A new interactive and user friendly 'R ESD Lab' to explore and evaluate (even downscale) large ensembles of climate model simulations

by Abdelkader Mezghani, Rasmus Benestad, Kajsa M. Parding, Andreas Dobler
@ Norwegian Meteorological Institute

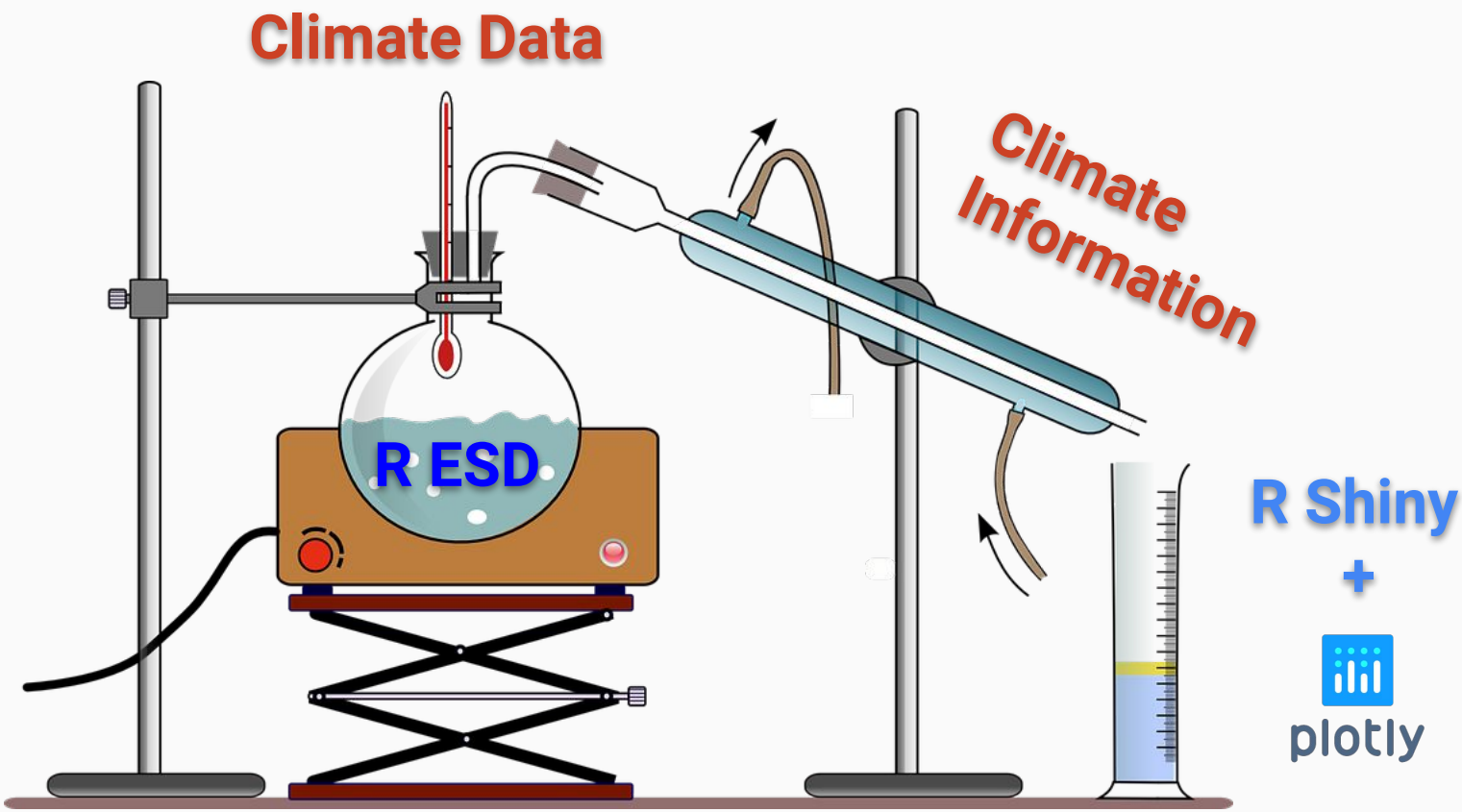
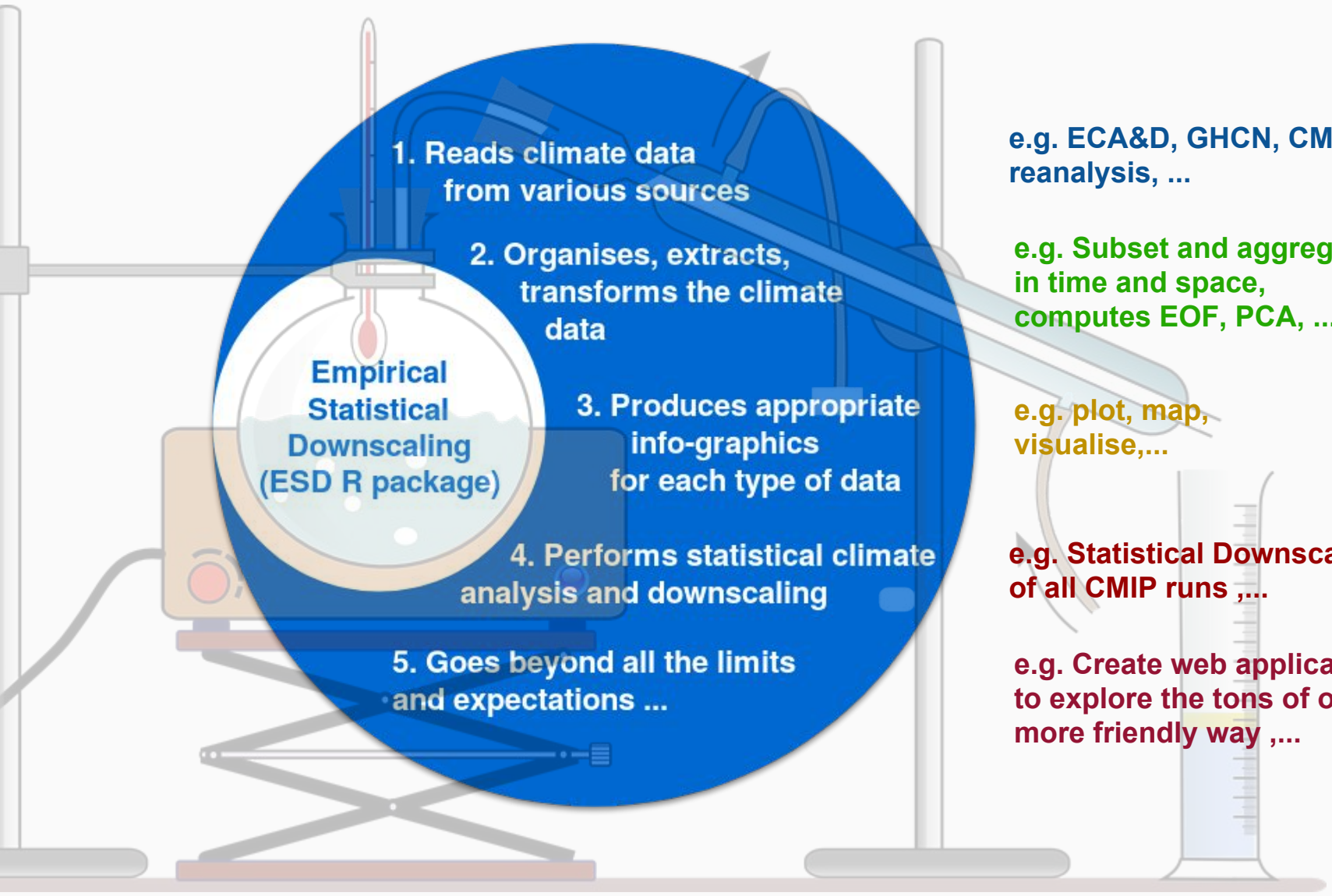


Image source: https://byjus.com/chemistry/distillation/

What does ESD R package do? All in 1!



e.g. ECA&D, GHCN, CMIP, reanalysis, ...

e.g. Subset and aggregate in time and space, computes EOF, PCA, ...

e.g. plot, map, visualise,...

e.g. Statistical Downscaling of all CMIP runs ,...

e.g. Create web applications to explore the tons of outputs in a more friendly way ,...

Data Evaluation for Climate Models (DECM prototype) - Mozilla Firefox

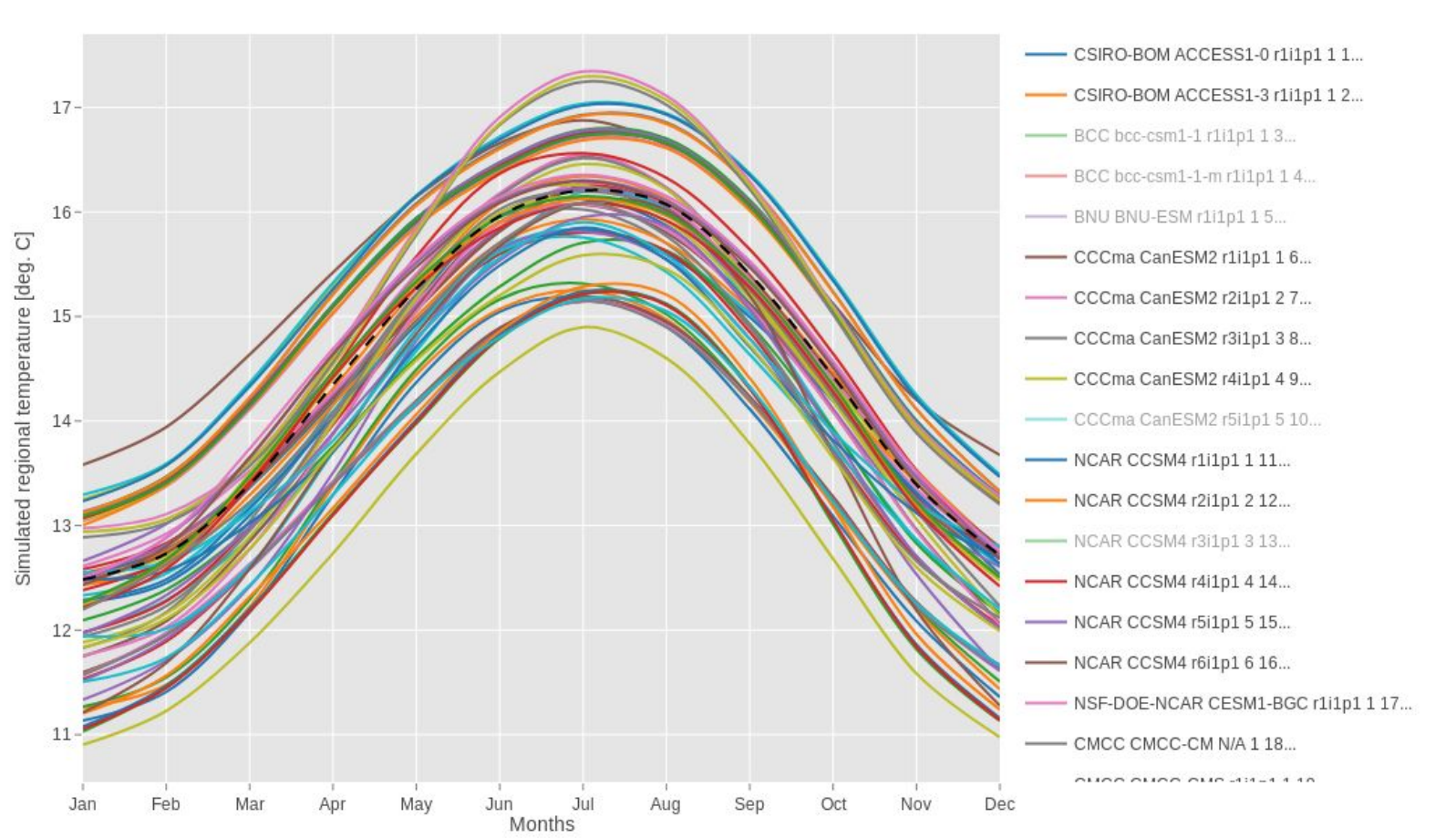
GitHub - metno/esd_Rsh | Data Evaluation for Climate

