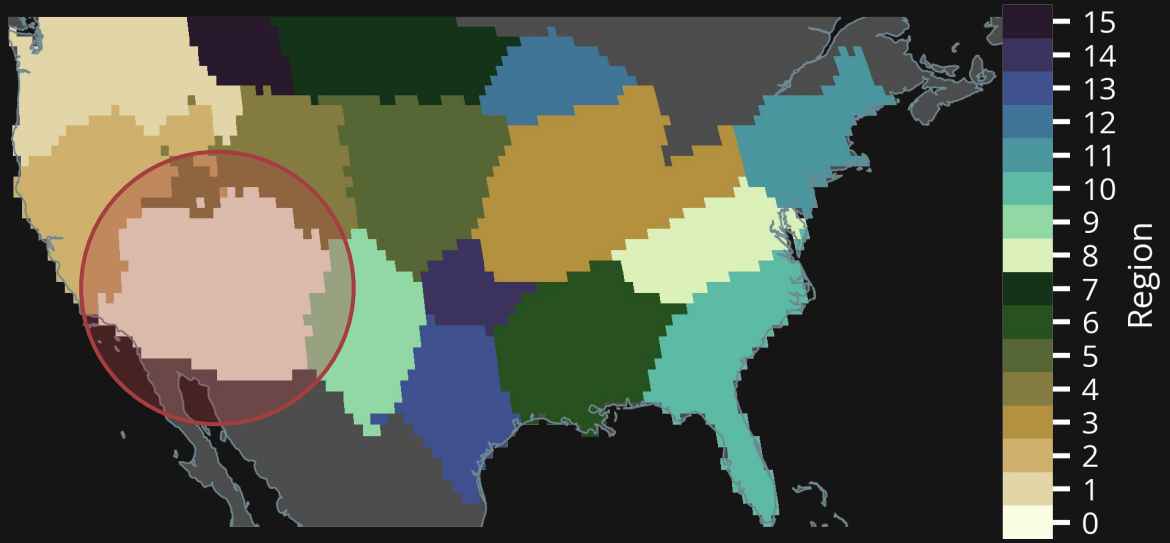
Meehl et al. Initialized Earth System prediction from subseasonal to decadal timescales. Nat Rev Earth Environ 2, 340–357 (2021)

Spatial aggregation of target regions

Consider sub-seasonal variations by looking at biweekly anomalies of precipitation

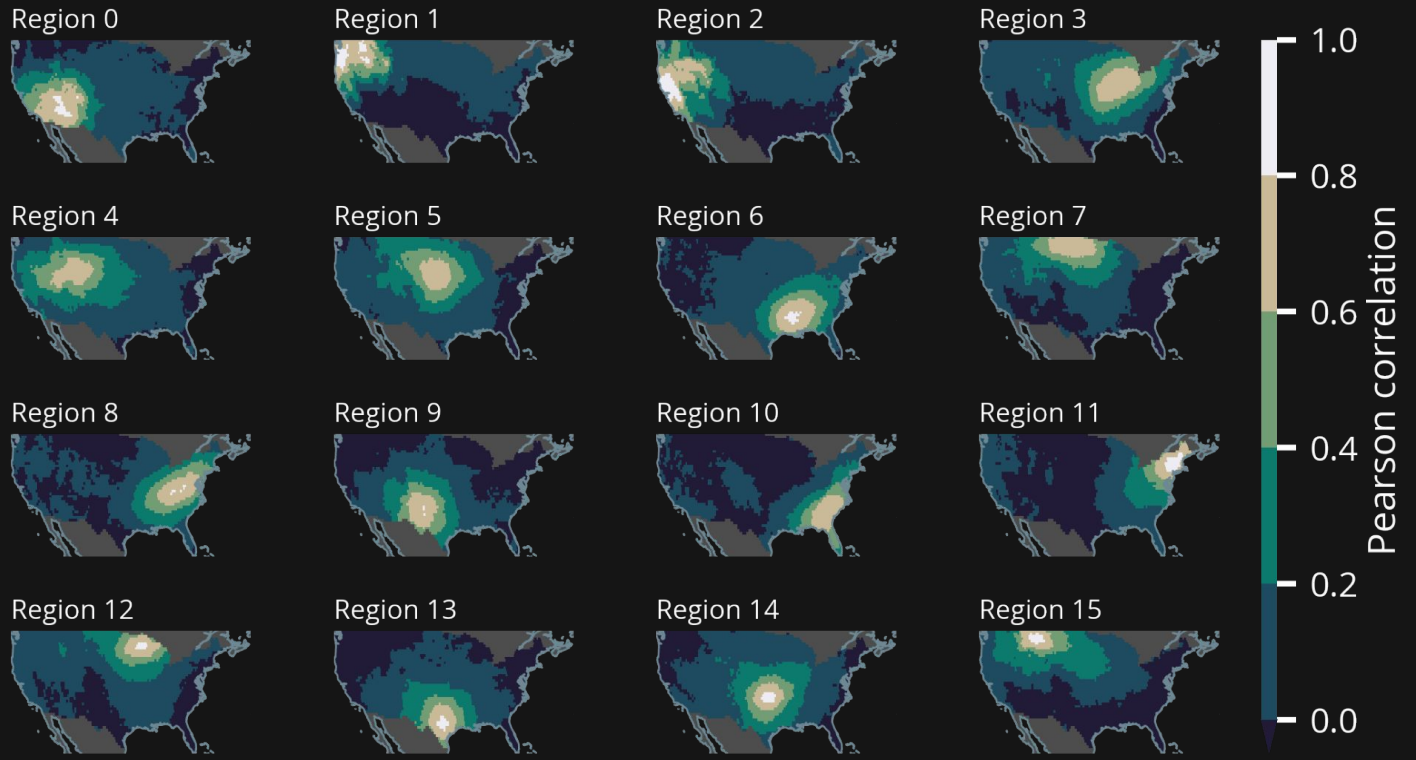
We use Varimax-rotated PCA to find clusters of similar precipitation patterns

