IsoriX

an R package for isoscape computation and inference of spatial origins using mixed models

The IsoriX core team

September 2022

Table of contents

Table of contents

Table of contents

Why IsoriX?

Benefits:

- fully reproducible methods & results

Why IsoriX?

Benefits:

- fully reproducible methods & results
- as simple as necessary (but not beyond that)

Why IsoriX?

Benefits:

- fully reproducible methods & results
- as simple as necessary (but not beyond that)
- each step of the analysis can be thoroughly studied

Why IsoriX?

Benefits:

- fully reproducible methods & results
- as simple as necessary (but not beyond that)
- each step of the analysis can be thoroughly studied
- cutting edge statistical methods accounting for many sources of uncertainty

Benefits:

- fully reproducible methods & results
- as simple as necessary (but not beyond that)
- each step of the analysis can be thoroughly studied
- cutting edge statistical methods accounting for many sources of uncertainty
- compatible with Geographic Information System within and without **R**

Benefits:

- fully reproducible methods & results
- as simple as necessary (but not beyond that)
- each step of the analysis can be thoroughly studied
- cutting edge statistical methods accounting for many sources of uncertainty
- compatible with Geographic Information System within and without **R**
- works on Windows, MacOS, Linux, Unix; localy or remotely; with or without Internet

Why IsoriX?

Benefits:

- fully reproducible methods & results
- as simple as necessary (but not beyond that)
- each step of the analysis can be thoroughly studied
- cutting edge statistical methods accounting for many sources of uncertainty
- compatible with Geographic Information System within and without **R**
- works on Windows, MacOS, Linux, Unix; localy or remotely; with or without Internet
- free & open source (anyone can use and improve IsoriX!)

Why IsoriX?

Benefits:

- fully reproducible methods & results
- as simple as necessary (but not beyond that)
- each step of the analysis can be thoroughly studied
- cutting edge statistical methods accounting for many sources of uncertainty
- compatible with Geographic Information System within and without **R**
- works on Windows, MacOS, Linux, Unix; localy or remotely; with or without Internet
- free & open source (anyone can use and improve IsoriX!)

Limits:

- **R** knowledge required
- a little stable isotope knowledge also required
- not multivariate (1 isotope only, but isoscapes or assignment maps can be combined)
- not Bayesian (prior information not considered)

Table of contents

An **R** package:

```
install.packages("IsoriX")  ## to install
```

```
library(IsoriX)  ## to load
```

```
## Registered S3 methods overwritten by 'registry':
##   method               from
##   print.registry_field proxy
##   print.registry_entry proxy
##
##  IsoriX version 0.8.3 is now loaded
##
##  Type:
##    * `?IsoriX` for a very short description
##    * `browseURL('https://bookdown.org/content/782/')` for a longer (online) documentation
##    * `help(package = 'IsoriX', help_type = 'html')` for a list of the package objects and help files
##    * `citation('IsoriX')` for how to cite IsoriX (i.e.  the papers you should read)
##    * `news(package = 'IsoriX')` for info on changed between versions of IsoriX
##
##  Please join the mailing list 'https://groups.google.com/g/IsoriX'
##  for help, news and discussions about IsoriX
```

An interface between **R** packages (without help files: only ca. 1000 lines of codes):

```
tools::package_dependencies("IsoriX") ## note: list dependencies, not suggested packages
```

```
## $IsoriX
##  [1] "elevatr"      "graphics"     "grDevices"    "grid"         "lattice"      "latticeExtra" "numDeriv"
##  [8] "raster"       "rasterVis"    "sp"           "spaMM"        "stats"        "tools"        "utils"
## [15] "viridisLite"
```

- elevatr (for data on elevation/bathymetry)
- numDeriv, **spaMM**, stats (for statistical computation)
- **raster**, sp (for GIS)
- graphics, grDevices, grid, lattice, latticeExtra, **rasterVis**, viridisLite (for plotting)
- tools, utils (for small geeky details)

What is IsoriX?