https://esdlab.met.no/decmapps/

2. Settings & Outputs : Modify the default settings and select the output type and values.

3. Evaluate the seasonal cycle in simulated Mean Air Temperature

Chart | Data | Metadata



4. Evaluate the seasonal cycle in Simulated Monthly Precipitation totals

Related abstracts on comparing ESD downscaled results with bias-corrected regional climate models and subsampling effect for two case studies <u>Norway & Poland</u>

EGU2018-18411 | Posters | CL5.05
Subsampling effect on the climate change signal based on simulations from statistical and dynamical downscaling
Abdelkader Mezghani, Rasmus E. Benestad, Andreas Dobler, Kajsa M. Parding, Jan E. Haugen, and Mikołaj Piniewski
Thu, 12 Apr, 17:30–19:00, Hall X5, X5.402

EGU2018-18952 | Posters | CL5.05
The curse of the law of small numbers haunts regional climate modelling
Rasmus Benestad, **Oskar Landgren**, Abdelkader Mezghani, Helene B Erlandsen, Andreas Dobler, Julia Lutz, Kajsa Parding, and Jan Erik Haugen
Thu, 12 Apr, 17:30–19:00, Hall X5, X5.401

EGU2018-18508 | Posters | CL5.06
Climate projections over Poland. Assessment of bias-corrected EURO-CORDEX simulations
Andreas Dobler, **Abdelkader Mezghani**, Rasmus E. Benestad, Kajsa M. Parding, Jan E. Haugen, Mikołaj Piniewski, and Zbyszek W. Kundzewicz
Fri, 13 Apr, 17:30–19:00, Hall X5, X5.488

Table of content

0. Introducing the ESD R package

- *What does the package do?*
- *What are the core concepts of the package?*

What does ‘esd’ do?

Processing tools

- Organising data for the purpose of the analysis.
- Dissecting Extracting a subset of the data.
- Transforming the data : Computing modes of variabilities, aggregating in time and space, ...

Analysing tools

- Applying statistics, describing results, hypothesis testing.
- Getting the essence from the data.
- Helps interpreting the transformation.

Core concepts of ‘esd R package’

- Generic methods (map, plot, subset, aggregate…) for different types of objects
  - R’s “S3 methods” selects appropriate functions (`map.station, map.field...`)

- attributes enhancing traceability and transparency

  e.g., `attr(x,'history'), attr(x,'source')`

- Conventions, standards, & common attributes
  - Climate and Forecast conventions (CF,CMIP)
  - csv, netCDF3/4, R binary (rda)

Meteorologisk institutt 150 år

1. ESD for reading and preparing climate data

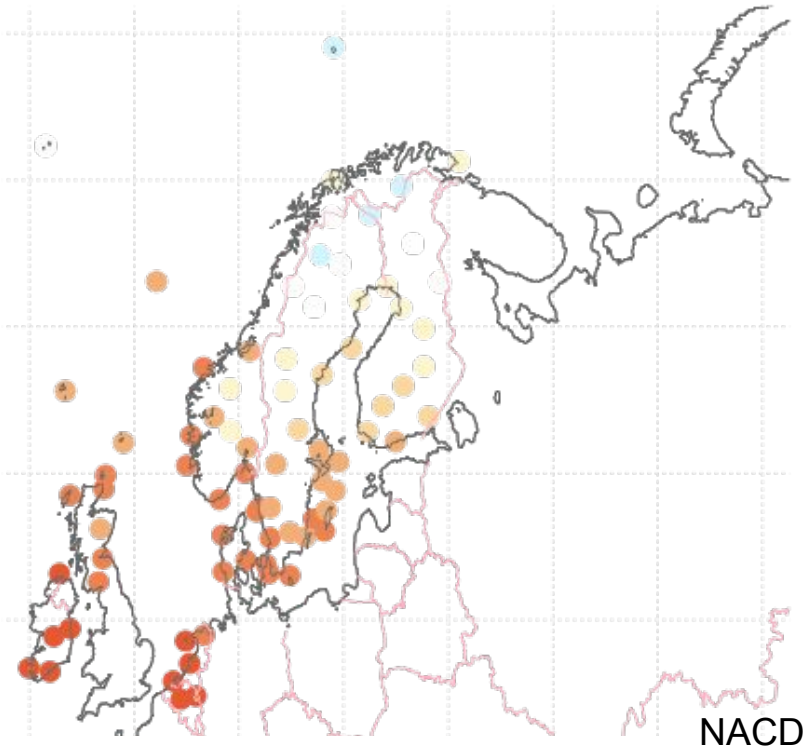
- *Read weather climate station and convert your data to a station object using 'as.station'*
- *Retrieve climate model data outputs and convert your data to a field object: as.field, as.trajectory*

‘esd’ data objects

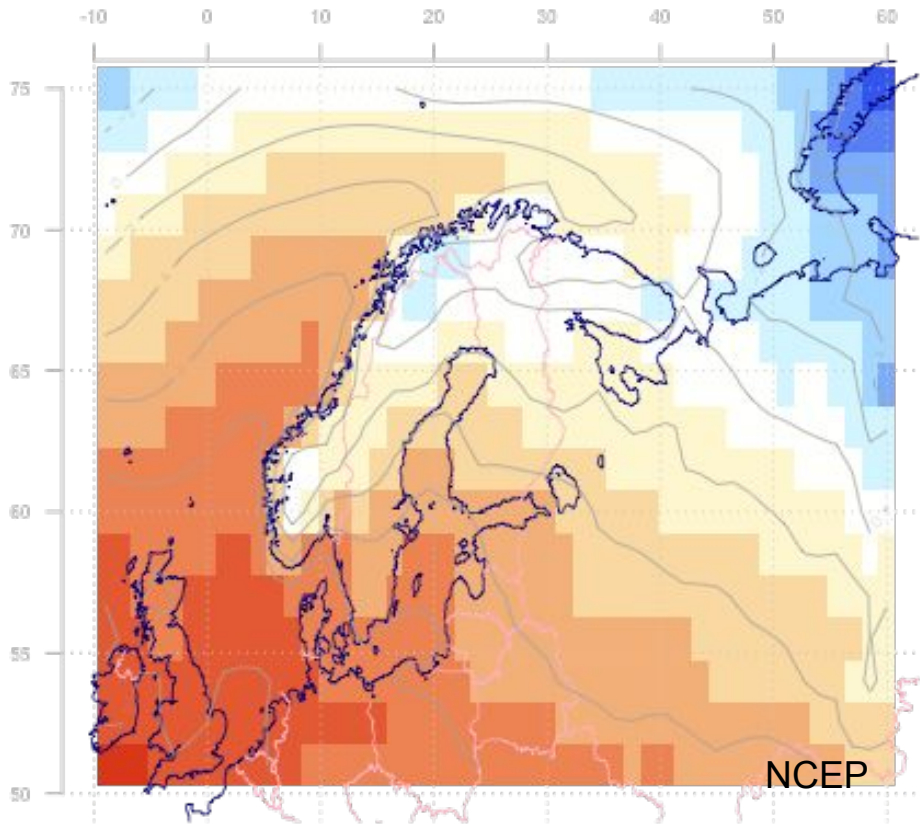
Object	Description	Example
‘station’	Single or a list of multiple time series	recorded values at weather stations
‘field’	Time series of gridded values	Reanalysis, GCMs, RCMs, EOBS
‘ds’	Time series of downscaled values	Downscaled results
‘events’	Spatio-temporal info about events	Cyclones, anticyclones
‘trajectory’	Trajectories of events	Storm tracks

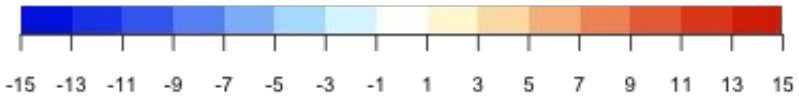
‘esd’ data objects

‘station’



‘field’





temperature (^{O}C)

Meteorologisk institutt 150 år

data

sample data included for testing

```
data(package='esd')
```

```
> Data sets in package ‘esd’:
>
> IPCC.AR5.Table.9.A.1    Sample data.
> NACD                    Sample data.
> NAOI                    Sample data.
> NARP                    Sample data.
> NINO3.4                 Sample data.
> Oslo                    Oslo monthly mean temperature time series
> Svalbard                Sample data.
> dse.Oslo                Sample data.
> dse.Svalbard
> dse.ferder              Sample data.
> eof.precip.ERAINT       Sample data.
> eof.slp.DNMI            Sample data.
…
```

Access the sample data:

```
x <- data(Oslo)
```

select.station metadata for stations

```
s <- select.station(src='GHCND', param='precip',
     nmin=50, lon=c(-60,65), lat=c(10,90))
```

src	data source
	'metnod', 'metnom' = met.no
	'ECAD' = European climate & assessment dataset
	'GHCND', 'GHCNM' = Global historical climatological network
	'NACD' = North Atlantic climatological dataset
	'NARP' = North Atlantic research program
	'NORDKLIM' = Climate data from Nordic countries
param	'`t2m`' = air temperature at 2m, '`precip`' = precipitation
cntr	country
lon, lat	longitude and latitude range
nmin	minimum length of time series (years)

Other variables and sources

- 'GLOSS' Global sea level data set
- Discharge data?

