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!

1. Reads climate data from various sources
2. Organises, extracts, transforms the climate data
3. Produces appropriate info-graphics for each type of data
4. Performs statistical climate analysis and downscaling
5. Goes beyond all the limits and expectations ...

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

3. Evaluate the seasonal cycle in simulated Mean Air Temperature

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 Norway & Poland

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. ESD package

1. ESD for reading and preparing climate data

2. ESD for processing the data

3. ESD for appropriate infographics

4. ESD for statistical analysis and downscaling

5. ESD Lab goes beyond the limits
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)
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

<table>
<thead>
<tr>
<th>Object</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘station’</td>
<td>Single or a list of multiple time series</td>
<td>recorded values at weather stations</td>
</tr>
<tr>
<td>‘field’</td>
<td>Time series of gridded values</td>
<td>Reanalysis, GCMs, RCMs, EOBS</td>
</tr>
<tr>
<td>‘ds’</td>
<td>Time series of downscaled values</td>
<td>Downscaled results</td>
</tr>
<tr>
<td>‘events’</td>
<td>Spatio-temporal info about events</td>
<td>Cyclones, anticyclones</td>
</tr>
<tr>
<td>‘trajectory’</td>
<td>Trajectories of events</td>
<td>Storm tracks</td>
</tr>
</tbody>
</table>
‘esd’ data objects

‘station’

‘field’

Temperature (°C)
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           Sample data.
> 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?
station retrieve station records from a given source

\[
x <- \text{station}(\text{src}='\text{ECAD}', \text{param}='\text{t2m}', \text{nmin}=50, \text{cntr}=\text{Norway}', \text{lon}=c(0,15), \text{lat}=c(60,70))
\]

**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?
**retrieve** retrieves global climate and reanalysis data from a netCDF file

\[
x \leftarrow \text{retrieve}(\text{file} = \text{“filename.nc”}, \\
\quad \text{it} = c(1980, 2010), \text{lon}=c(0, 15), \text{lat}=c(60, 70))
\]

- **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
...

Meteorologisk institutt 150 år
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

\[ y <- \text{subset}(x, \]
\[ \quad \text{it="djf"}, \]
\[ \quad \text{is=list(lon=c(-30,30),lat=c(50,75))}, \]
\[ \quad \text{ip=1,} \]
\[ \quad \text{im=9}) \]

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

\text{is} \quad \text{spatial range, list of the longitude and latitude range}

\text{ip} \quad \text{index for pattern and pca selection}

\text{im} \quad \text{index for climate model selection}
aggregate

Aggregate in space

\[ y \leftarrow \text{aggregate.area}(x, \text{FUN}='\text{mean}') \]

Aggregate in time

seasonal cycle:

\[ y \leftarrow \text{aggregate}(x, \text{by} = \text{month}(x), \text{FUN}='\text{mean}') \]

annual mean:

\[ y \leftarrow \text{aggregate}(x, \text{by} = \text{year}(x), \text{FUN}='\text{mean}') \]

\[ y \leftarrow \text{annual}(x, \text{FUN}='\text{mean}') \]

seasonal and monthly mean:

\[ y \leftarrow \text{as.4seasons}(x, \text{FUN}='\text{mean}') \]

\[ y \leftarrow \text{as.monthly}(x, \text{FUN}='\text{mean}') \]

The argument \text{FUN} can be, e.g., ‘\text{mean}’, ‘\text{sd}’, ‘\text{sum}’, ‘\text{trend}’, ‘\text{wetfreq}’, ‘\text{wetmean}’, ‘\text{count}’, ‘\text{spell}’, ‘\text{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

\[ X \leftarrow \text{EOF}(x, n=20) \]

\[ \text{plot}(X) \]

\( x \) a ‘field’ object

\( n \) number of EOFs

spatial patterns

amplitudes

time series

NCEP reanalysis, spring (MAM)
What is PCA?

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

Original data

Principal components (PCs)
**EOF**  
empirical orthogonal functions

\[ X \leftarrow \text{EOF}(x, \text{n}=20) \]

\[ \text{plot}(X) \]

\( x \)  
a ‘field’ object  
\( n \)  
number of EOFs

**Spatial patterns**

**Amplitudes**

**Time series**

NCEP reanalysis, spring (MAM)
3. ESD for visualising, producing appropriate infographics

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

empirical orthogonal functions

\[ X \leftarrow \text{EOF}(x, \ n=20) \]

\[ \text{plot}(X) \]

- **x**: a ‘field’ object
- **n**: number of EOFs

**spatial patterns**

**amplitudes**

**time series**

NCEP reanalysis, spring (MAM)
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!