Focus on mode 1 (= region 0) which describes most of the sub-seasonal variance



Data: CPC Global Unified Gauge-Based Analysis of Daily Precipitation

Cluster verification



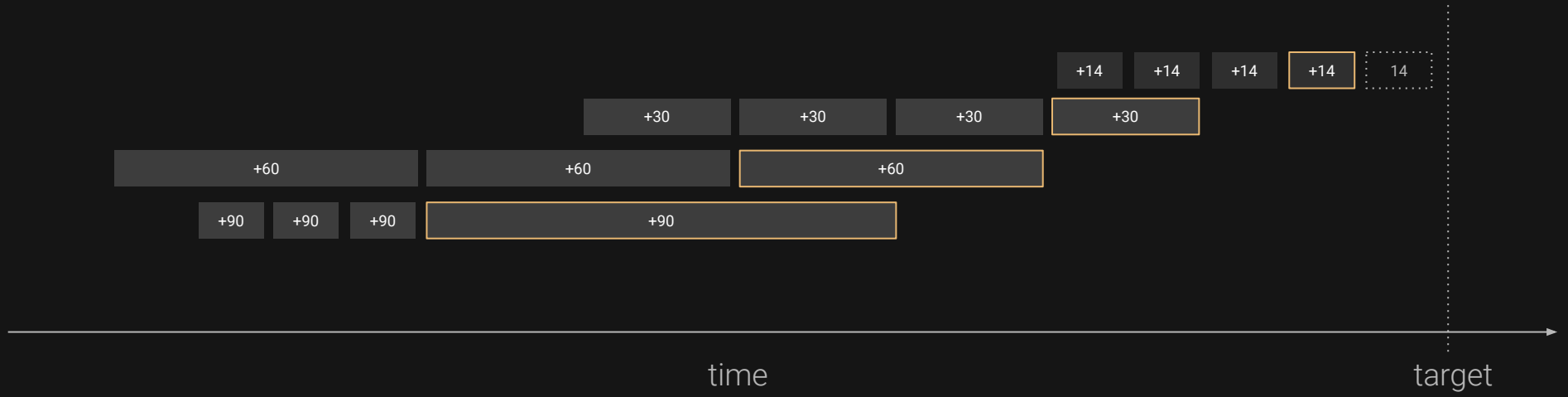
Find potential predictors

Compute correlation maps for:

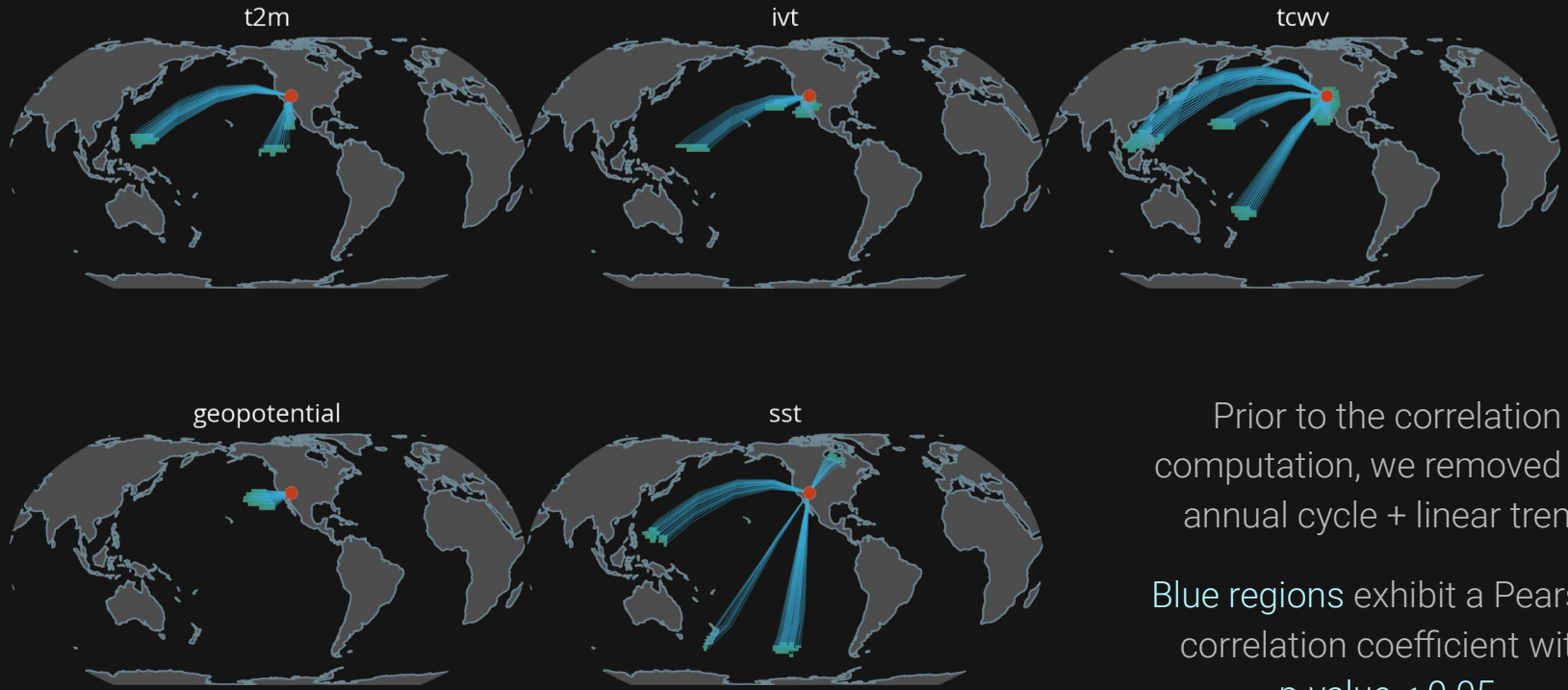
geopotential, ivt, sic, sm, snow, sss, sst, t2m, tcwv

