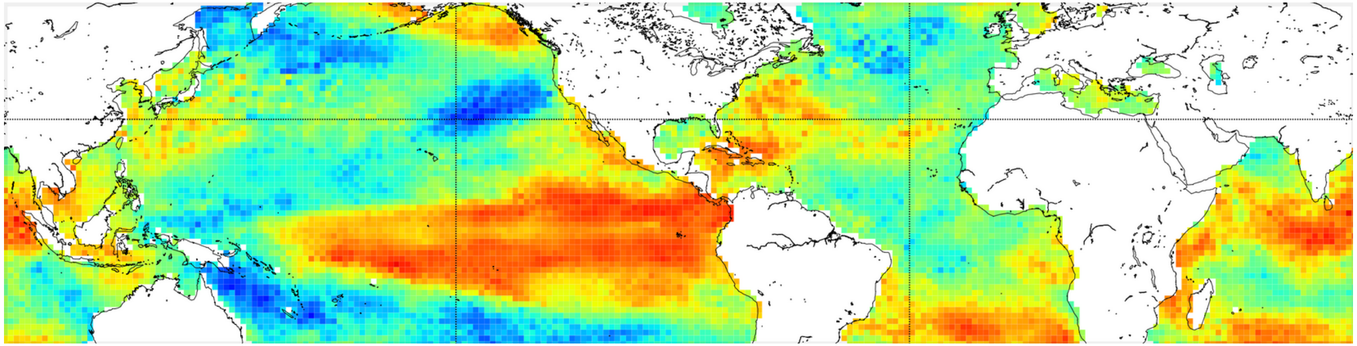
Improving sub-seasonal drought forecasting with machine learning and climate indices

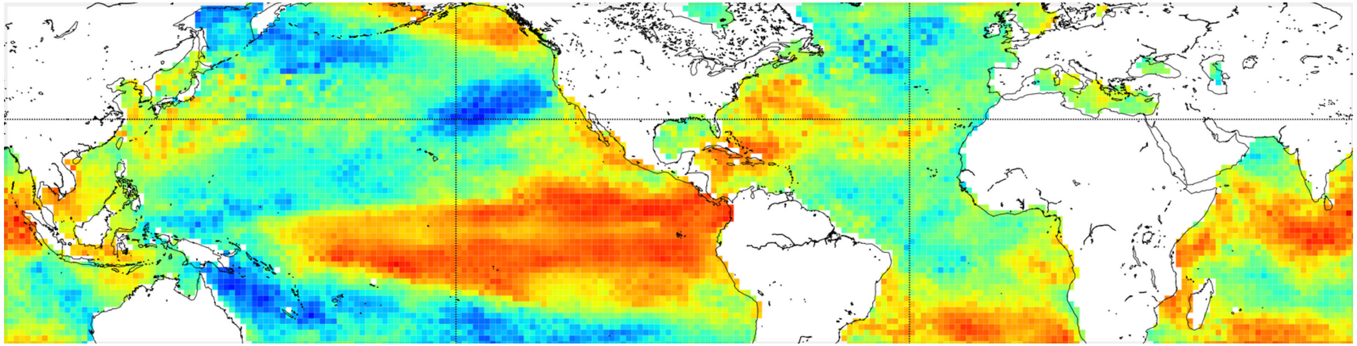
BOSSO FRANCESCO

IHE Delft, 03/11/2022



you can find the slides here!

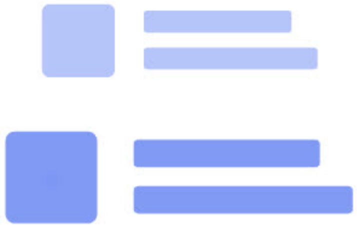
BOSSO FRANCESCO



IHE Delft, 03/11/2022

Today's Agenda

this presentation will go through the following stages:



01

Intro

02

Context

03

Framework

Intro

- 01 What is drought
- 02 ML for Drought
- 03 The gap

Intro

Meteorological Drought

a period of time in which a region experiences below-normal precipitation

---

Reduced soil moinsture,Reduced stream flow,Crop damage

Water shortage

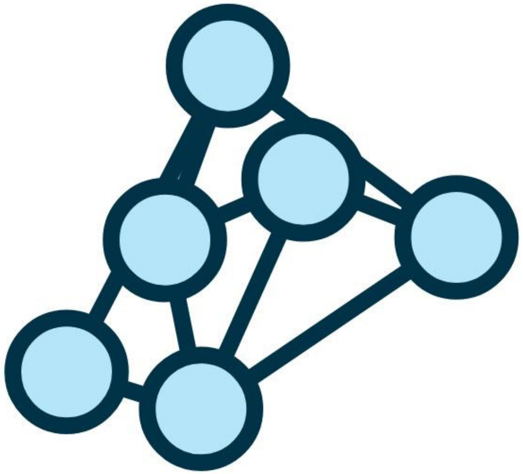
Intro

The onset, extent and duration of drought are difficult to define

different stakeholders have varying degrees of tolerance and resilience to these events (Slette et al., 2019)

Being able to forecast them is crucial

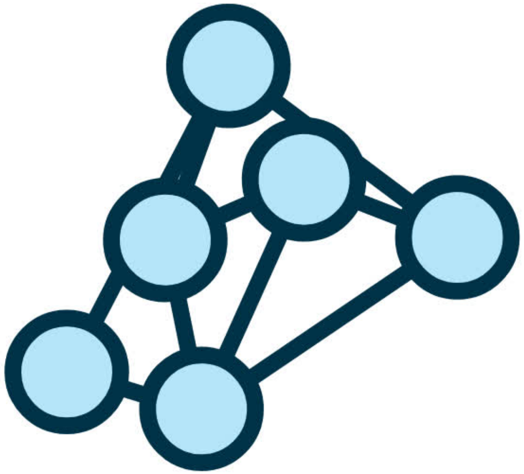
Intro



exploitation of *statistic* and *dynamic techniques* for droughts forecasting has been and is widely studied

sub-seasonal forecasting

Intro



Earth observation data
Artificial Intelligence
Hardware (GPU,TPU)

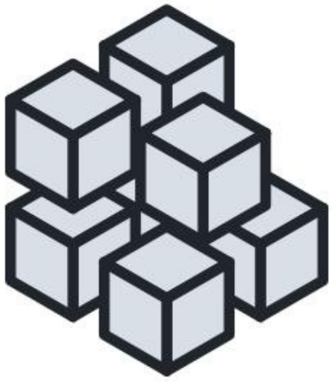
→ AI-based prediction models

McGovern et al. (2017)

Learn from past data
Integrate physical understanding into the models
Discover additional knowledge from the data
Handle large amounts of input variables

Intro

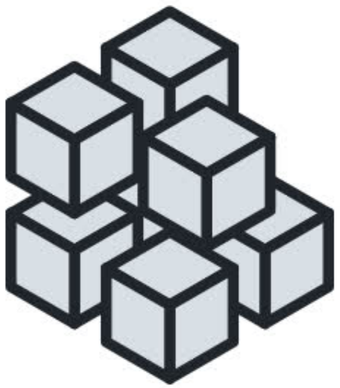
- 01 What is drought
- 02 ML for Drought
- 03 The gap



sub-seasonal drought forecasting ⟷ AI

Why to focus on sub-seasonal lead times?