Meteorologisk institutt 150 år

station retrieve station records from a given source

```
x <- station(src='ECAD', param='t2m', nmin=50,
     cntr='Norway', lon=c(0,15), lat=c(60,70))
```

src	data source
	'metnod', 'metnom' = met.no
	'ECAD' = European climate & assessment dataset
	'GHCND', 'GHCNM' = Global historical climatological network
	'NACD' = North Atlantic climatological dataset
	'NARP' = North Atlantic research program
	'NORDKLIM' = Climate data from Nordic countries
param	'`t2m`' = air temperature at 2m, '`precip`' = precipitation
cntr	country
lon, lat	longitude and latitude range
nmin	minimum length of time series (years)

Other variables and sources

- 'GLOSS' Global sea level data set
- Discharge data?

Meteorologisk institutt 150 år

retrieve retrieves global climate and reanalysis data from a netCDF file

```
x <- retrieve(file = "filename.nc",
 it = c(1980,2010), lon=c(0,15), lat=c(60,70))
```

file	path to file
it	time index (range of years)
lon	longitude range
lat	latitude range

retrieve.rcm retrieves regional climate data from a netCDF file

as.station convert your data to a ‘station’ object

```
x <- read(file)
y <- as.station(x)
```

as.field convert your data to a ‘field’ object

```
y <- as.field(x)
```

x	object to be transformed into a station or field object

Other input may be needed depending on the class and structure of x:

lon	longitude(s)
lat	latitude(s)
param	short variable name
longname	long variable name
unit	unit name
…	

exploring the data structure

```
str(x)
View(x)
class(x)
dim(x)
length(x)

names(attributes(x))
attr(x,'variable'), attr(x,'source'), …

summary(x)
history(x)
```

Use the argument `verbose=TRUE` to print results while running a function, e.g., `plot(x, verbose=TRUE)` or `as.field(x, verbose=TRUE)`

2. ESD Lab for organising, extracting, transforming data

- *‘subset()’ a specific weather station, time coverage or a region of interest,*
- *‘aggregate’ in time e.g. monthly ‘as.monthly()’, seasons ‘as.4seasons()’ and space e.g. ‘aggregate.area()’, …*
- *transform into ‘EOF()’ or ‘PCA()’, ...*

subset

```
y <- subset(x,
            it=”djf”,
            is=list(lon=c(-30,30),lat=c(50,75)),
            ip=1,
            im=9)
```

it time range, e.g., dates `it = c(´1980-11-10´, ´1980-12-10´)`
years `it = c(1980,2010)`
months `it = c(´august´, ´september´)`
season `it = ´mam´`

is spatial range, list of the longitude and latitude range

ip index for pattern and pca selection

im index for climate model selection

aggregate

Aggregate in space

```
y <- aggregate.area(x, FUN='mean')
```

Aggregate in time

seasonal cycle:

```
y <- aggregate(x, by=month(x), FUN='mean')
```

annual mean:

```
y <- aggregate(x, by=year(x), FUN='mean')
y <- annual(x, FUN='mean')
```

seasonal and monthly mean:

```
y <- as.4seasons(x, FUN='mean')
y <- as.monthly(x, FUN='mean')
```

The argument `FUN` *can be, e.g., ´mean´, ´sd´, ´sum´, ´trend´, 'wetfreq', 'wetmean', 'count', 'spell', 'your own function', ...*

What is PCA/EOF?

A way of organising the data according to information contents

Similar to

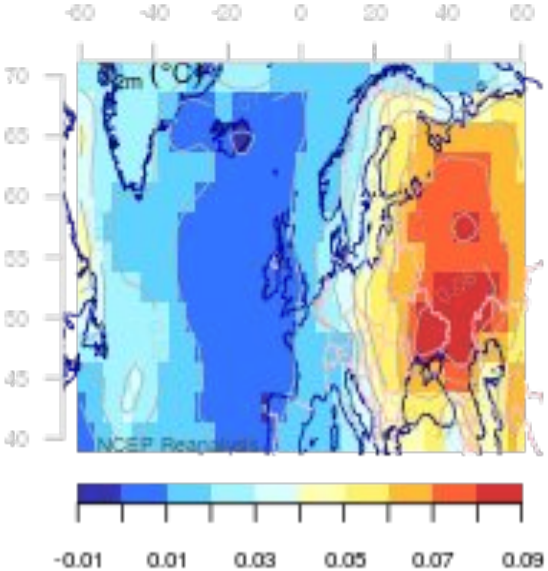
- Eigenvalues
- Empirical Orthogonal Functions (EOFs)
- Fourier series

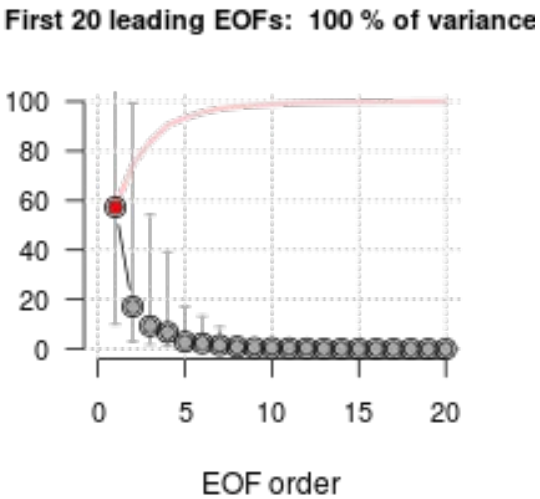
EOF empirical orthogonal functions

```
X <- EOF(x, n=20)
plot(X)
```

x	a ‘field’ object
n	number of EOFs

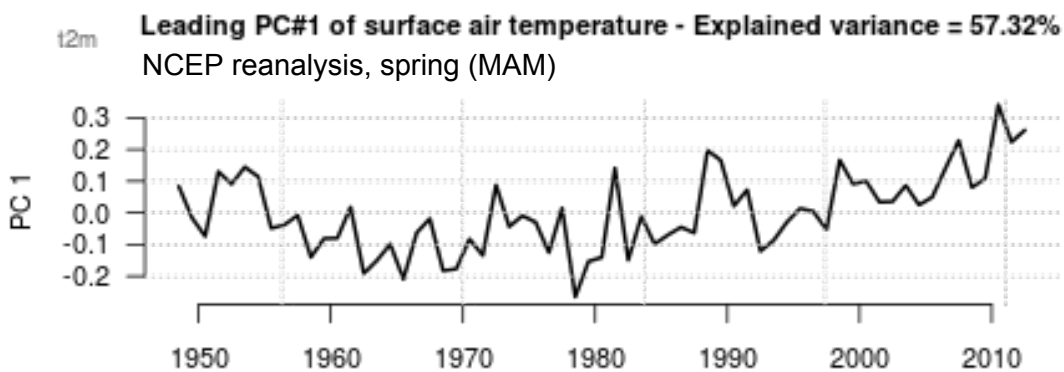
spatial patterns





amplitudes

time series

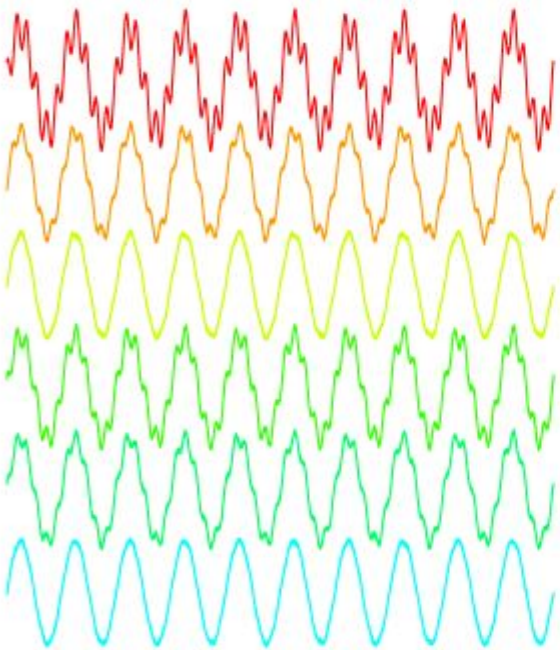


Meteorologisk institutt 150 år

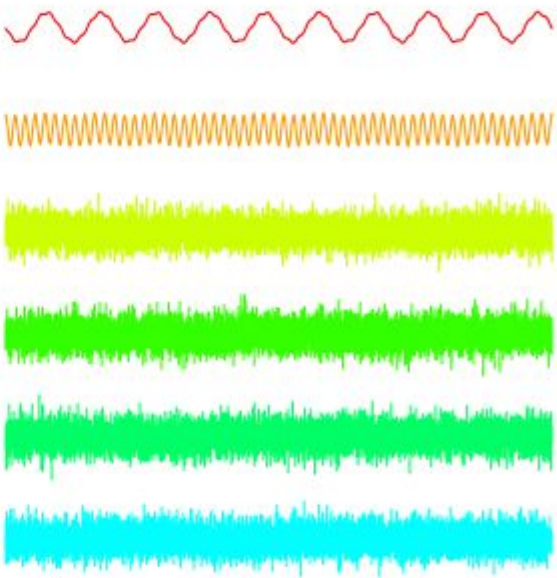
What is PCA?

$$y = \sum_{i=1}^{n} \beta_i x_i$$

$$x_i$$



Original data



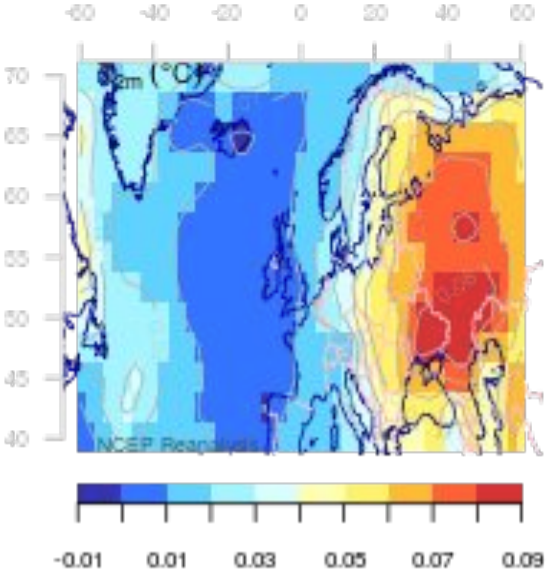
Principal components (PCs)