smoothing windows

90 **60** **30** **14**



Potential predictor regions



Prior to the correlation computation, we removed the annual cycle + linear trend

Blue regions exhibit a Pearson correlation coefficient with p-value < 0.05

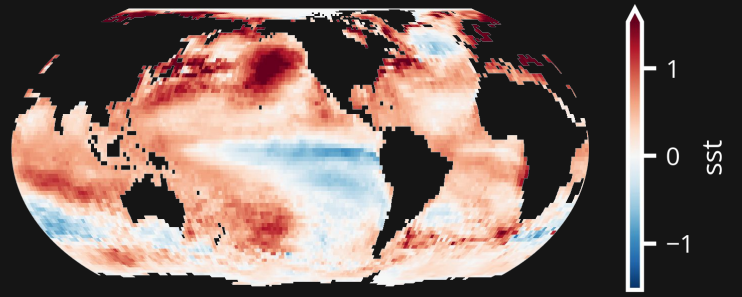
Forecasting using a LSTM-RNN

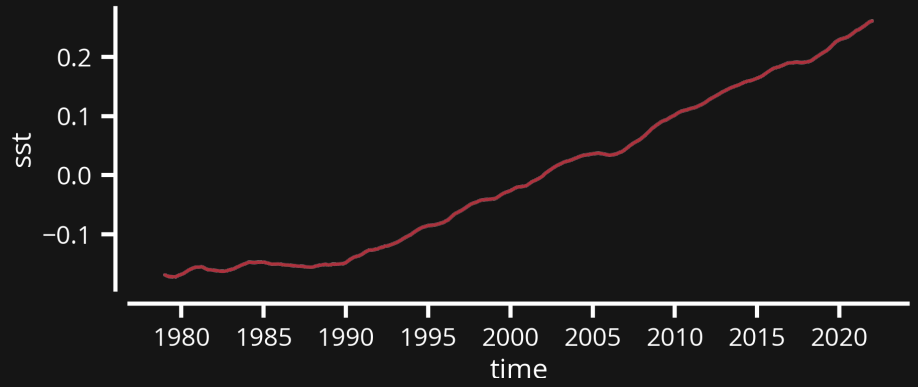
We build a LSTM-RNN to predict probability of a precipitation tercile using the identified predictors.

Additional predictors:

1. ONI, EMI, NAO
2. QBO, MJO amplitude & phase
3. Solar cycle
4. Seasonal cycle
5. Global warming index

Training period: 1979-2017
Validation: leave-one-year-out cross validation
Test period: 2018-2021





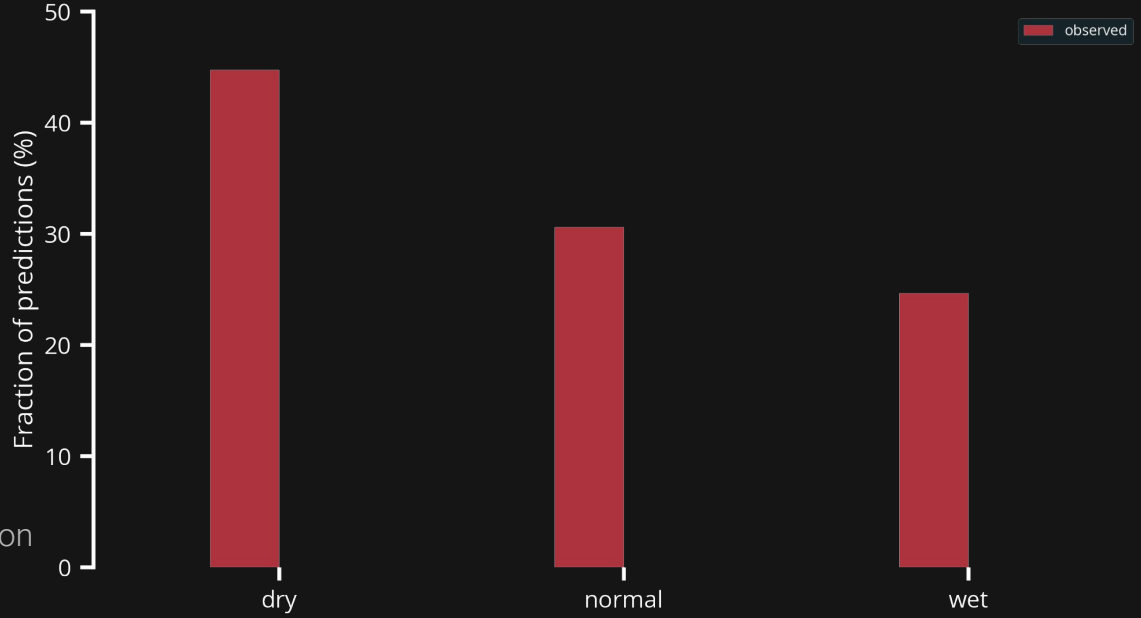
Model evaluation

We build a LSTM-RNN to predict probability of a precipitation tercile using the identified predictors.