Intro



Informative predictors

seasonal:

climate indices and large scale teleconnection patterns

short-medium term:

local variable (precipitation, temperature)

sub-seasonal?

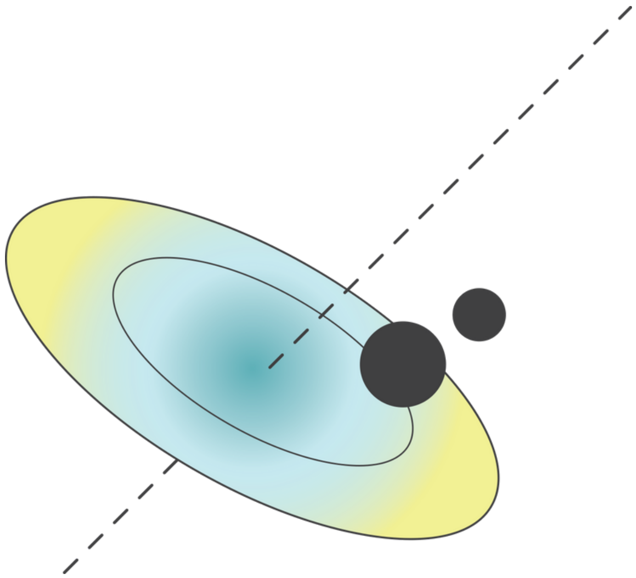
- short enough that the atmosphere still has memory of its initial conditions
- long enough to allow atmospheric circulation to affect the evolution of weather conditions

Context

01 What (our goal)

02 Where (study area)

03 How (the framework)



Context

Machine Learning model for sub-seasonal drought classification

Based on SPI classes

Machine Learning model for sub-seasonal precipitation forecasting

precipitation forecasting

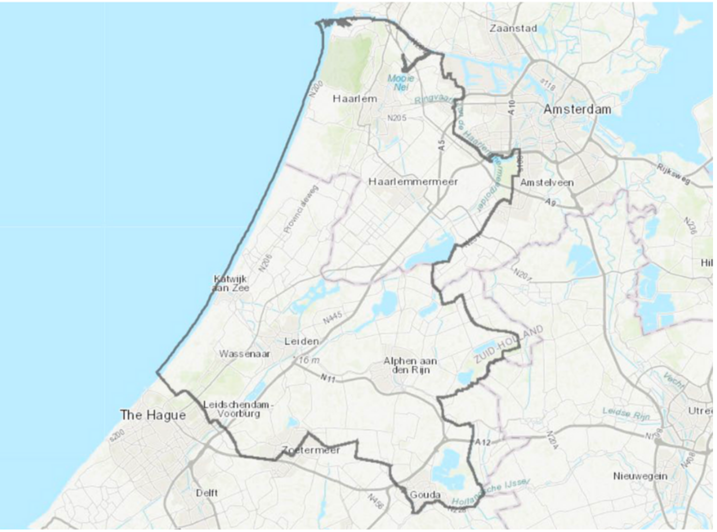
↓

SPI

↓

drought forecating

Context



Rijnland

small sub-catchment of 1000 km2 at the very end of the Rhine delta in the Netherlands

water board of Rijnland is able to forecast drought at bi-weekly lead times. The goal is to extend it to a month

Context

Extreme Event definition → **Input variables selection** → **ML algorithm** → **Performance assessment**

Extreme Event definition

- Meteorological drought
- Occurrence of drought (binary or multi-class or precipitation prediction)
- Decision trigger: SPI
- Lead time: 30 days

Input variables selection

- Climate State Intelligence (Giuliani et al., 2019) extended for sub-seasonal forecasts to identify teleconnections (NIPA)

ML algorithm

- NN for drought prediction on a single cell (precipitation forecasting or SPI classification) – input form NIPA and target in a single cell

Performance assessment

- Depending on the kind of model: MSE, hitrate, …

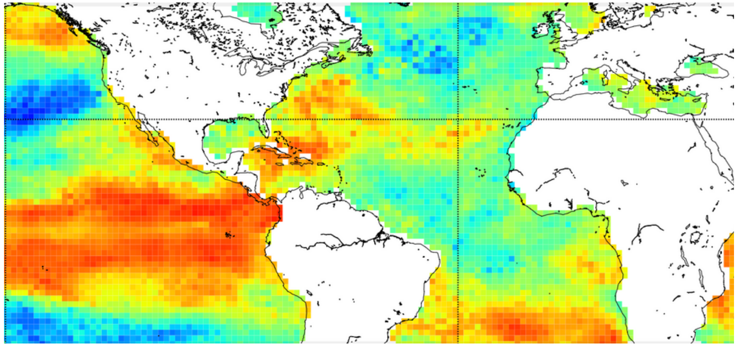
Framework

- 01 NIPA
- 02 Neural Network

Framework

- 01 NIPA
- 02 Neural Network

Nino Index Phase Analysis



Zimmerman et al. (2016)

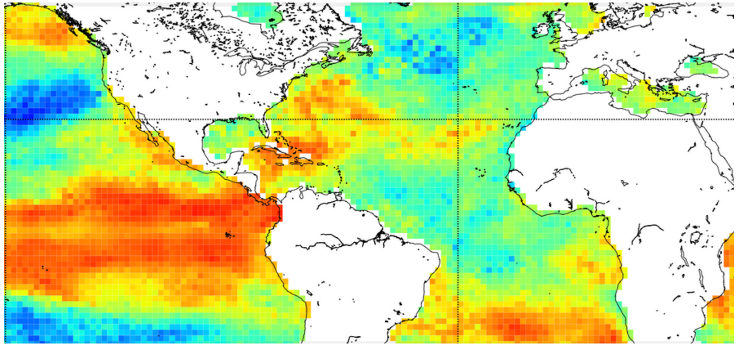
↓

Giuliani et al. (2019)

↓

Our readaptation

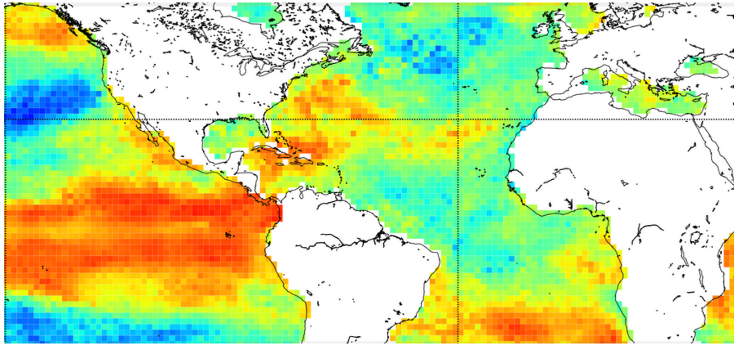
Framework



NIPA is a framework that searches for links between Global and Local variables exploiting the phases of teleconnection patterns materialized by climate indices

Framework

- 01 NIPA
- 02 Neural Network



climate indices

El Niño Southern Oscillation (ENSO)

North Atlantic Oscillation (NAO)

SCAndinavian oscillation (SCA)

East Atlantic oscillation (EA)