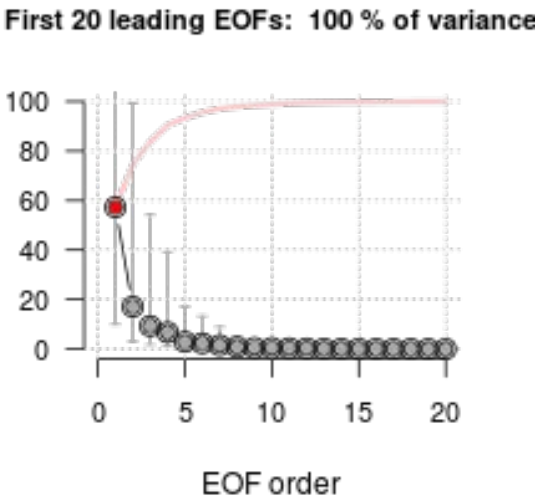
EOF empirical orthogonal functions

```
X <- EOF(x, n=20)
plot(X)
```

x	a ‘field’ object
n	number of EOFs

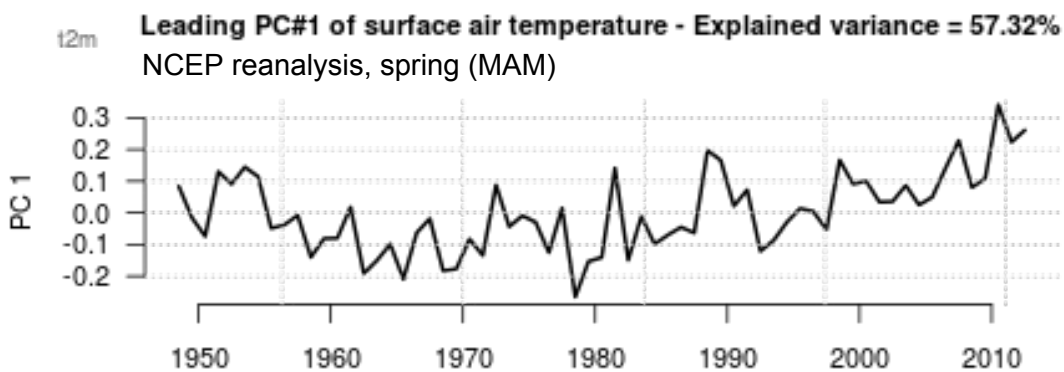
spatial patterns





amplitudes

time series



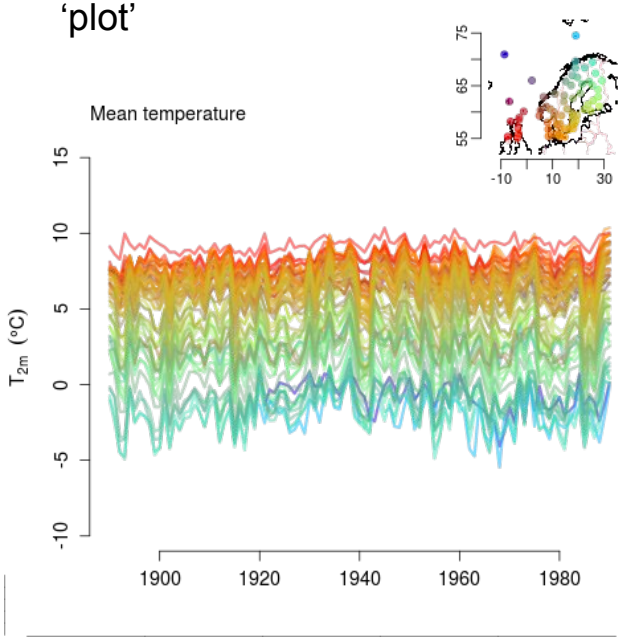
Meteorologisk institutt 150 år

3. ESD for visualising, producing appropriate infographics

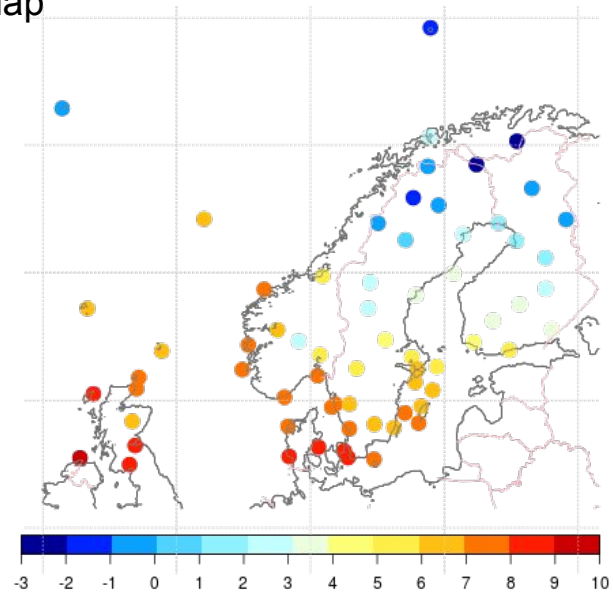
e.g.,

- *‘plot()’ stations, fields, an EOF, a PCA, ... objects*
- *‘map()’ stations, fields, EOFs, PCAs, ... objects*
- *‘vis()’*ualise e.g. wheel charts, visualise trends, ...*

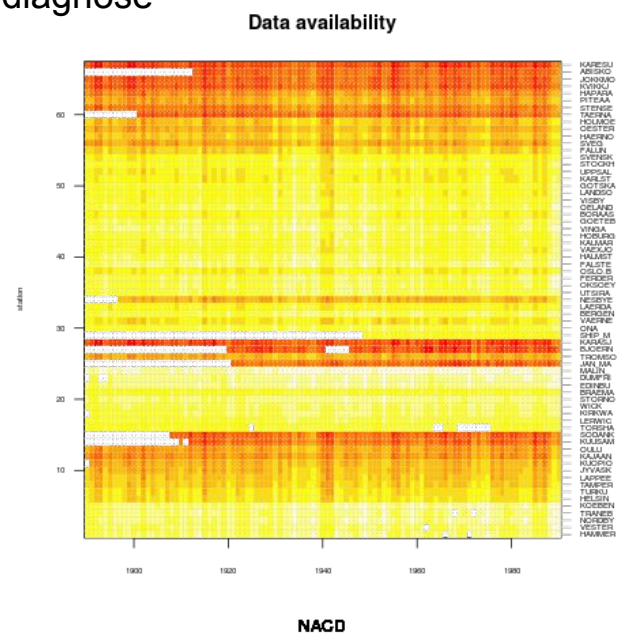
‘plot’



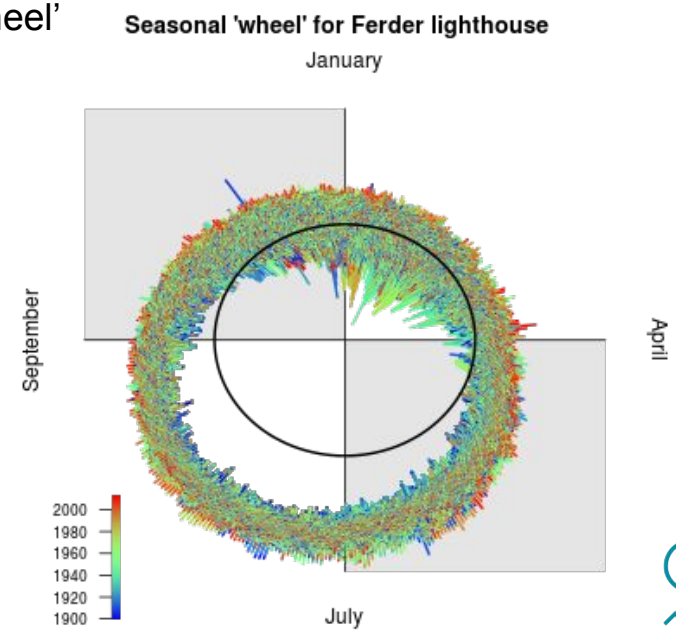
‘map’



‘diagnose’



‘wheel’



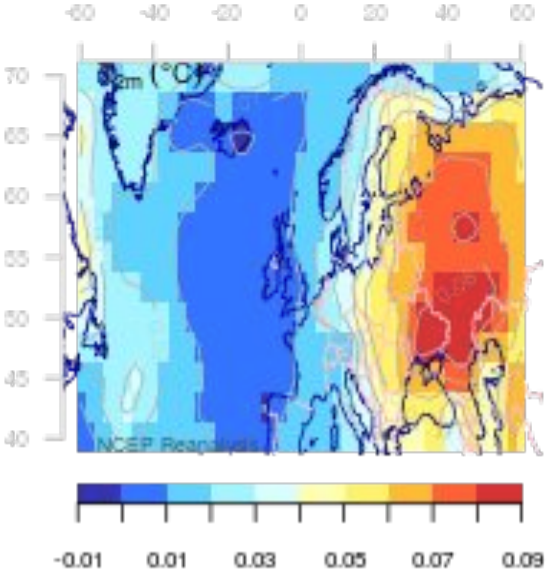
Meteorologisk institutt 150 år

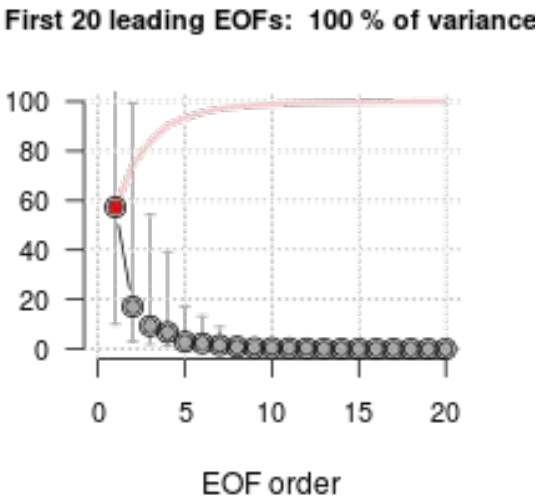
EOF empirical orthogonal functions

```
X <- EOF(x, n=20)
plot(X)
```