Additional predictors:

1. ONI, EMI, NAO
2. QBO, MJO amplitude & phase
3. Solar cycle
4. Seasonal cycle
5. Global warming index

Training period: 1979-2017
Validation: leave-one-year-out cross validation
Test period: 2018-2021



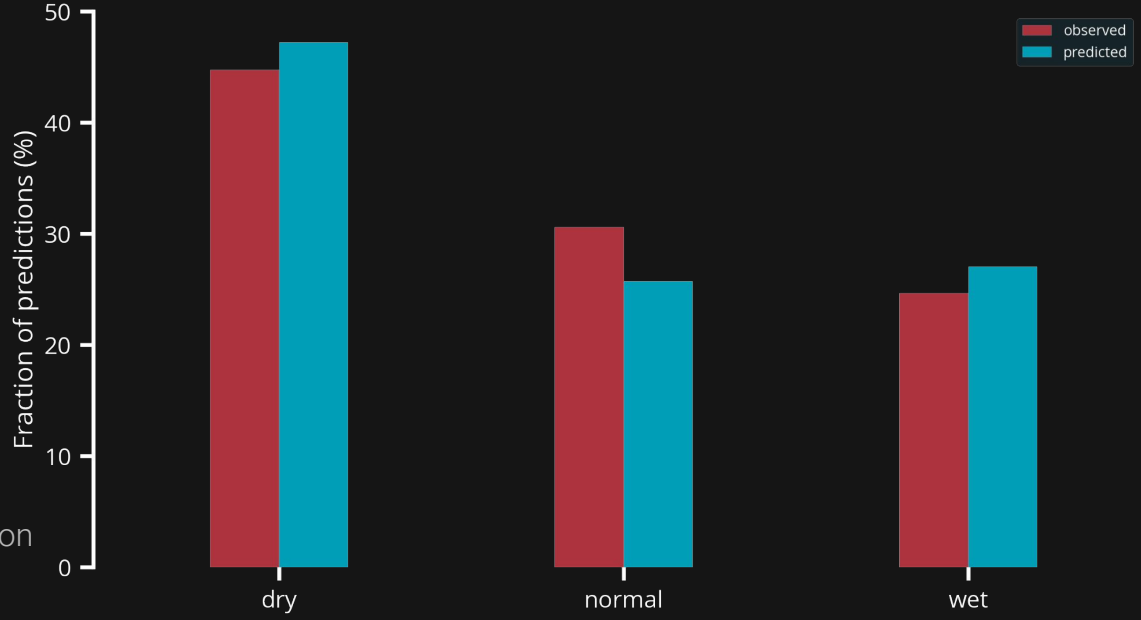
Model evaluation

We build a LSTM-RNN to predict probability of a precipitation tercile using the identified predictors.

Additional predictors:

1. ONI, EMI, NAO
2. QBO, MJO amplitude & phase
3. Solar cycle
4. Seasonal cycle
5. Global warming index

Training period: 1979-2017
Validation: leave-one-year-out cross validation
Test period: 2018-2021



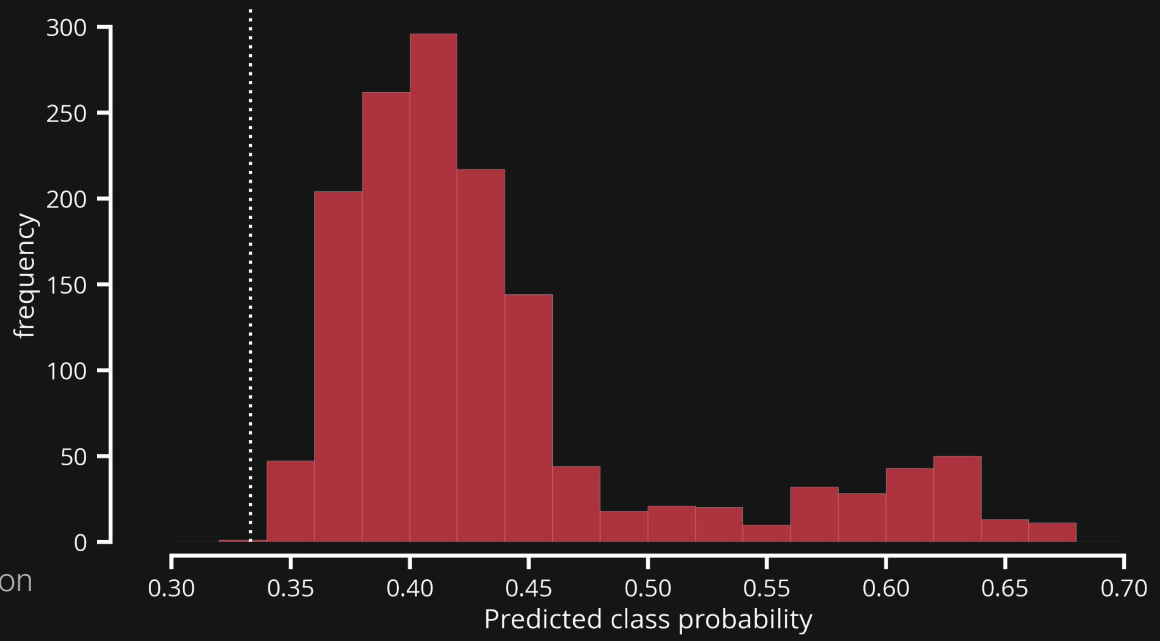
Model evaluation

We build a LSTM-RNN to predict probability of a precipitation tercile using the identified predictors.