Framework

- above/below-normal temperatures in eastern United States and northern Europe
- above/below-normal temperatures in Greenland and southern Europe
- above/below-normal precipitation over northern Europe and Scandinavia
- above/below-normal precipitation over southern and central Europe

climate indices

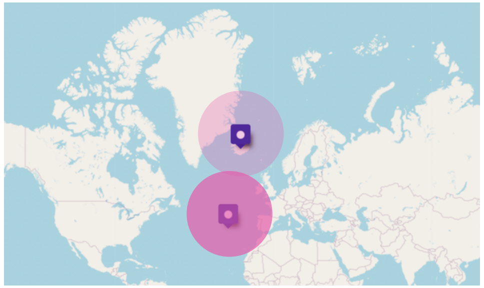
North Atlantic Oscillation (NAO)

Framework

- above/below-normal temperatures in eastern United States and northern Europe
- above/below-normal temperatures in Greenland and southern Europe
- above/below-normal precipitation over northern Europe and Scandinavia
- above/below-normal precipitation over southern and central Europe

climate indices

North Atlantic Oscillation (NAO)



Phase Neg

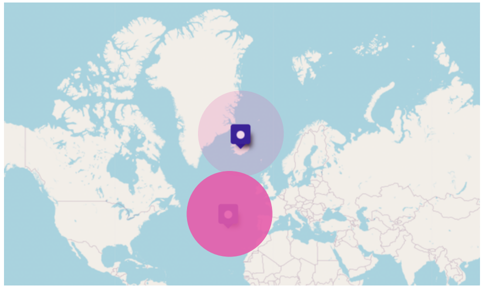
Framework

- 01 NIPA
- 02 Neural Network

- above/below-normal temperatures in eastern United States and northern Europe
- above/below-normal temperatures in Greenland and southern Europe
- above/below-normal precipitation over northern Europe and Scandinavia
- above/below-normal precipitation over southern and central Europe

climate indices

North Atlantic Oscillation (NAO)



Phases Pos

Framework

DATA

- Local precipitation (monthly timeseries) - cumulative
- Global variable (monthly timeseries) - SLP,SST,Z500 - mean
- Climate Index (monthly timeseries) - ENSO, NAO,SCA,EA

SETTING PARAMETERS

- Month (of local precipitation)
- Aggregation level (of pre-month global data)

Framework

SETTING PARAMETERS

- Month (of local precipitation)
- Aggregation level (of pre month global data)

Example:

- Month 1
- Aggregation level 1

local precipitation of January and the global variable of December

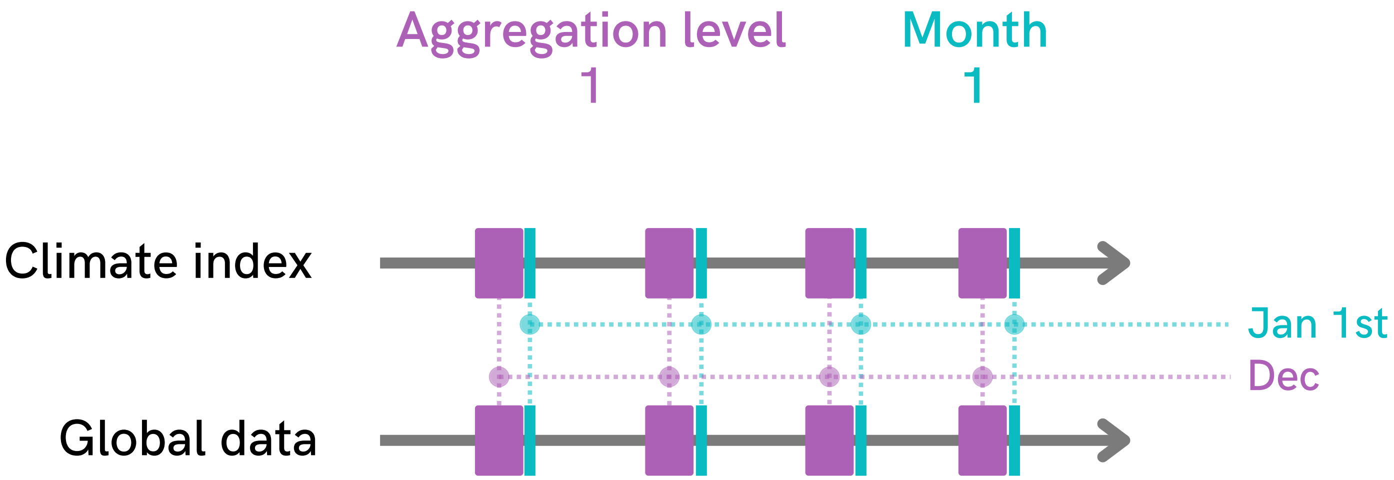
- Month 1
- Aggregation level 2

local precipitation of January and the global variable of November + December

Framework

- 01 NIPA
- 02 Neural Network

Input
Data extraction
Phase segmentation
Correlation
PCA
output

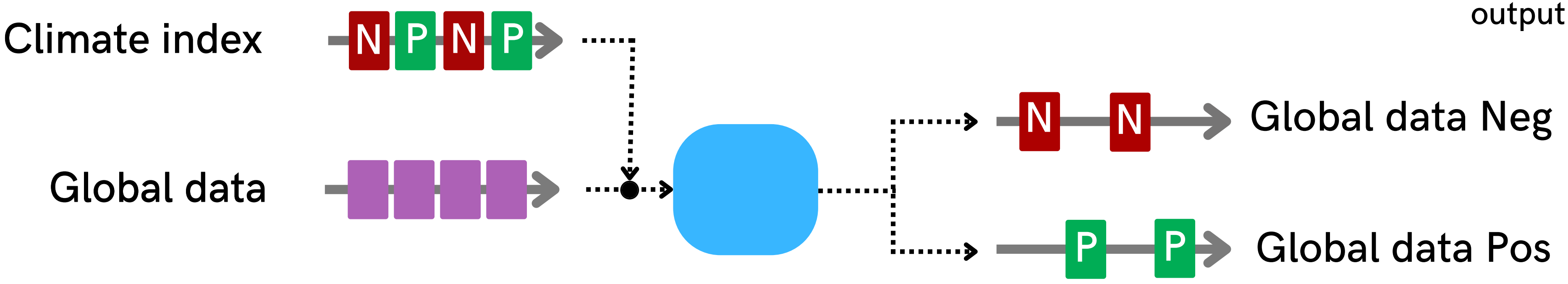


NOTE: this is an year-based operation. NIPA will extract the data for the December of each year