1. Training framework
   - Scenario
   - Global
   - Climate
   - Weather
   - Reanalysis and Historical
   - GCM
   - ESD
   - ESD
   - Weather Generator
   - Observed Local Climate (Stations)

2. Simulation framework
   - Few RCPs (~4)
   - All GCMs (~Hundred)
   - ESD Seasonal Statistics of daily values
   - Weather stations (~thousands)
   - Hydrological Simulations

ESD can be used in different parts of the simulation framework to improve accuracy and resolution.
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.

d$s <-$ 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.

d$s <- DS(y, X, eofs=1:5)$

plot(ds)
DS empirical-statistical downscaling

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

d$s \leftarrow \text{DS}(y, X, \text{eofs}=1:5)$

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

\[ t = [t_A, t_B] = 1, \ldots, N \]
Two (or more) field objects are combined before EOF analysis is applied. This is done to identify **common spatial patterns** in the data sets.

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

\[
Z \leftarrow \text{combine}(X, Y)
\]
\[
eof.Z \leftarrow \text{EOF}(Z)
\]

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

```r
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
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)
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/](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
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 the seasonal cycle in mean air temperature. The following graphs show an example of this comparison.

1. **Select** your data source from the climatedatasite.net.
2. **Settings & Outputs** choose the type of climatic simulation you want to evaluate.
3. **Evaluate the seasonal cycle in simulated Mean Air Temperature**
   - 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 output from the “Settings & Outputs” tab box into, for example, individual simulations, or the ensemble model simulations. Box plots of both, transform the values into anomalies, group the models by attributes, etc. The user can also double-click on specific climate models from the legend (once displayed) or the metadata table to isolate one or a group of simulations or modify the displayed 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 table for further details about the simulations.

4. **Evaluate the seasonal cycle in Simulated Monthly Precipitation totals**
   - The interactive figure shows the seasonal cycle of pseudo-observed (dashed) and modeled monthly precipitation by the multi-model ensemble of simulations assuming the intermediate emission scenario (RCP4.5). The user can modify the type of output from the “Settings & Outputs” tab box into, for example, individual simulations, or the ensemble model simulations. Box plots of both, transform the values into anomalies, group the models by attributes, etc. The user can also double-click on specific climate models from the legend (once displayed) or the metadata table to isolate one or a group of simulations or modify the displayed 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 table for further details about the simulations.
5. Scatter Plots of Simulated Climate Variables

![Scatter plot showing changes in annual mean regional precipitation values versus changes in annual mean regional temperature values.](https://climatedatasite.net)

- CSIRO-BOM ACCESS1.0 r1i1p1 1 1...
- CSIRO-BOM ACCESS1.3 r1i1p1 1 2...
- BCC bcc-csm1-1 r1i1p1 1 3...
- BCC bcc-csm1-1-m r1i1p1 1 4...
- BNU BNU-ESM r1i1p1 1 5...
- CCCma CanESM2 r1i1p1 1 6...
- CCCma CanESM2 r2i1p1 2 7...
- CCCma CanESM2 r3i1p1 3 8...
- CCCma CanESM2 r4i1p1 4 9...
- CCCma CanESM2 r5i1p1 5 10...
- NCAR CCSM4 r1i1p1 1 11...
- NCAR CCSM4 r2i1p1 2 12...
- NCAR CCSM4 r3i1p1 3 13...
- NCAR CCSM4 r4i1p1 4 14...
- NCAR CCSM4 r5i1p1 5 15...
- NCAR CCSM4 r6i1p1 6 16...
- NSF-DOE-NCAR CESM1-BGC r1i1p1 1 17...
- CMCC CMCC-CM N/A 1 18...
- CHES CMCC-CM N/A 1 19...
Climate Mapping Toolkit
Climate Projections Toolkit for the Nordic Countries

Projected changes for the Nordic countries

+ 2.6 C
Averaged change in Temperature

Intermediate
Emission Scenario (RCP4.5)

108
Global Climate Models

Disclaimer

Explore Map

ENSEMBLE MEAN of projected changes in mean annual (all seasons) temperature by 2070–2099 with regards to the base period 1980–2010 assuming Intermediate emissions (RCP4.5) scenarios.
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