Additional predictors:

1. ONI, EMI, NAO
2. QBO, MJO amplitude & phase
3. Solar cycle
4. Seasonal cycle
5. Global warming index

Training period: 1979-2017
Validation: leave-one-year-out cross validation
Test period: 2018-2021



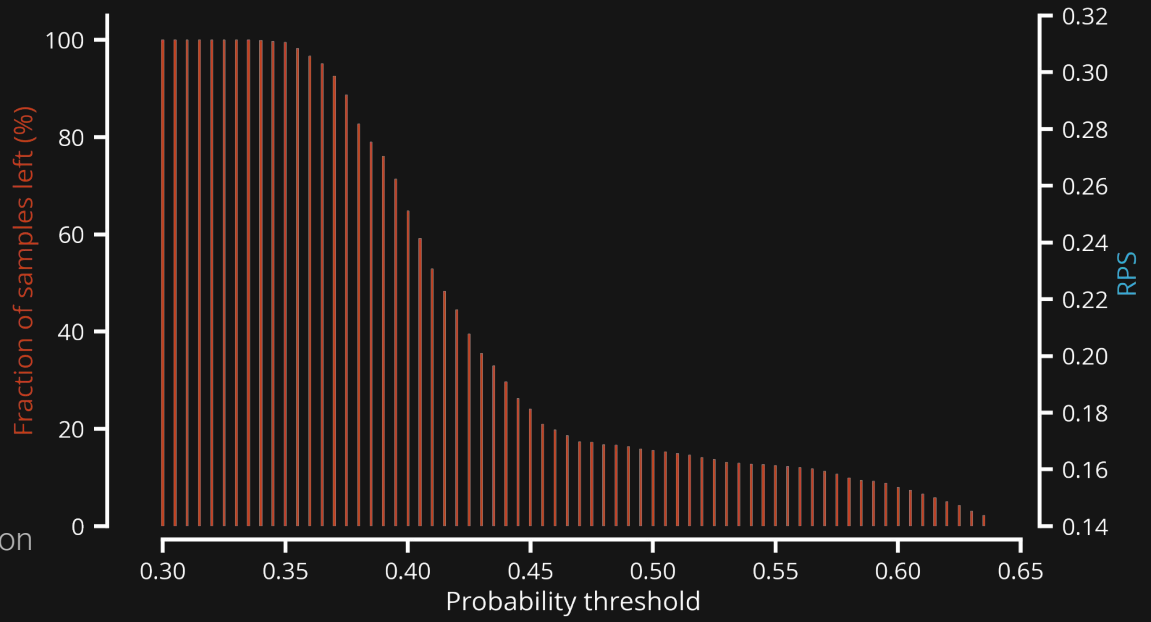
Model evaluation

We build a LSTM-RNN to predict probability of a precipitation tercile using the identified predictors.

Additional predictors:

1. ONI, EMI, NAO
2. QBO, MJO amplitude & phase
3. Solar cycle
4. Seasonal cycle
5. Global warming index

Training period: 1979-2017
Validation: leave-one-year-out cross validation
Test period: 2018-2021



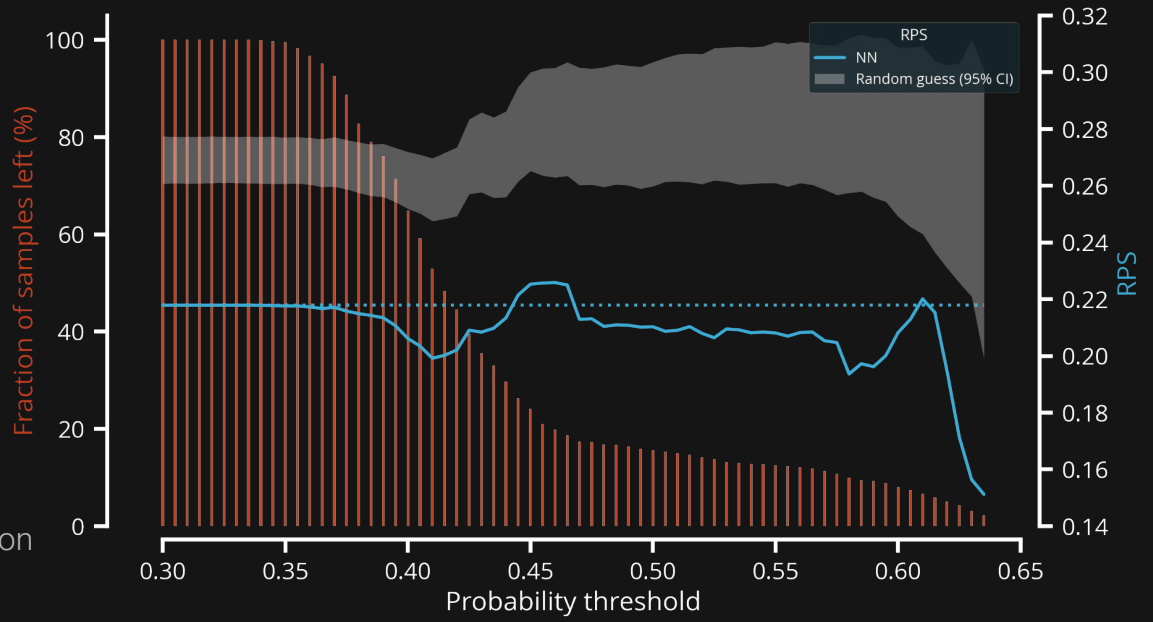
Model evaluation

We build a LSTM-RNN to predict probability of a precipitation tercile using the identified predictors.

Additional predictors:

1. ONI, EMI, NAO
2. QBO, MJO amplitude & phase
3. Solar cycle
4. Seasonal cycle
5. Global warming index

Training period: 1979-2017
Validation: leave-one-year-out cross validation
Test period: 2018-2021