Framework

- 01 NIPA
- 02 Neural Network

Input
Data extraction
Phase segmentation
Correlation
PCA
output



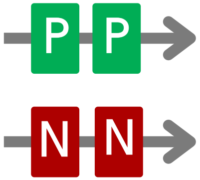
Framework

01 NIPA

02 Neural Network

Input
Data extraction
Phase segmentation
Correlation
PCA
output

Global data Pos
Global data Neg



Jan

Local data

Jan 1st

01 NIPA

02 Neural Network

Input

Data extraction

Phase segmentation

Correlation

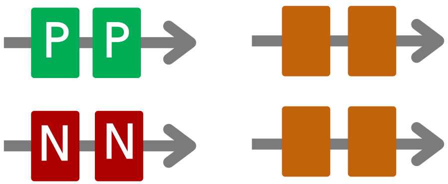
PCA

output

Framework

Global data Pos

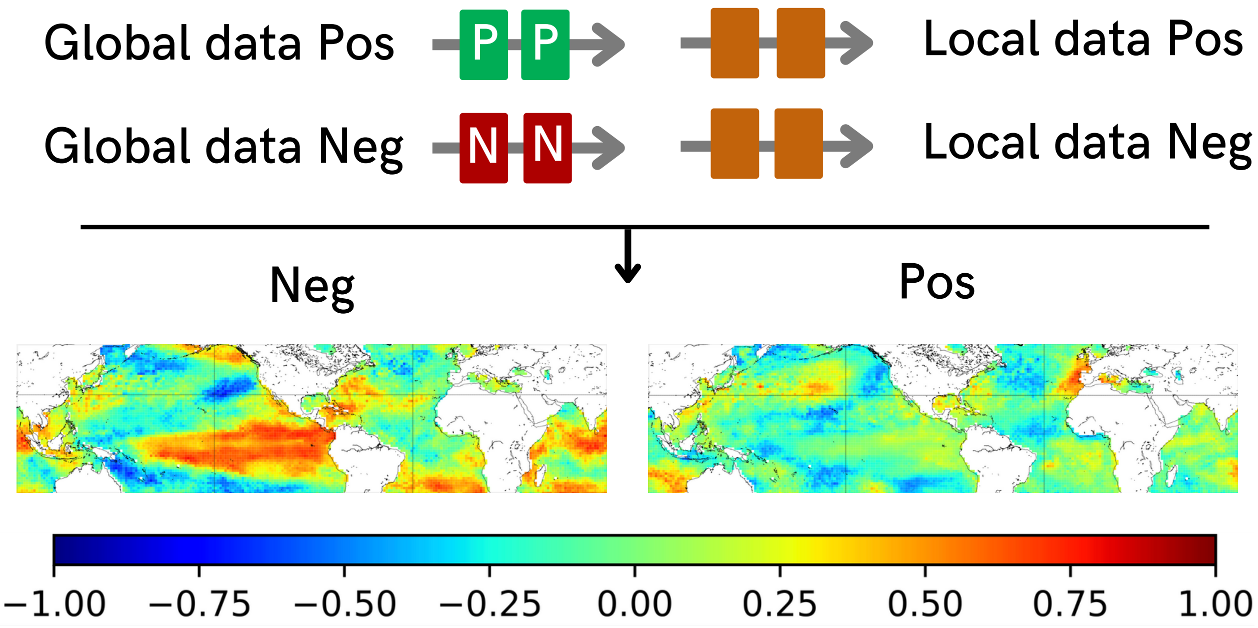
Global data Neg



Local data Pos

Local data Neg

Framework

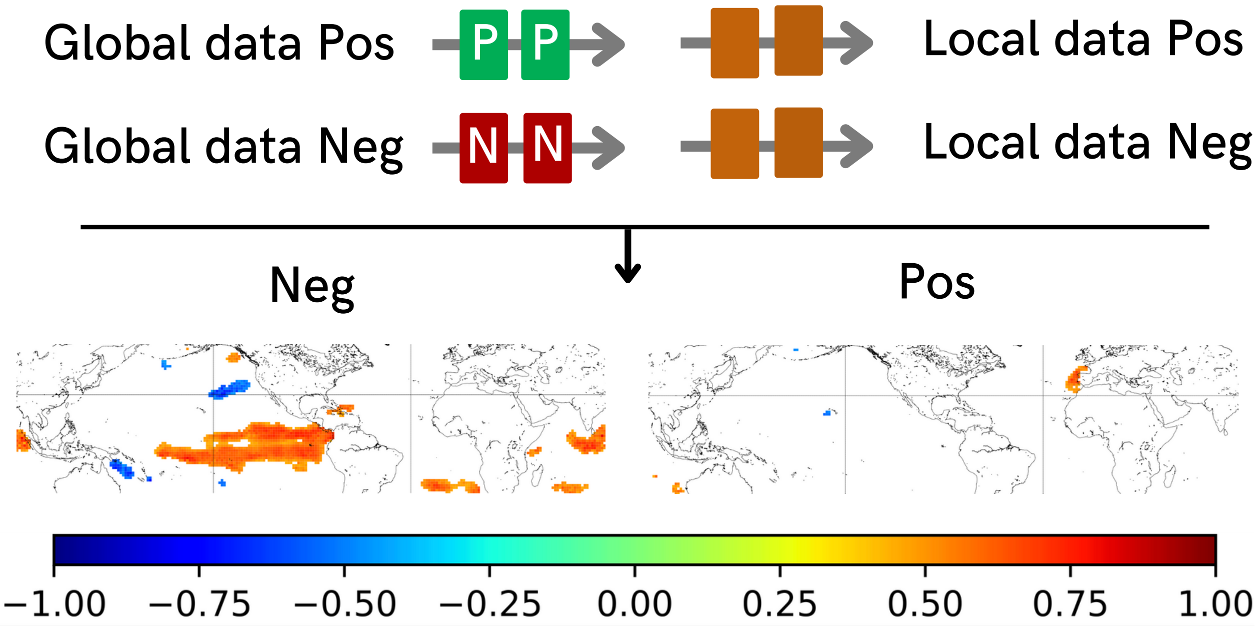


Correlation maps

Framework

- 01 NIPA
- 02 Neural Network

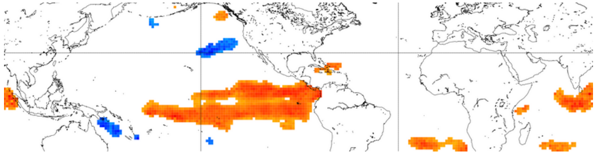
Input
Data extraction
Phase segmentation
Correlation
PCA
output