x	a ‘field’ object
n	number of EOFs

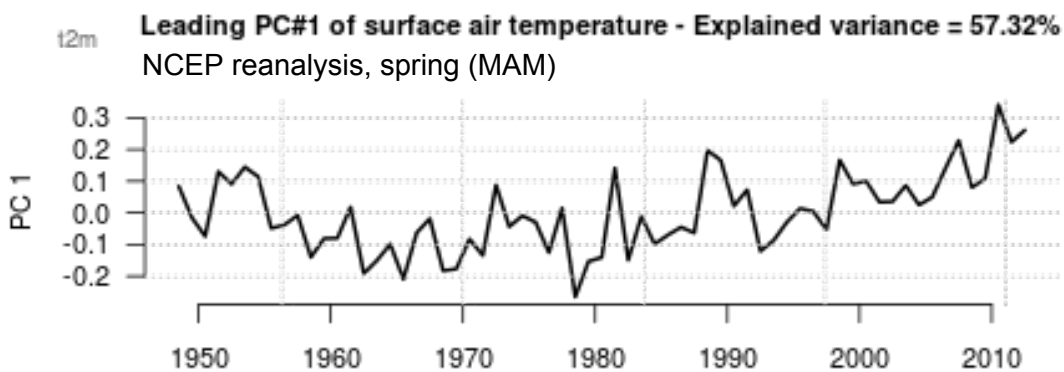
spatial patterns





amplitudes

time series

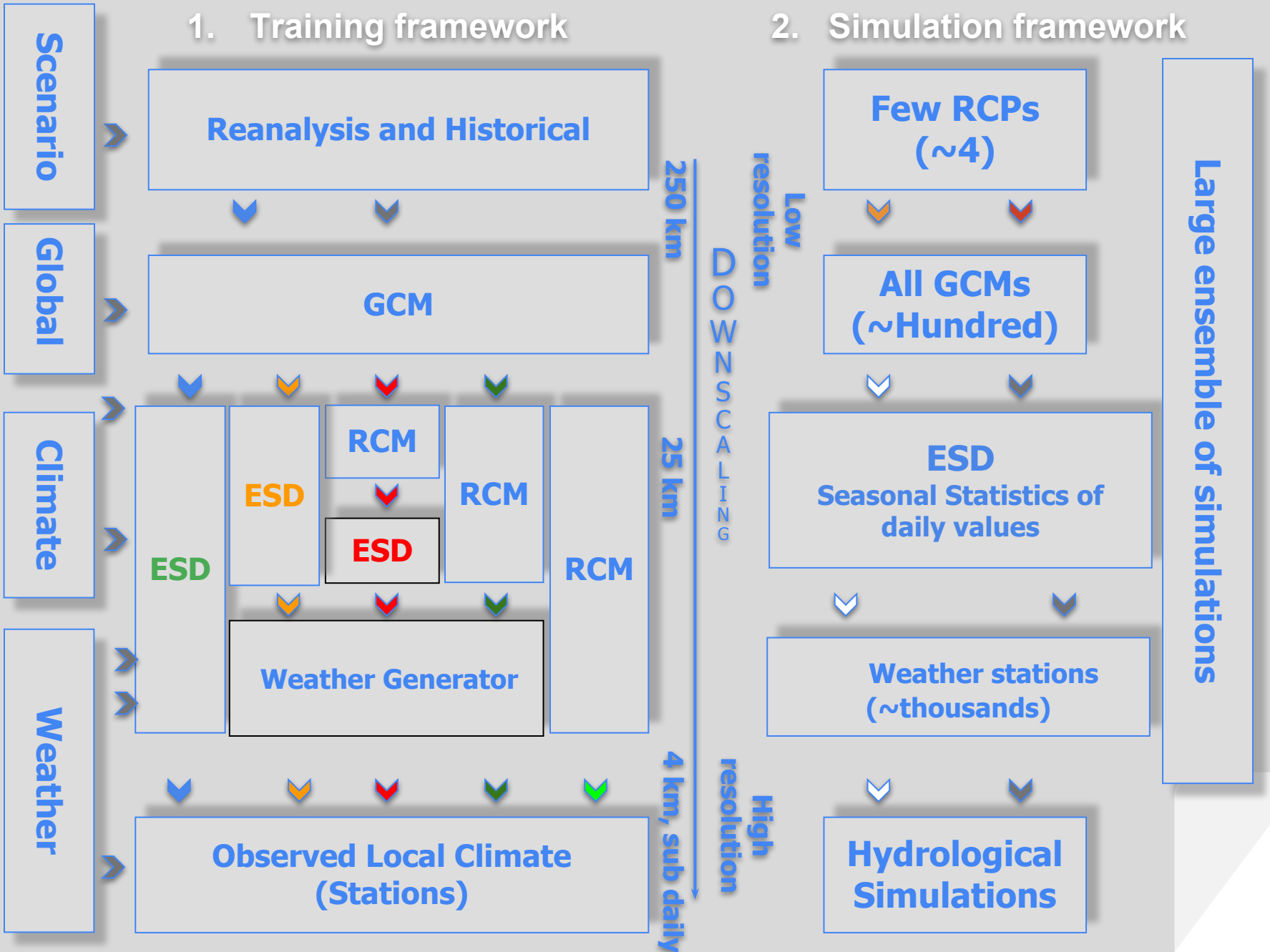


Meteorologisk institutt 150 år

4. Statistical analysis and downscaling

- *Empirical orthogonal functions ‘EOF function’,*
- *Principal component Analysis ‘PCA function’,*
- *Empirical statistical Downscaling ‘DS function’*

Examples of where ESD can be implemented in your simulation framework !



Statistical models usually consist of an empirical relationship for which a set of parameters are estimated so that it best represents the patterns found in the data.

Local-Scale = f(Geography + Large-Scale) + Noise

$$Y = f(G,X) + N$$

Y is the predictand (Quantification of the local scale patterns)

X is the predictor (Quantification of the large or regional scale patterns)

N is a Noise term (Stochastic process)

DS empirical-statistical downscaling

Establish a connection between local observations and large scale climate patterns (EOFs of reanalysis data) through multiple linear regression.

```
ds <- DS(y, X, eofs=1:5)

attr(ds,'model')
```

```
> Call:
>  lm(formula = y ~ X.1 + X.2 + X.5, data = caldat)
>
>  Coefficients:
>  (Intercept)          X.1          X.2          X.5
>  -8.487e-19    7.518e+00    3.939e+00    5.102e+00
```

DS empirical-statistical downscaling

Establish a connection between local observations and large scale climate patterns (EOFs of reanalysis data) through multiple linear regression.

```
ds <- DS(y, X, eofs=1:5)
```

y	**predictand**: object to be downscaled, typically a station object (can also be a PCA or EOF object).
X	**predictor**: an EOF object that represents the large scale climate (if X is a field, EOF analysis is applied within the DS function)
eofs	predictor patterns to use for downscaling

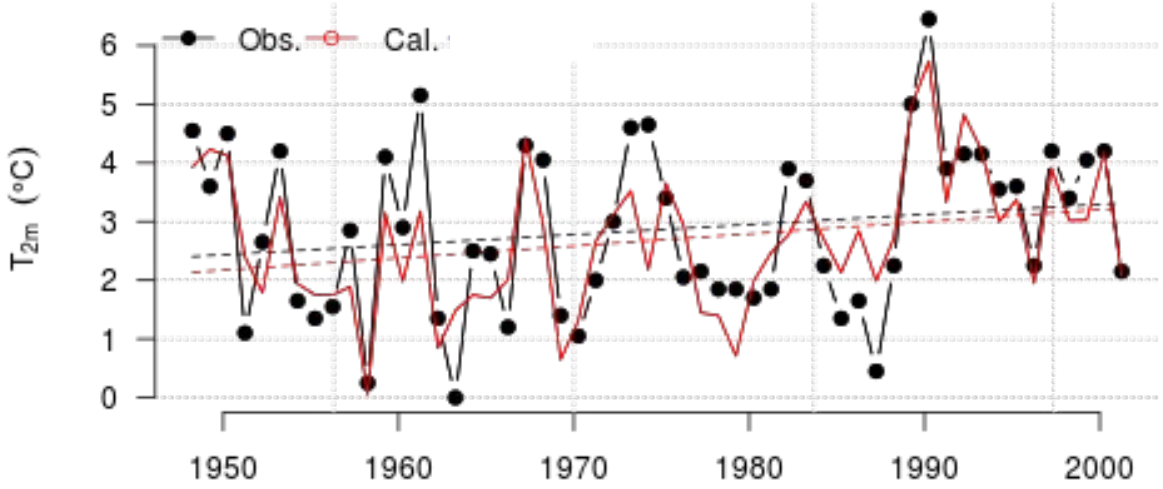
DS empirical-statistical downscaling

Establish a connection between local observations and large scale climate patterns (EOFs of reanalysis data) through multiple linear regression.

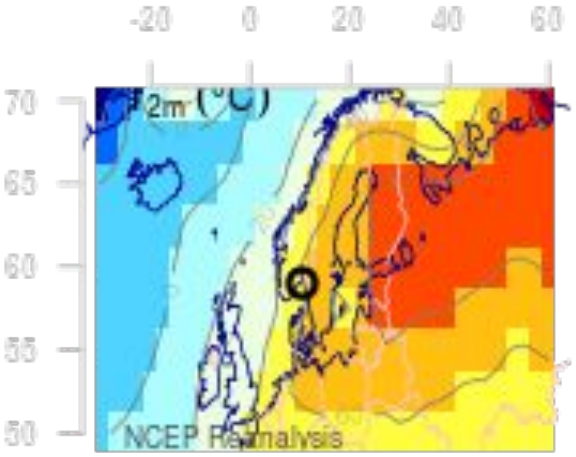
```
ds <- DS(y, X, eofs=1:5)

plot(ds)
```

CALCULATED TIME SERIES



1st PREDICTOR PATTERN



Meteorologisk institutt 150 år

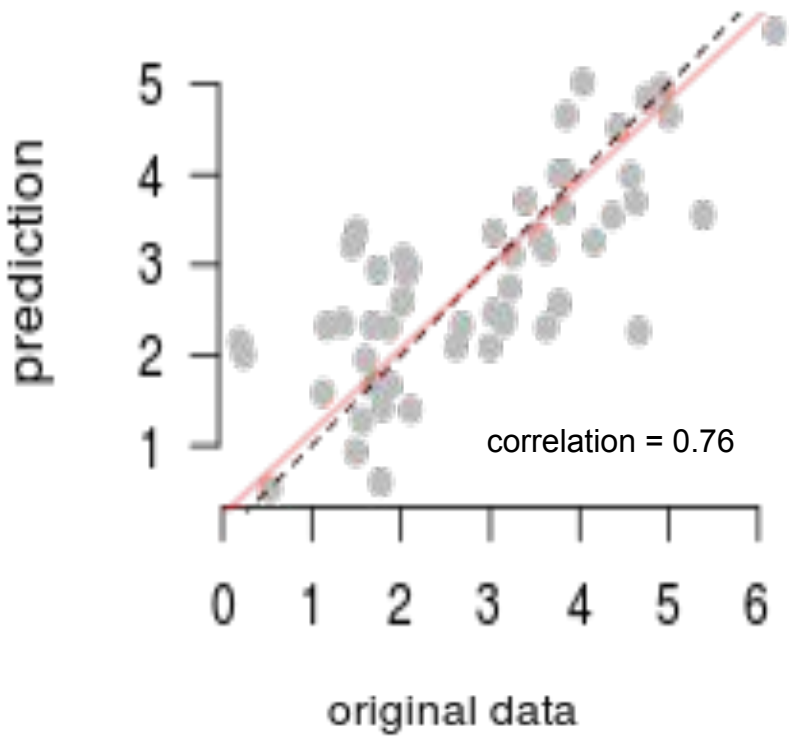
DS empirical-statistical downscaling

Establish a connection between local observations and large scale climate patterns (EOFs of reanalysis data) through multiple linear regression.

```
ds <- DS(y, X, eofs=1:5)

plot(ds)
```