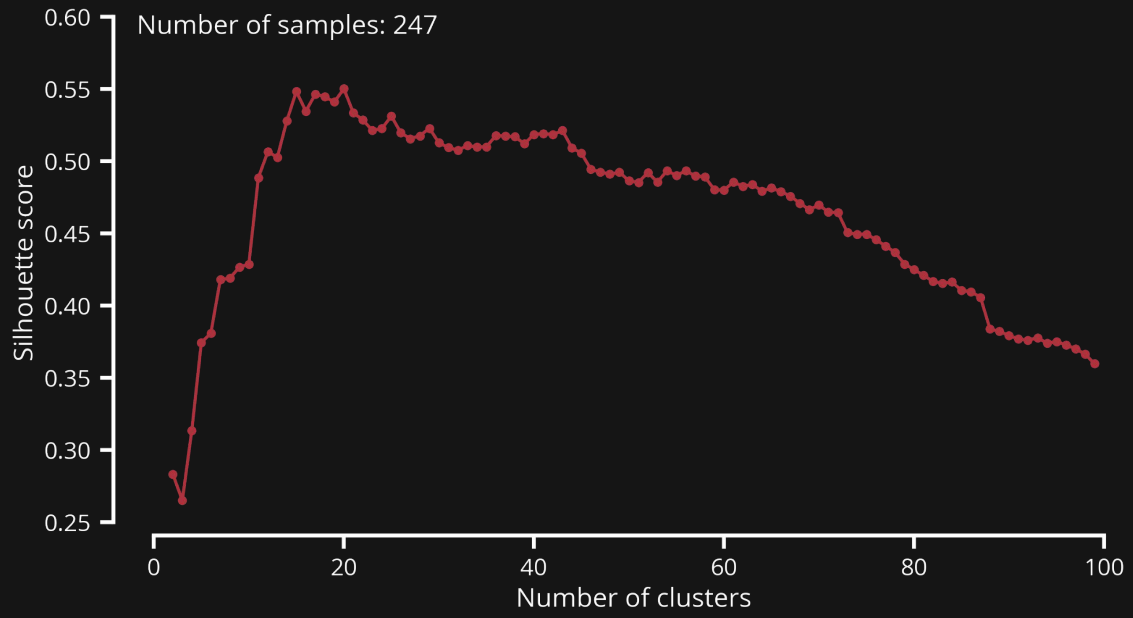
Relevance clustering

To get insight into the models' reasoning we apply layerwise relevance propagation (LRP)

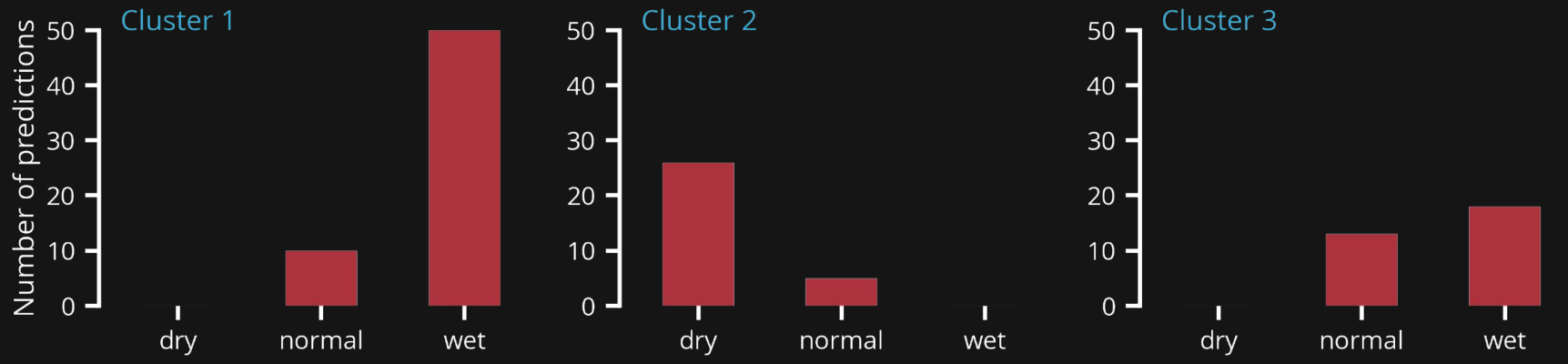
We want to explain predictions which

1. correctly predicted the target
2. are confident ($p > 0.45$)

We use hierarchical clustering to identify 15 groups of LRP scores as indicated by the silhouette score