95% of significance

Framework

Neg

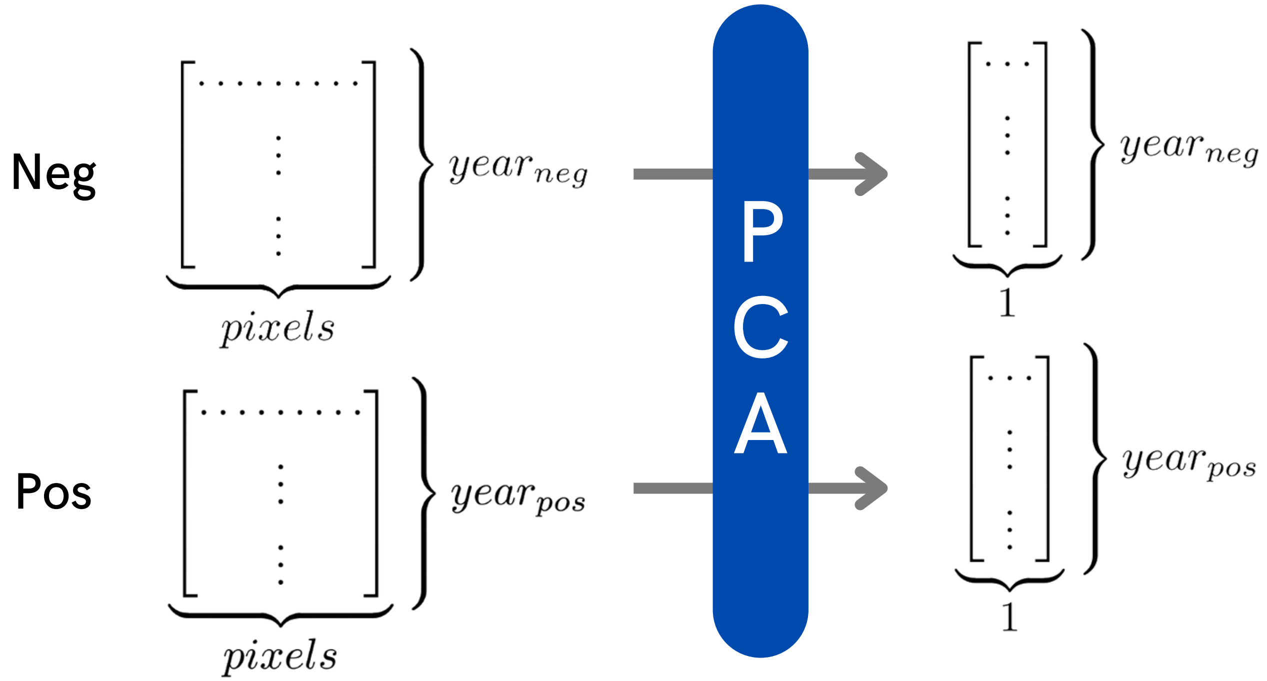


Pos

$$\underbrace{\begin{bmatrix} \cdots\cdots\cdots \\ \vdots \\ \vdots \end{bmatrix}}_{pixels} \left.\vphantom{\begin{bmatrix} \cdots \\ \vdots \\ \vdots \end{bmatrix}}\right\} year_{neg}$$

$$\underbrace{\begin{bmatrix} \cdots\cdots\cdots \\ \vdots \\ \vdots \end{bmatrix}}_{pixels} \left.\vphantom{\begin{bmatrix} \cdots \\ \vdots \\ \vdots \end{bmatrix}}\right\} year_{pos}$$

Framework


Neg
$year_{neg}$
$pixels$
Pos
$year_{pos}$
$pixels$
PCA
$year_{neg}$
1
$year_{pos}$
1

Framework

PC1	phase_label
PC1 1979	1
PC1 1980	2
...	...
...	...
PC1 2021	2

Dataset for 1 month

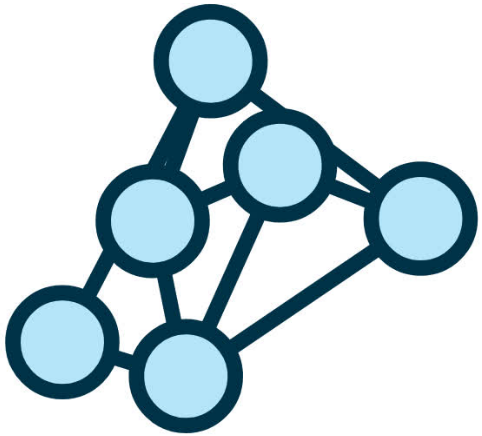
Input
Data extraction
Phase segmentation
Correlation
PCA
output

Framework

This procedure can be applied

- for each Month
- for each combination of:
  - Local Precipitation
  - Global Variable (SST/SLP/Z500)
- for each aggregation level of SST/SLP/Z500 (1/2/3 month)