CROSS-VALIDATION

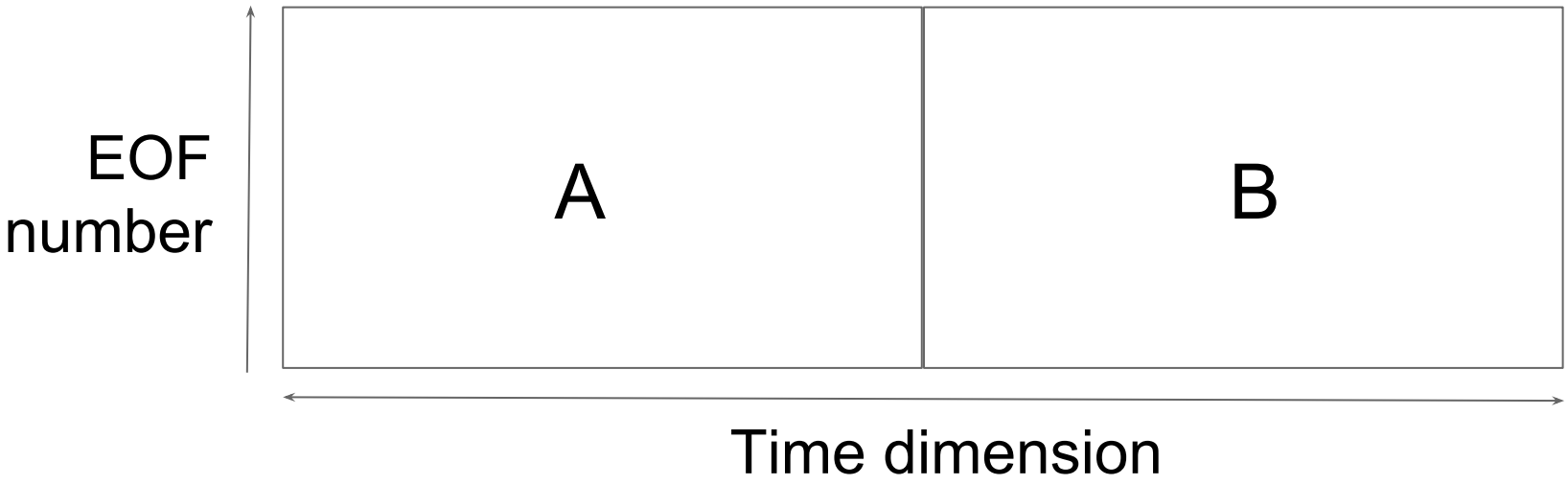


Meteorologisk institutt 150 år

combine & EOF.comb

common EOFs

Two (or more) field objects are combined before EOF analysis is applied. This is done to identify common spatial patterns in the data sets.

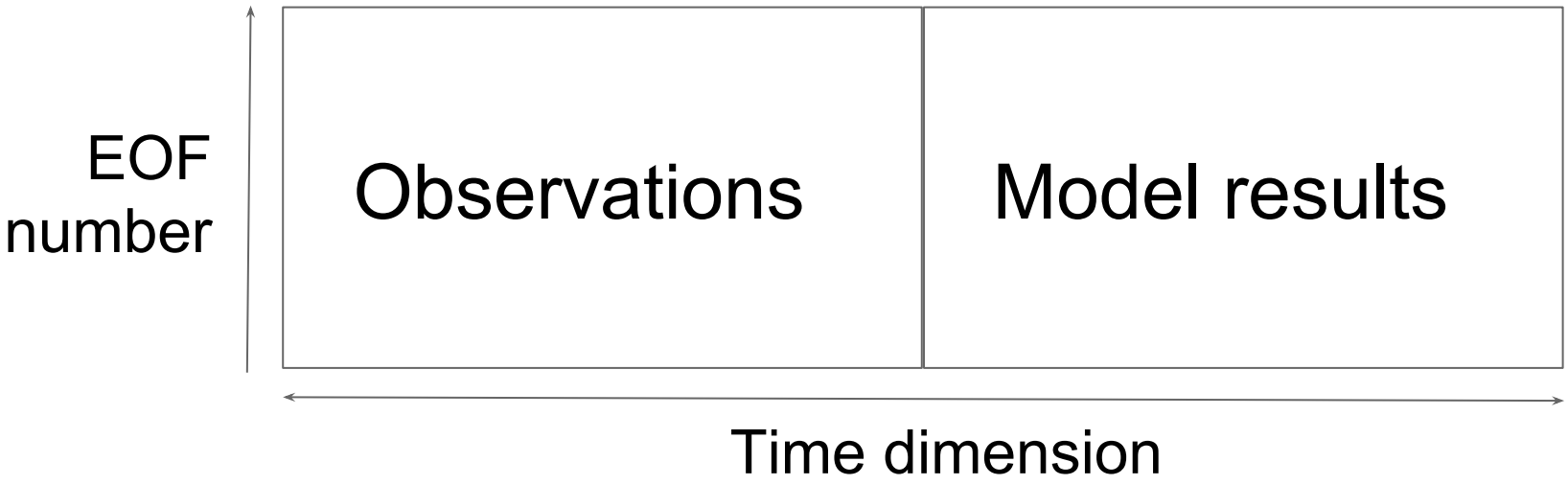


$$t = [t_A, t_B] = 1, \ldots, N$$

combine & EOF.comb

common EOFs

Two (or more) field objects are combined before EOF analysis is applied. This is done to identify common spatial patterns in the data sets.



$$t = [t_{obs}, t_{model}] = 1, \ldots, N$$

combine & EOF.comb

common EOFs

Two (or more) field objects are combined before EOF analysis is applied. This is done to identify common spatial patterns in the data sets.

```
Z <- combine(X,Y)
eof.Z <- EOF(Z)
```

X	a field object (e.g., reanalysis data)
Y	another field object (e.g., GCM results)

DSensemble downscale a GCM ensemble

Establish a connection between local observations and large scale climate patterns (EOFs of reanalysis data) through multiple linear regression.

Apply the empirical model to GCM data.

```
dse <- DSensemble(y, predictor=X, eofs=1:8, select=NULL,
    path='CMIP5/', rcp='rcp45', pattern='psl_Amon_ens_',
    lon=c(-20,20), lat=c(-10,10), rel.cord=TRUE, ...)
```

Y	predictand, e.g., a station object
predictor	predictor, either a field object or a path and filename of a netCDF file with reanalysis data
lon, lat	predictor domain (longitude and latitude range)
rel.cord	if rel.cord = TRUE, lon and lat is interpreted as relative to the location of the predictand
eofs	eofs to use as predictors
pattern, path, rcp	the function looks for GCM data in the directory 'path/rcp/' in files with names following the defined 'pattern'
select	GCMs to include, NULL = all available models

Meteorologisk institutt 150 år

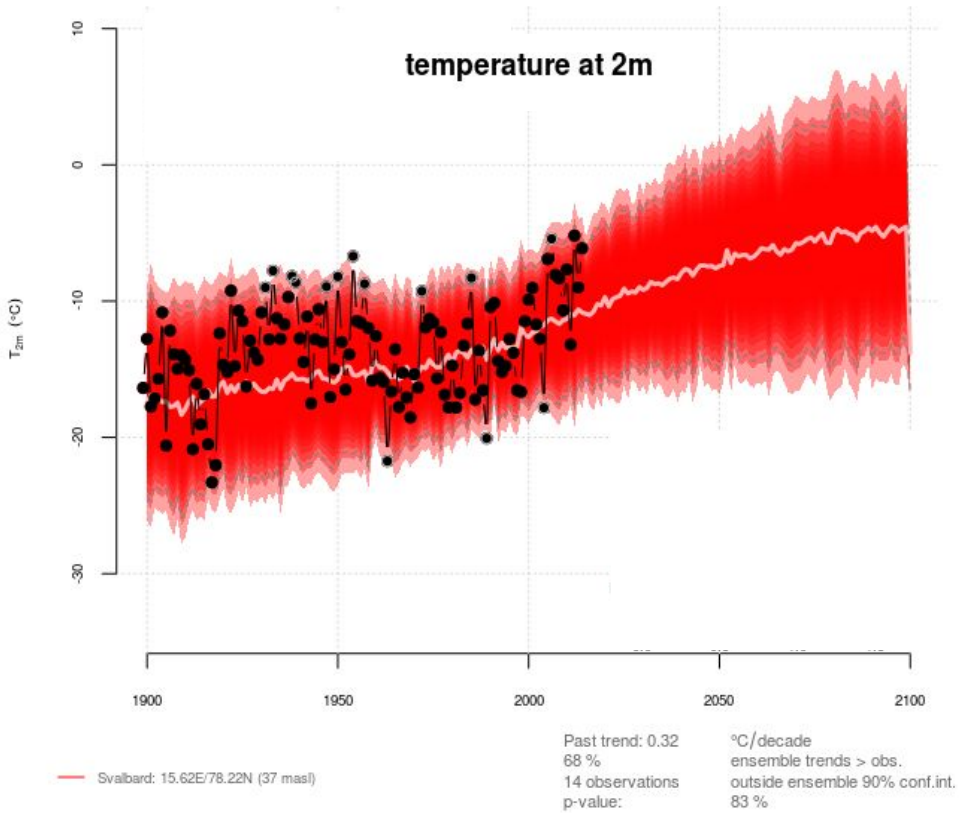
DSensemble downscale a GCM ensemble

Establish a connection between local observations and large scale climate patterns (EOFs of reanalysis data) through multiple linear regression.