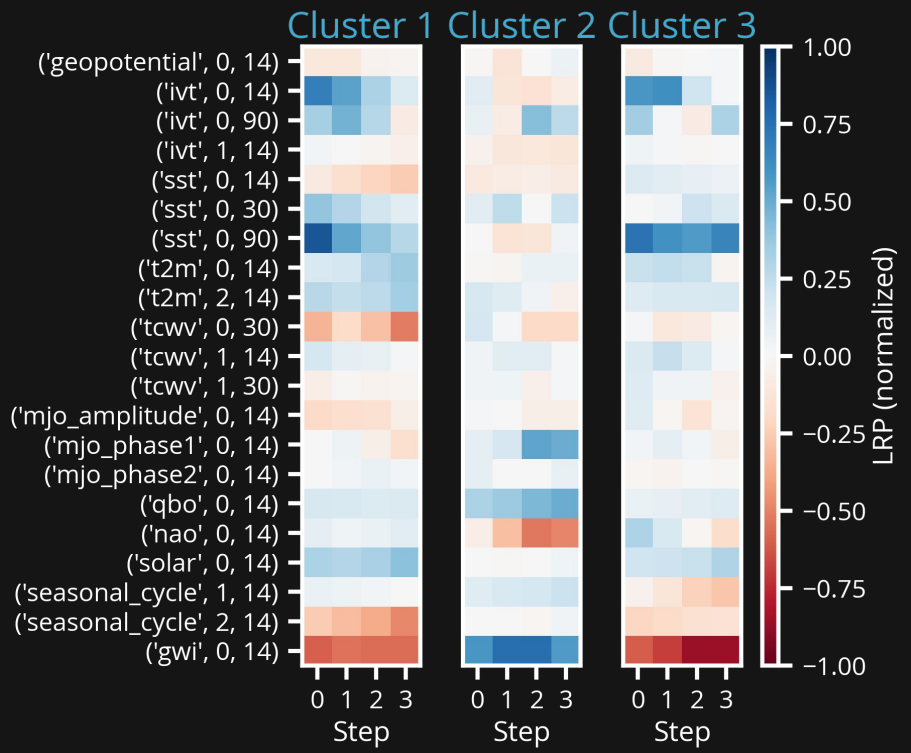


Relevance clustering

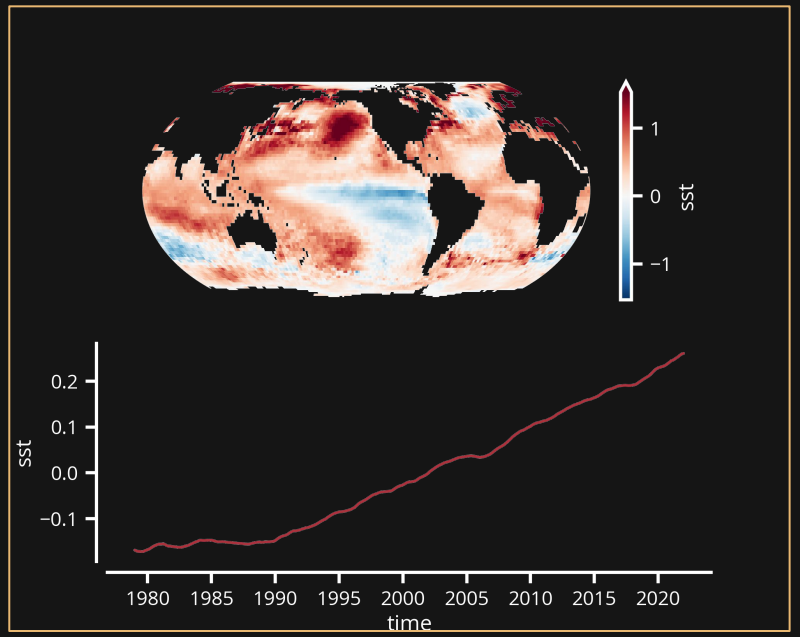


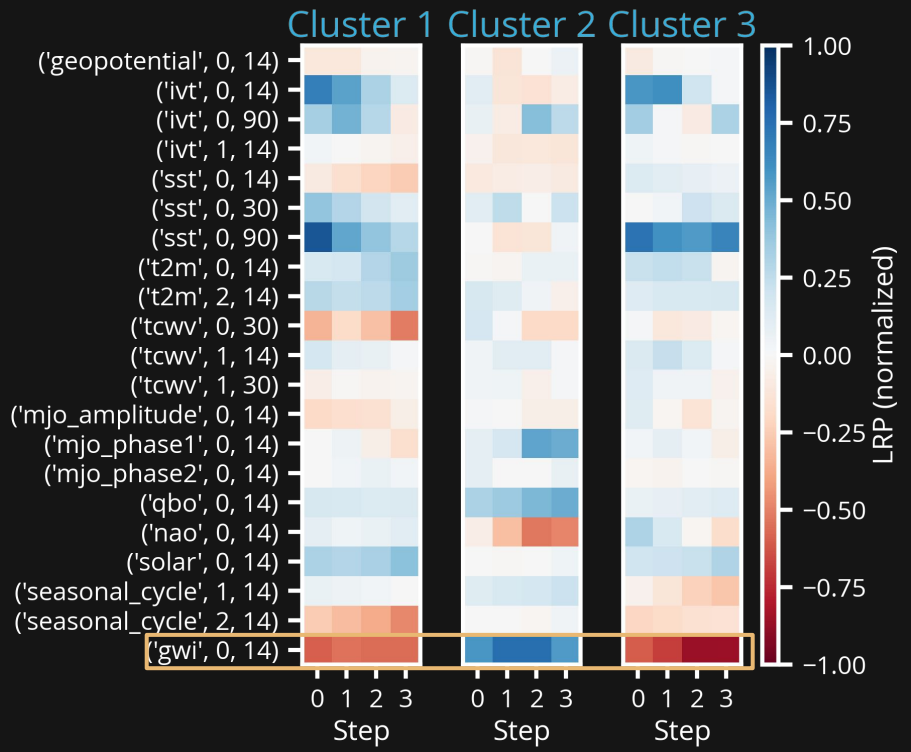


Relevance scores

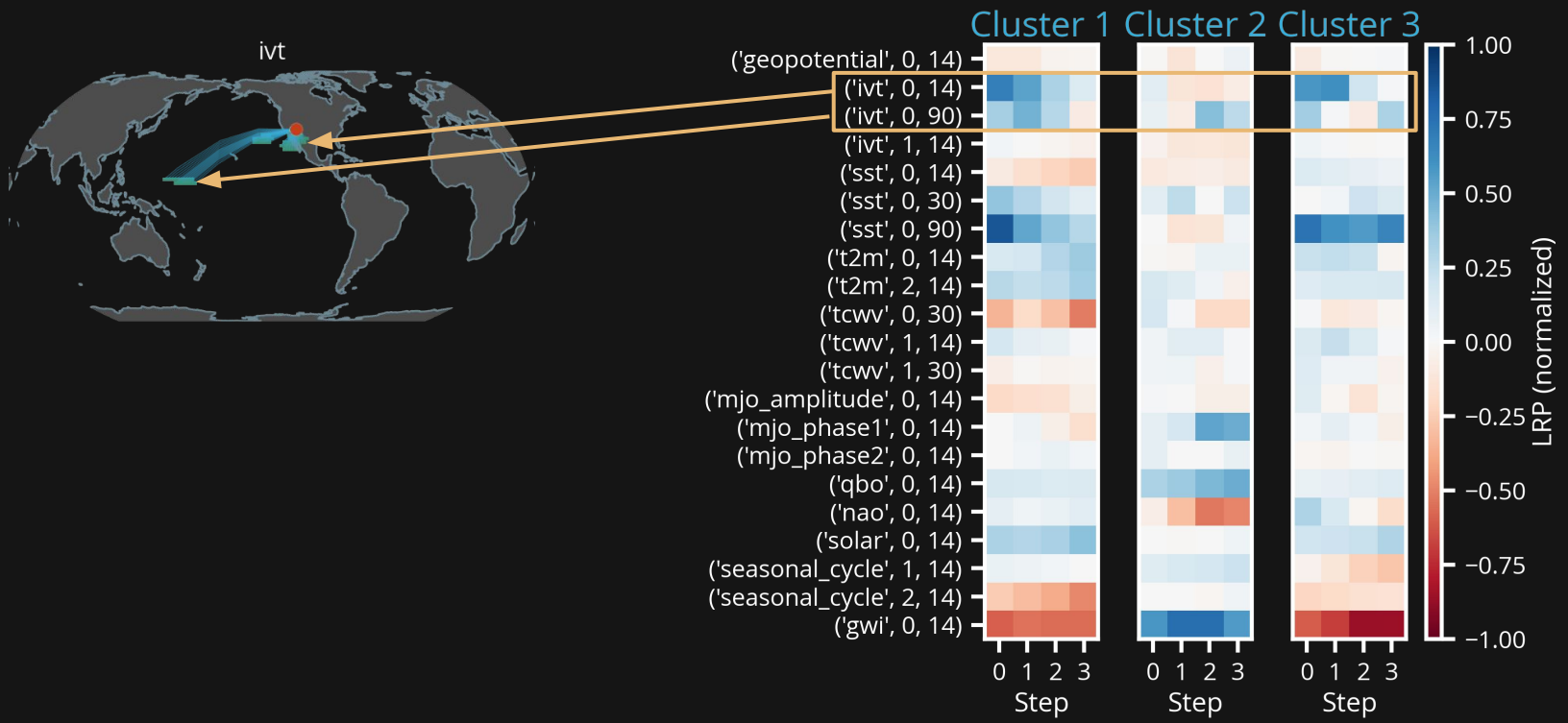


Relevance scores

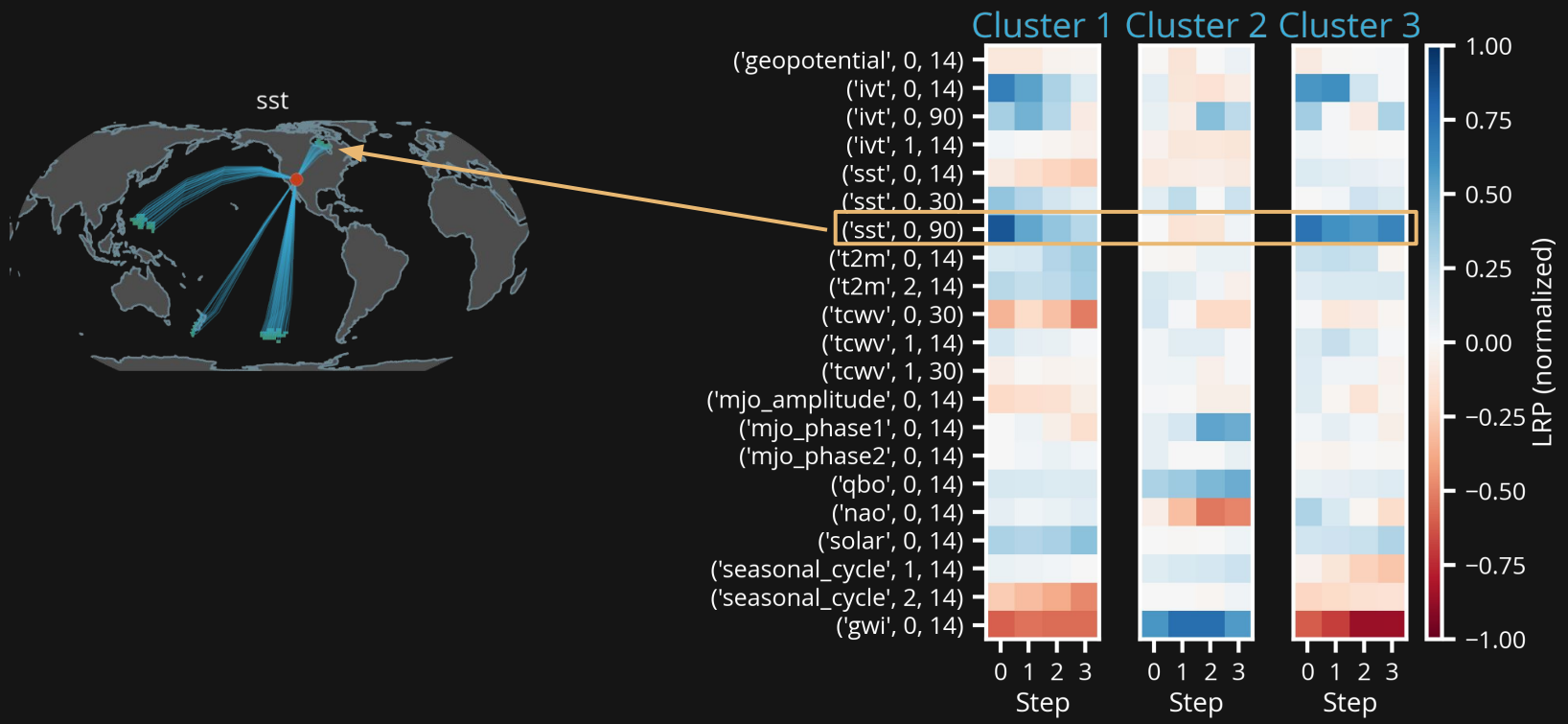




Relevance scores



Relevance scores



Conclusions



Correlation maps at different lead times to identify potential predictors



We build a RNN to predict week 3+4 precipitation in the southwestern US to be in a given tercile of the climatological distribution



Explainable AI suggests water vapour transport (~2 weeks), SST (~ 3 months) and global warming trend as most relevant predictors