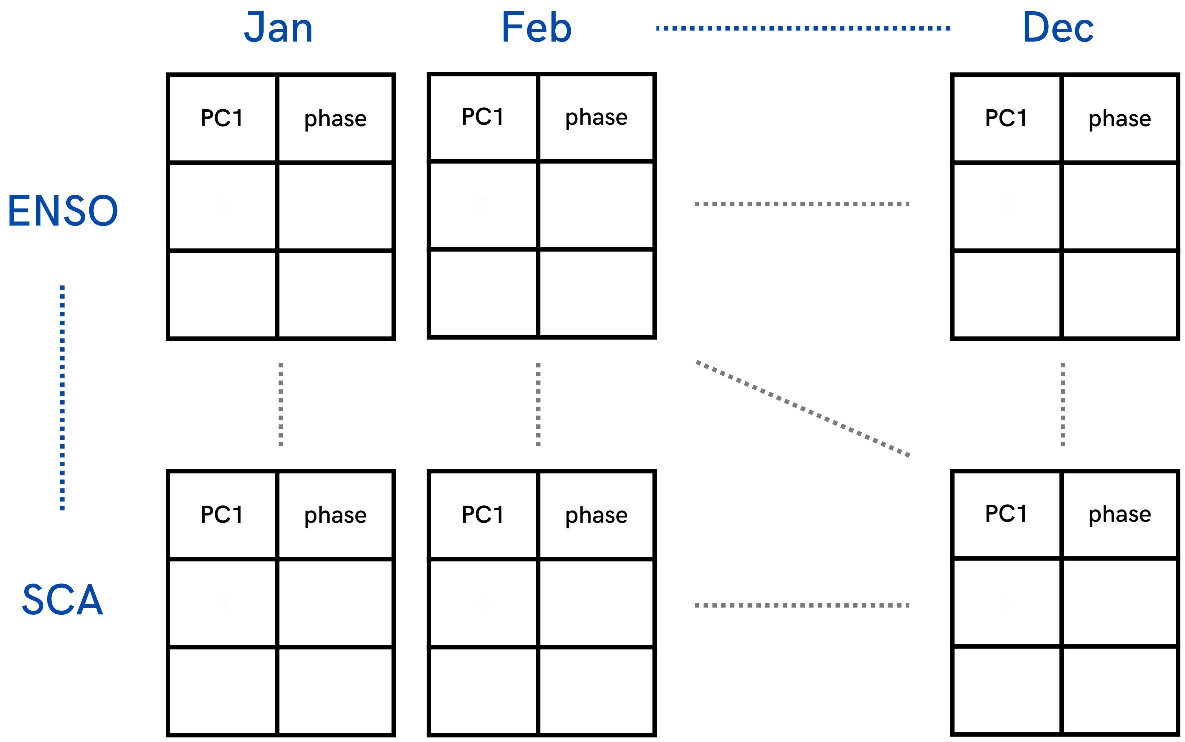
Framework



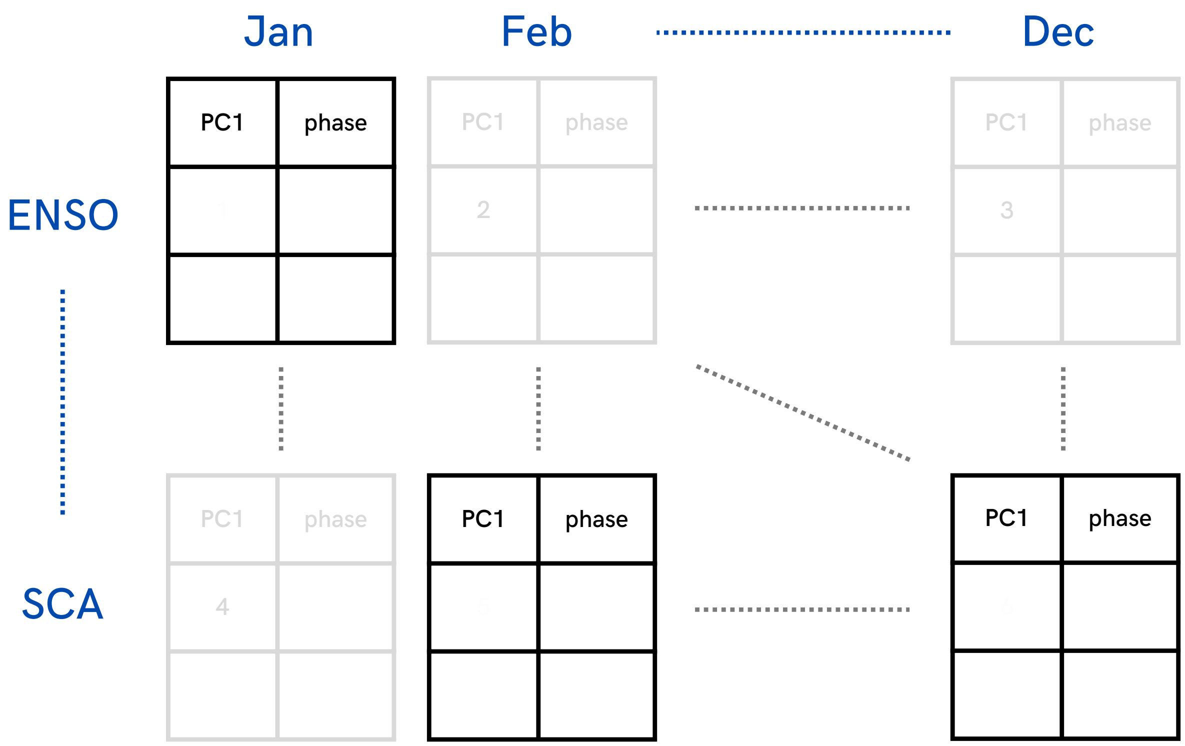
Just entered in this step

- which are our thoughts on how to proceed
- what has emerged from the test

Framework


Jan
Feb
Dec
ENSO
SCA
PC1
phase
PC1
phase
PC1
phase
PC1
phase
PC1
phase
PC1
phase

Framework


Jan
Feb
Dec
ENSO
SCA
PC1
phase
PC1
phase
2
PC1
phase
3
PC1
phase
4
PC1
phase
PC1
phase

Framework

- Skim some of the features based on the pearson coefficients of a linear regression between PC1 and Local Precipitation

- Skim some of the features by imposing a minimum correlation threshold

- Consider the skimmed set of features and build N different models for each month and compare the N different Leave One Out validation errors to choose the best one

Framework

Inputs: (PC1, phase label) **Target**: (Local Precipitation)

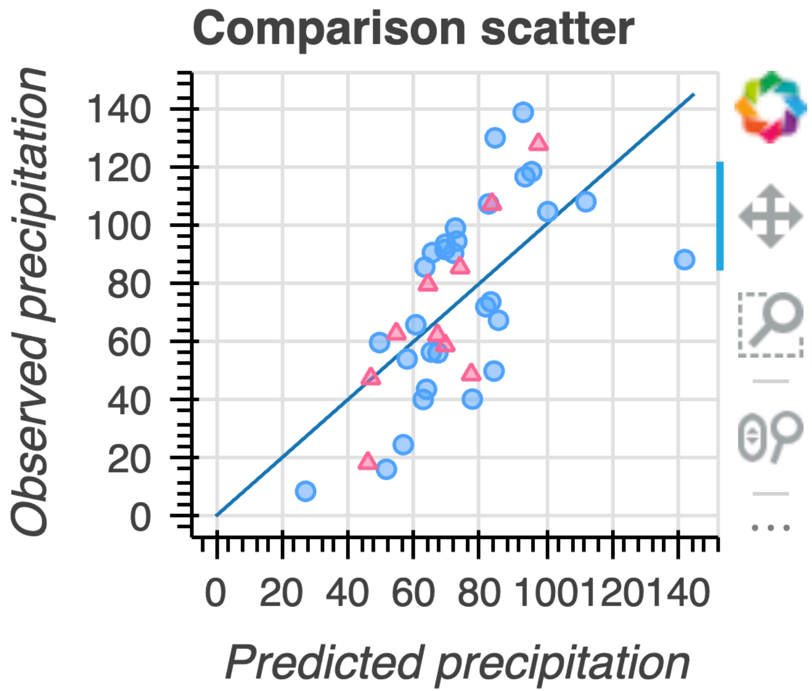
Inputs: (PC1_1, PC1_2, climate state);
Target: (Local Precipitation)

Climate index 1	Climate index 2	Climate state
1	1	1
1	2	2
2	1	3
2	2	4

Framework

- Input features:
  - SCA-SLP-1-1,
  - EA-SST-1-1,
  - climate state
- Target: Cumulative precipitation
- Hidden layers: 2
- Neurons: (3, 2)
- Activation function: ReLU
- Loss function: MSE

Framework



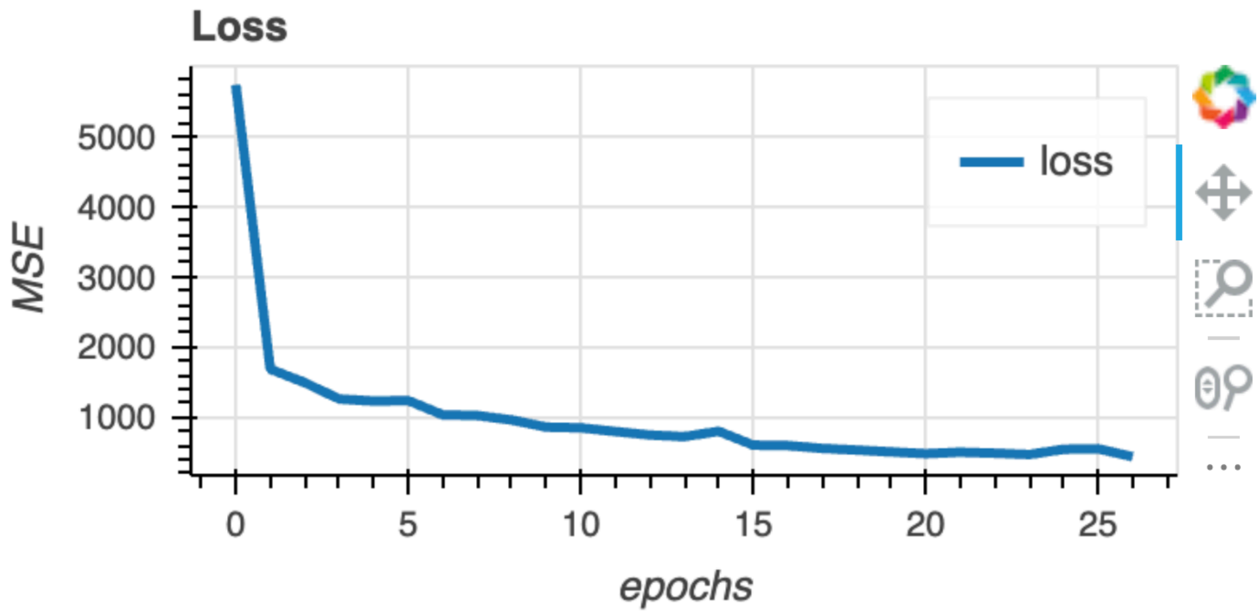
Training set

Validation set

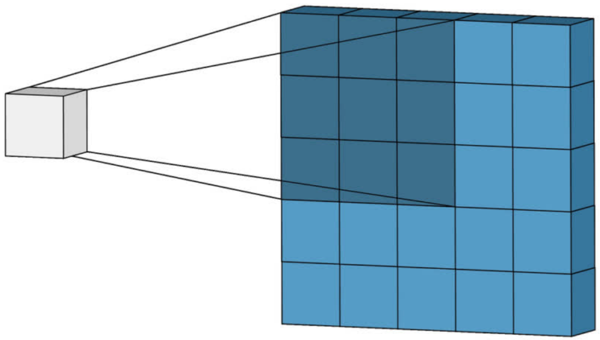
MSE of 319.0494 on the validation set

comparable with giuliani et al with ELM (374.905)