Apply the empirical model to GCM data.

```
dse <- DSensemble(y, predictor=X, ...)

plot(dse)
```



Meteorologisk institutt 150 år

5. ESD Lab goes beyond the limits

- *Interactive mapping and plotting using ‘plotly’ package*
- *Creating ‘Web Apps’ using ‘R Shiny & R Shinydashboard’ packages*
- *Publish and share your findings*
- *Track the usage and collect feedback using ‘Google Analytics’*

A new interactive and user friendly 'R ESD Lab' to explore and evaluate (even downscale) large ensembles of climate model simulations

by Abdelkader Mezghani, Rasmus Benestad, Kajsa M. Parding, Andreas Dobler
@ Norwegian Meteorological Institute

Twofold recipe for successful climate application based on ESD R and other R packages

1. **ESD R package (back-end)**
2. **Shiny & Plotly R packages (front-end) application**

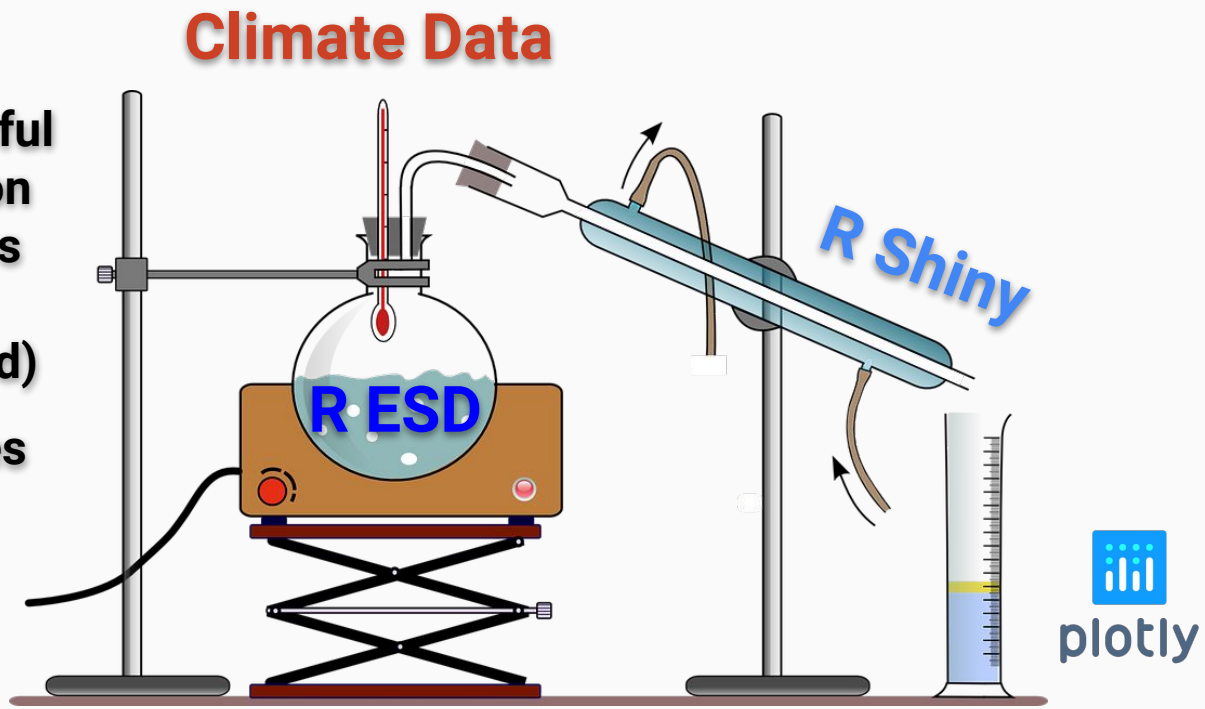


Image source: https://byjus.com/chemistry/distillation/

R-shiny apps

https://esdlab.met.no/

- Set of sample apps
- Uses dynamic dashboards
- SSL certificate for the shiny server (https://...)
- Uses back-end and the esd package as an engine
- User perspective: different requirements for data users and product users based on conditional panels and dynamic web pages
- Information from data (sophisticated statistical methods) using esd R and other packages

https://climatedatasite.net 60% Search

PROOF OF CONCEPT

Live example from the climatedatasite.net ! Try it out on your computer.

Copernicus Europe's eyes on Earth — Climate Change Service

HOME DEMO VIDEOS ESD LAB ABOUT

To evaluate climate models on a global and regional scales, we propose a straightforward comparison of climate models' ability to reproduce statistical properties such as th figure shows an example w ons.

You can open each w

1. Select your

Data Eval

1. Select a reg

2. Settings &

3. Evaluate th

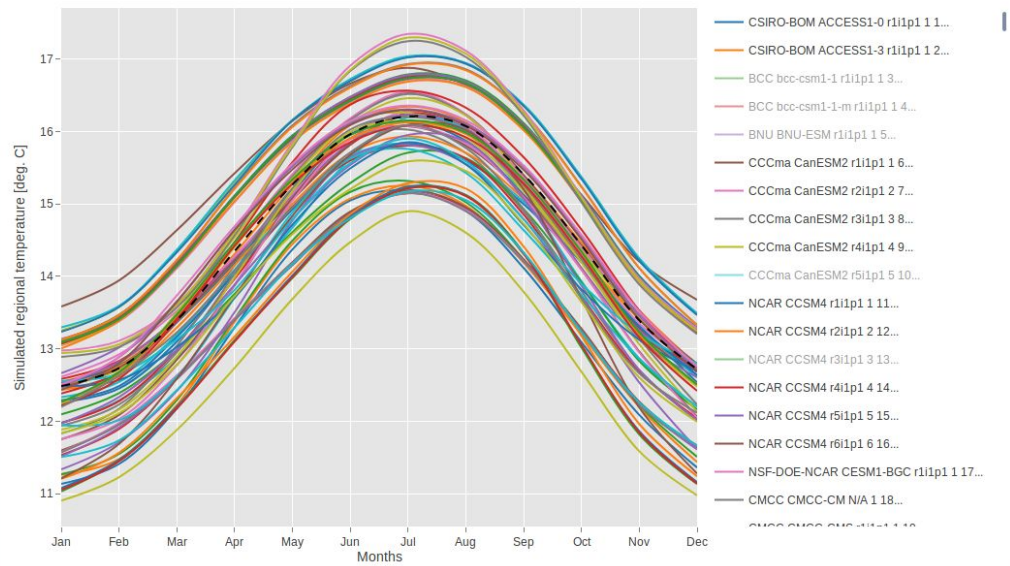
Chart Da

The interactive fig intermediate emi the ensemble mo climate models fr standard deviatio compare data bet details about the s

3. Evaluate the seasonal cycle in simulated Mean Air Temperature

Chart Data Metadata

The interactive figure shows the seasonal cycle of pseudo-observed (dashed) and modeled air mean temperature by the multi-model ensemble of simulations assuming the intermediate emission scenario (RCP4.5). The user can modify the type of the output from the "Settings & Outputs" tab box into, for example, individual simulations, envelope of the ensemble model simulations, box plots of both, transform the values into anomalies, group the models by attributes, etc. The user can additionally double click on specific climate models from the legend (once displayed) or the meta data table to isolate one or a group of simulations or modified the displyed statistic to, for example, spatial standard deviation and correlations instead of the mean. Other options are also included such as zoom in/out, show closest data by pointing with the mouse on the simulations, compare data between simulations, and download the plot as png by taking a snapshot. The user can also check and download both the data and meta data tabs for furhter details about the simulations.



4. Evaluate the seasonal cycle in Simulated Monthly Precipitation totals

https://climatedatasite.net 60% Search

PROOF OF CONCEPT

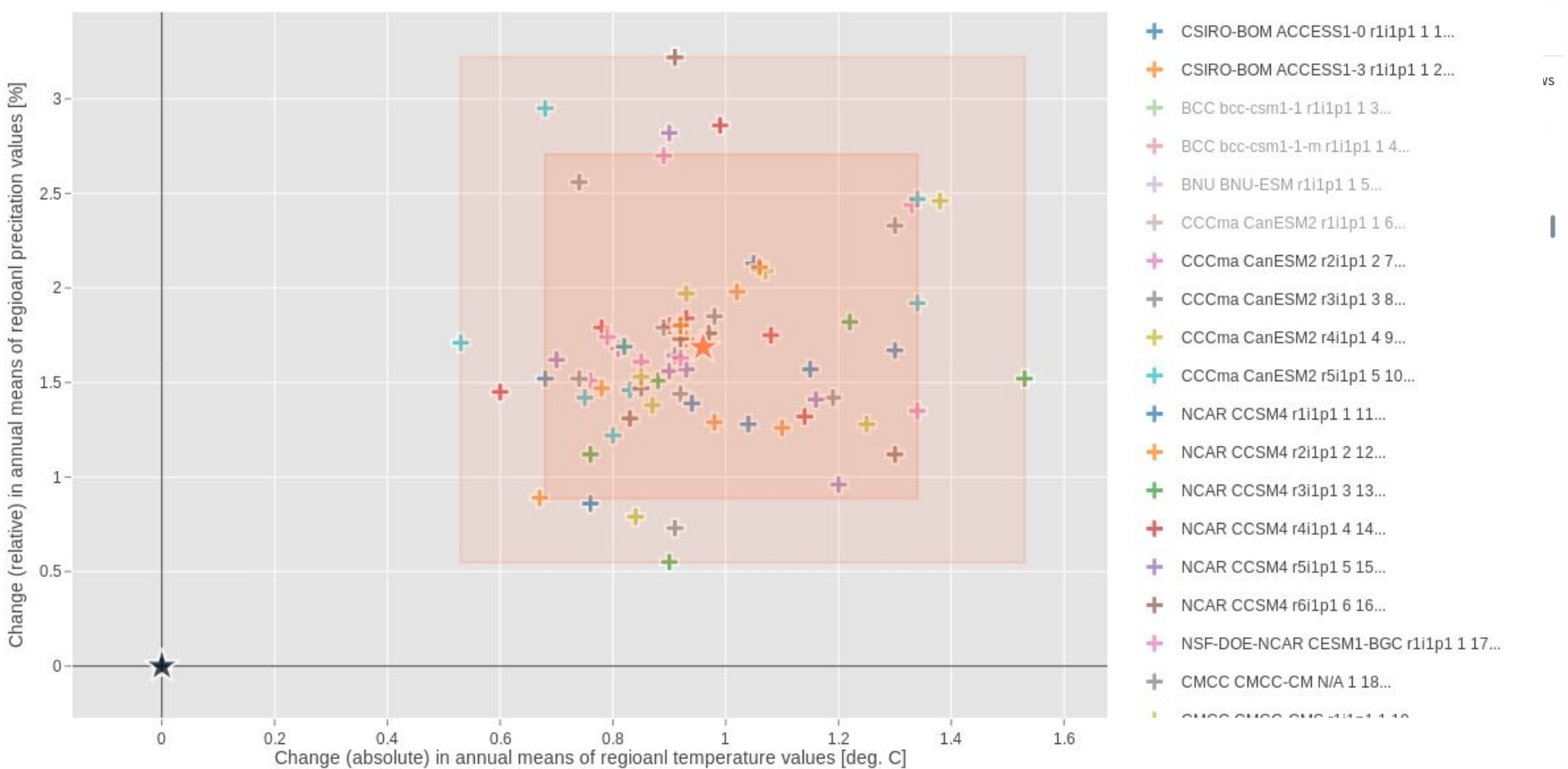
Copernicus Europe's eyes on Earth

Climate Change Service

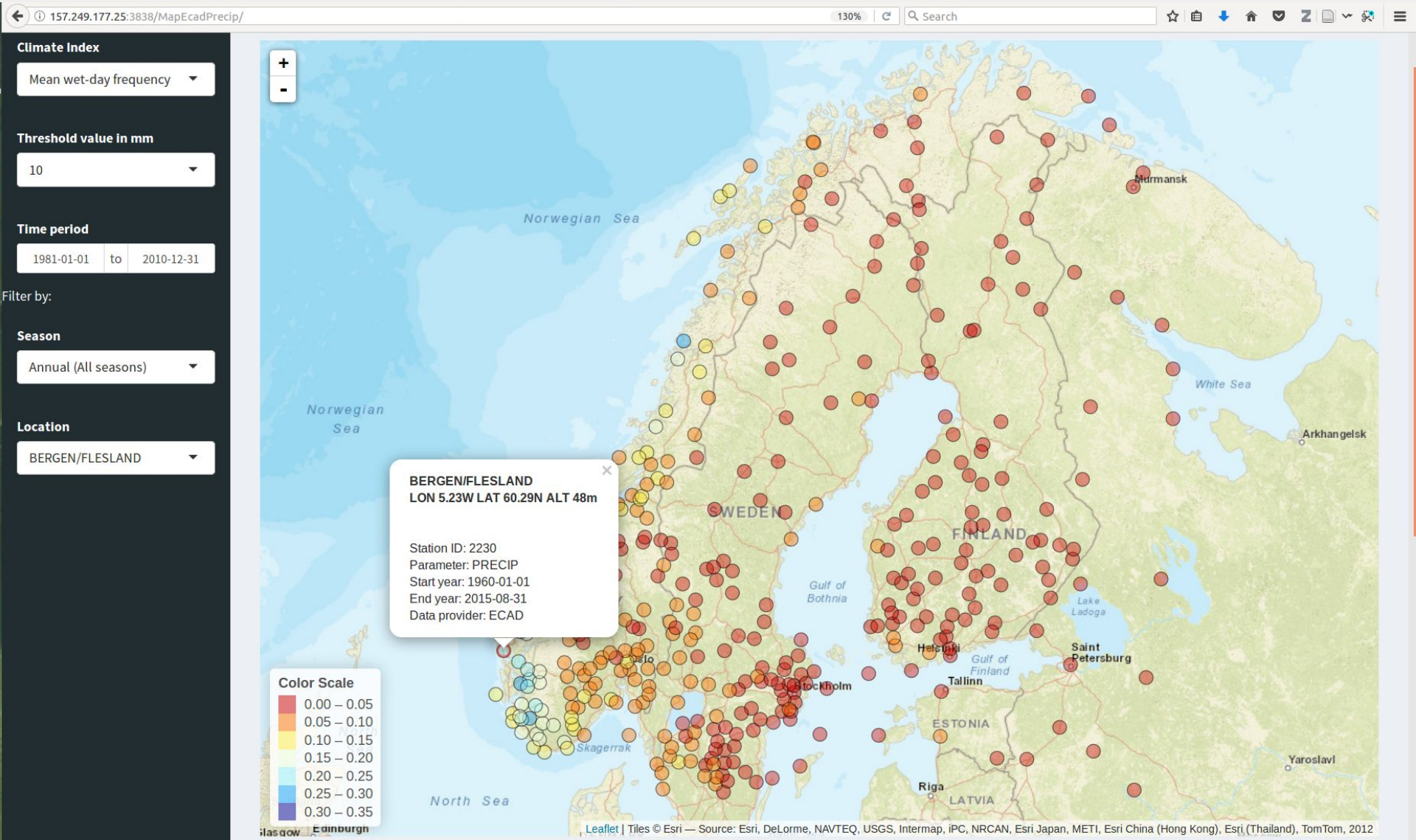
HOME DEMO VIDEOS ESD LAB ABOUT

https://esdlab.met.no/decmapps/ 130% Search

5. Scatter Plots of Simulated Climate Variables

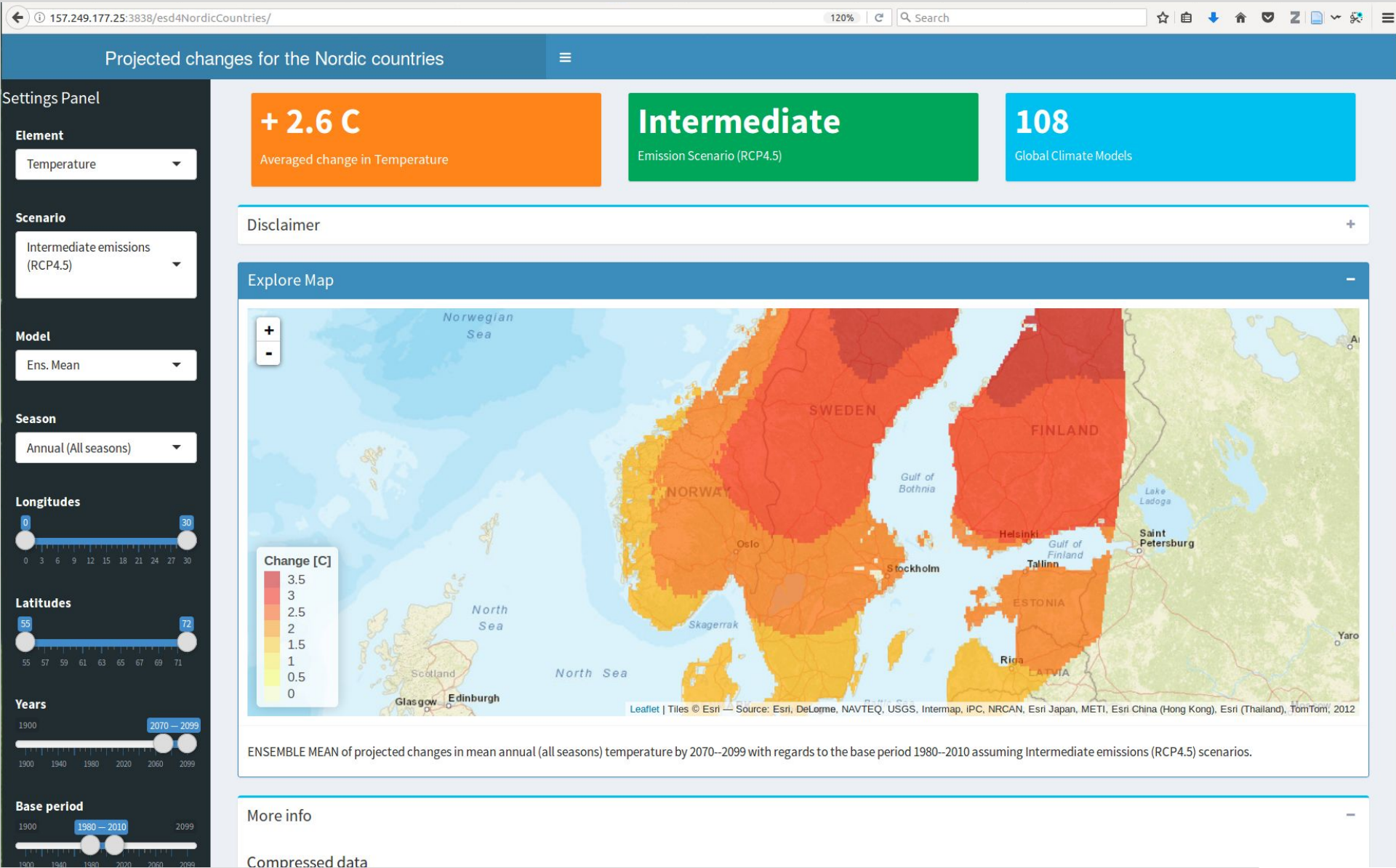


Climate Mapping Toolkit



Annual mean of Wet-day mean

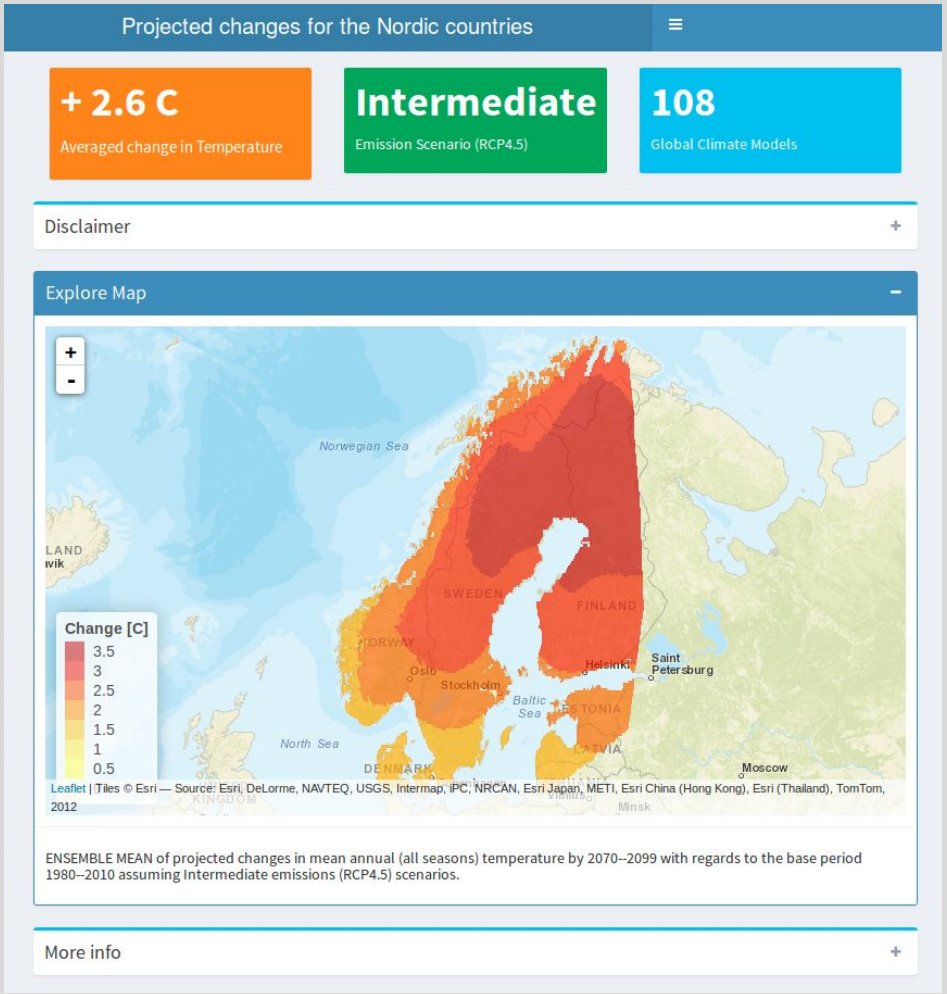
[ESD] **Climate Projections Toolkit for the Nordic Countries**



Web applications for non advanced users

http://157.249.177.25:3838/esd4NordicCountries/

- New ways of exploring climate change
- Dynamic web page ‘leaflet’ interactive maps
- Concise information using value info-boxes
- All settings in dashboards
- Pop up window for metadata
- Collapsible boxes



More information and resources

The esd-package www.github.com/metno/esd/wiki

www.facebook.com/Rclimateanalysis/

`manual(), ABC4ESD()`

The esd laboratories:

- ESD Lab within the C3S DECM project at https://esdlab.met.no/
- ESD Lab within the eSACP project at http://157.249.177.25:3838