Framework



Future ideas



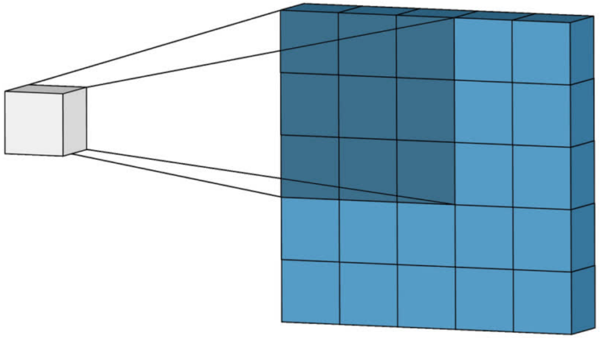
NN

CNN

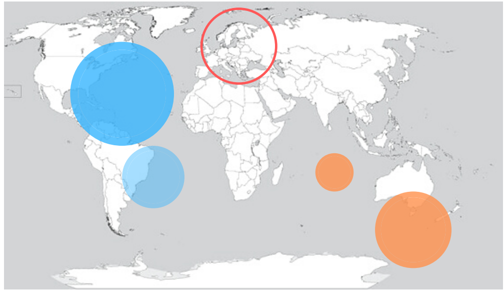
PC1

Correlation Map

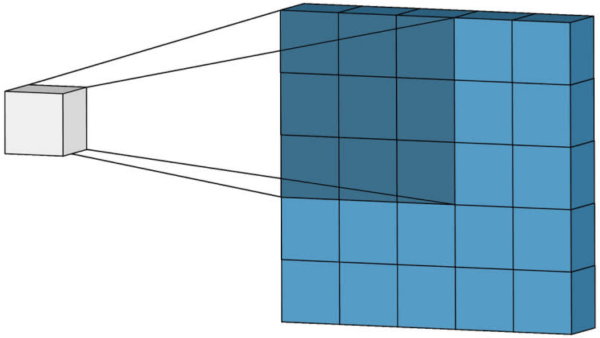
Future ideas

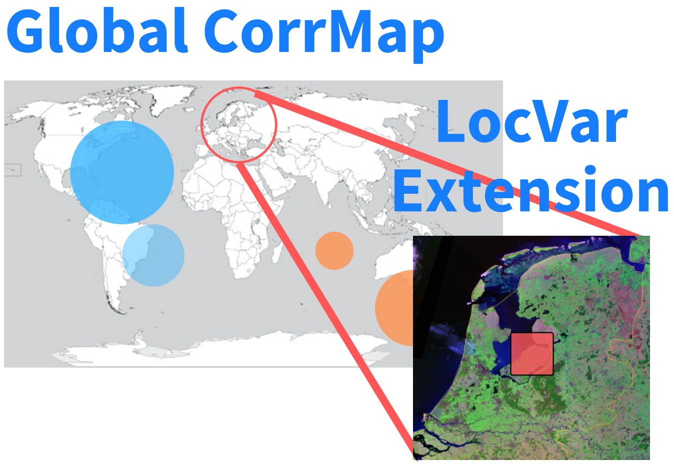


Global CorrMap

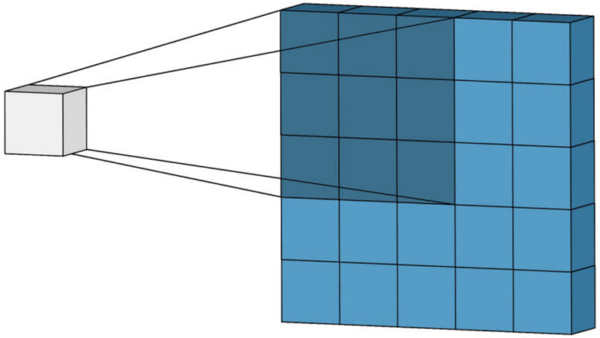


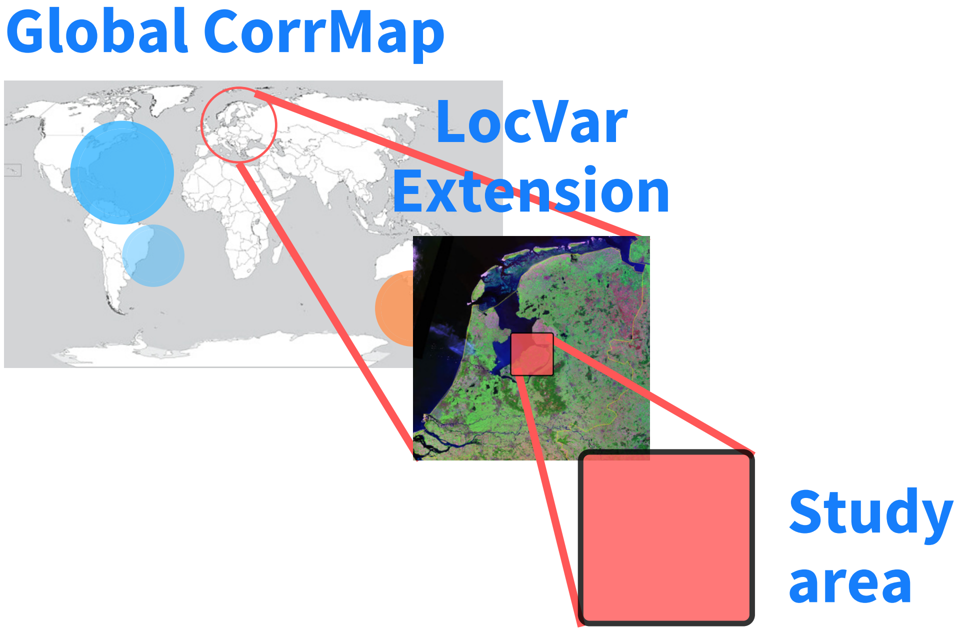
Future ideas



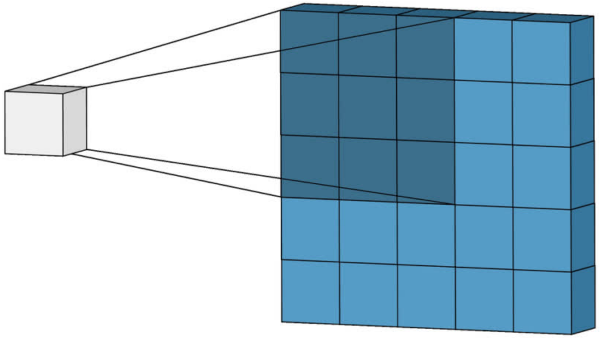


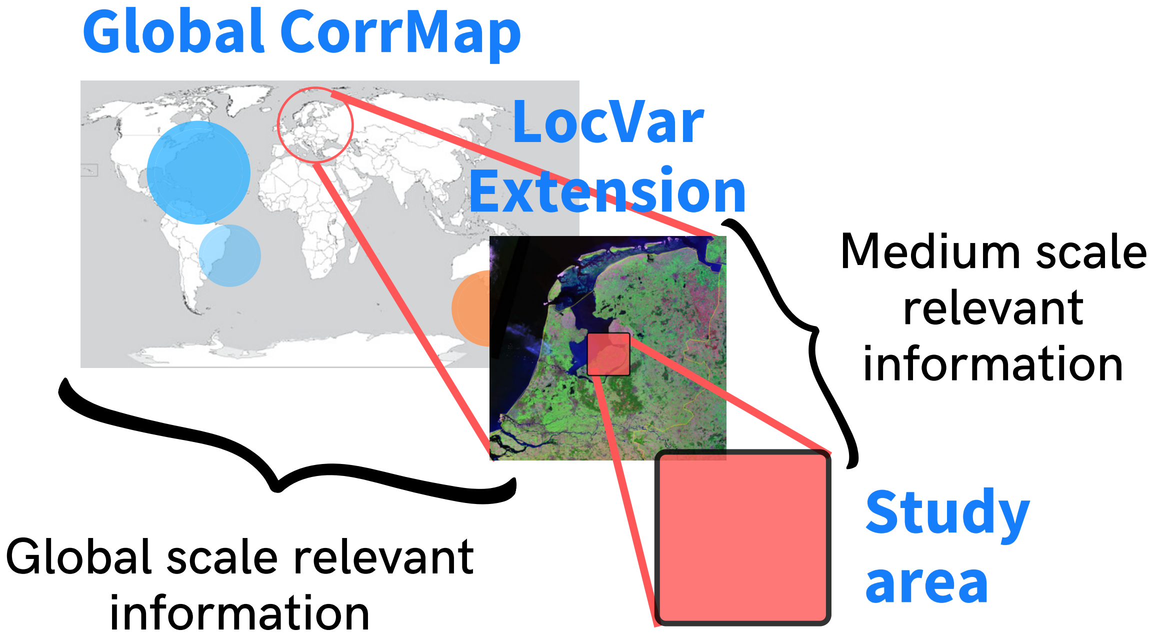
Future ideas

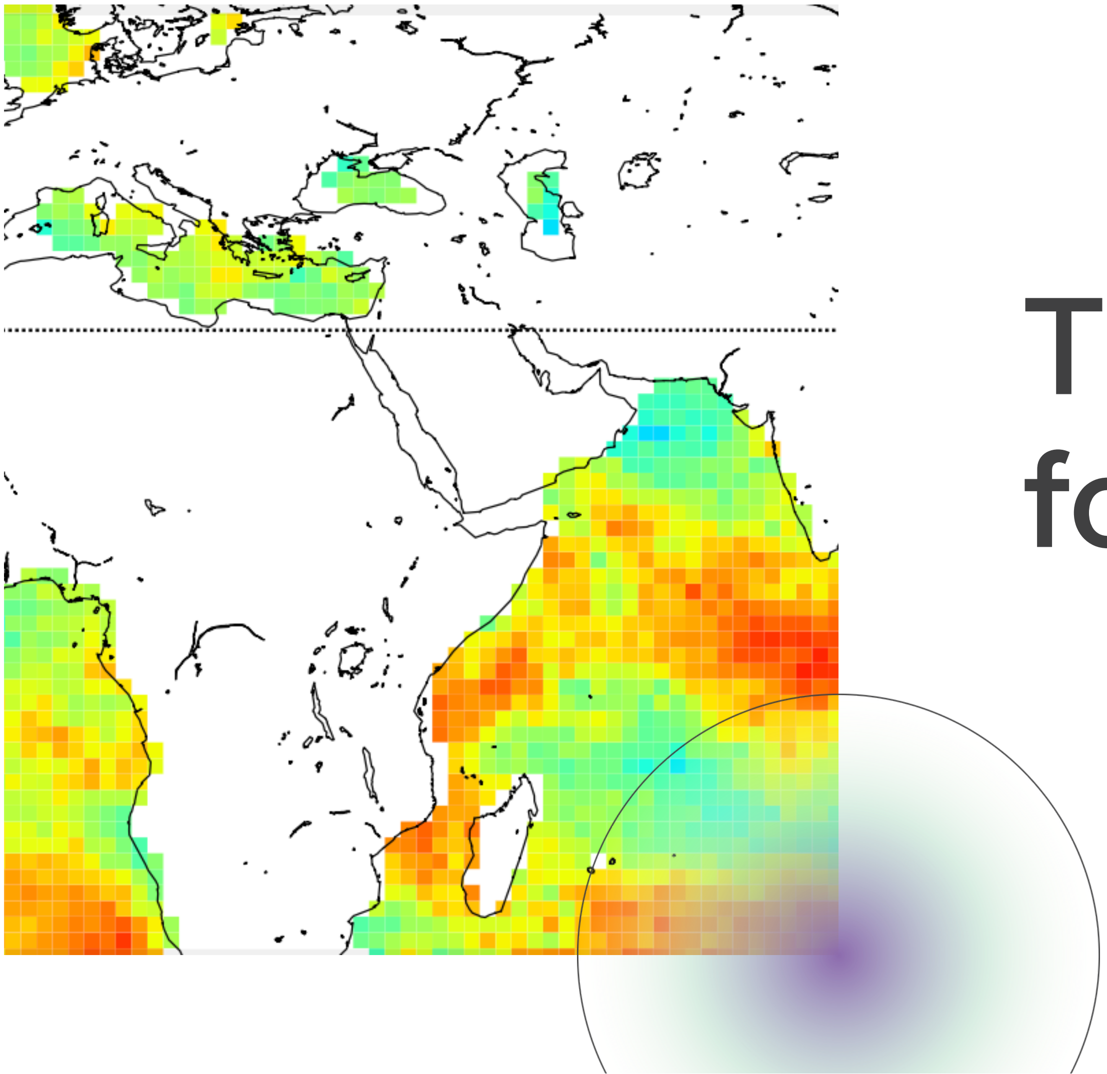




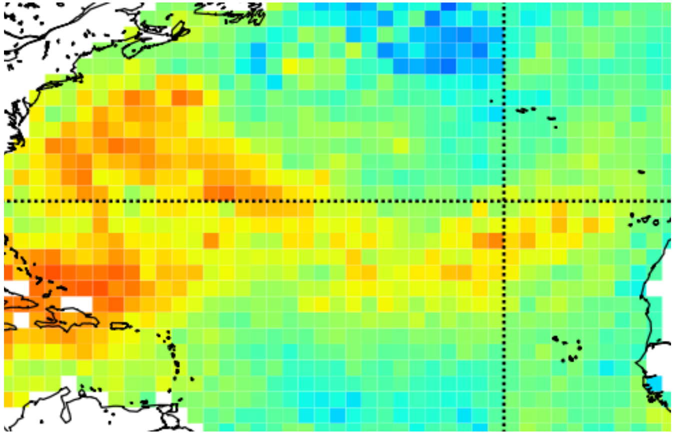
Future ideas







Thank you for attending!



you can find the slides here!

Zimmerman et al. (2016)

Giuliani et al. (2019)

Our readaptation