A collection of functions:

```
setdiff(ls("package:IsoriX"), data(package = "IsoriX")$results[, "Item"])  ## note: does not show methods and unexported functions
```

```
##  [1] "area"             "calibfit"         "create_aliens"    "CRS"              "downloadfile"     "extent"
##  [7] "extent<-"         "extract"          "get_ranPars"      "getelev"          "getOption_IsoriX" "getprecip"
## [13] "gpar"             "grid.text"        "isofind"          "isofit"           "isomultifit"      "isomultiscape"
## [19] "isoscape"         "layer"            "levelplot"        "options_IsoriX"   "panel.points"     "prepcipitate"
## [25] "prepraster"       "prepsources"      "projection<-"     "raster"           "RdBuTheme"        "shift"
## [31] "sp.points"        "sp.polygons"      "values"           "xyplot"
```

What is IsoriX?

A collection of functions:

```
setdiff(ls("package:IsoriX"), data(package = "IsoriX")$results[, "Item"])  ## note: does not show methods and unexported functions
```

```
##  [1] "area"             "calibfit"         "create_aliens"    "CRS"              "downloadfile"     "extent"
##  [7] "extent<-"         "extract"          "get_ranPars"      "getelev"          "getOption_IsoriX" "getprecip"
## [13] "gpar"             "grid.text"        "isofind"          "isofit"           "isomultifit"      "isomultiscape"
## [19] "isoscape"         "layer"            "levelplot"        "options_IsoriX"   "panel.points"     "prepcipitate"
## [25] "prepraster"       "prepsources"      "projection<-"     "raster"           "RdBuTheme"        "shift"
## [31] "sp.points"        "sp.polygons"      "values"           "xyplot"
```

The main ones being:

1. `prepsources()`
2. `isofit()`
3. `getelev()`
4. `prepraster()`
5. `isoscape()`
6. `calibfit()`
7. `isofind()`

All these functions (as well as the methods behind `plot()`) are documented!
So you can type e.g. `?prepsources()` to get help.

What is IsoriX?

A collection of data to try things out or to make plots nicer:

```
data(package = "IsoriX")$results[, c("Item", "Title")]
```

```
##       Item              Title
##  [1,] "AssignDataAlien" "Simulated assignment dataset"
##  [2,] "AssignDataBat"   "Assignment dataset for bat species"
##  [3,] "AssignDataBat2"  "Assignment dataset for bat species"
##  [4,] "CalibDataAlien"  "Simulated calibration dataset"
##  [5,] "CalibDataBat"    "Calibration dataset for bat species"
##  [6,] "CalibDataBat2"   "Calibration dataset for bat species"
##  [7,] "CountryBorders"  "Borders of world CountryBorders"
##  [8,] "ElevRasterDE"    "The raster of elevation for Germany"
##  [9,] "GNIPDataALLagg"  "Hydrogen delta values in precipitation water (aggregated per location)"
## [10,] "GNIPDataDE"      "Hydrogen delta values in precipitation water, Germany"
## [11,] "GNIPDataEUagg"   "Hydrogen delta values in precipitation water (aggregated per location)"
## [12,] "OceanMask"       "Mask of world oceans"
## [13,] "PrecipBrickDE"   "The precipitation monthly amounts for Germany"
## [14,] "isopalette1"     "Colour palettes for plotting"
## [15,] "isopalette2"     "Colour palettes for plotting"
```

What is IsoriX?

A collection of data to try things out or to make plots nicer:

```
data(package = "IsoriX")$results[, c("Item", "Title")]
```

```
##       Item              Title
##  [1,] "AssignDataAlien" "Simulated assignment dataset"
##  [2,] "AssignDataBat"   "Assignment dataset for bat species"
##  [3,] "AssignDataBat2"  "Assignment dataset for bat species"
##  [4,] "CalibDataAlien"  "Simulated calibration dataset"
##  [5,] "CalibDataBat"    "Calibration dataset for bat species"
##  [6,] "CalibDataBat2"   "Calibration dataset for bat species"
##  [7,] "CountryBorders"  "Borders of world CountryBorders"
##  [8,] "ElevRasterDE"    "The raster of elevation for Germany"
##  [9,] "GNIPDataALLagg"  "Hydrogen delta values in precipitation water (aggregated per location)"
## [10,] "GNIPDataDE"      "Hydrogen delta values in precipitation water, Germany"
## [11,] "GNIPDataEUagg"   "Hydrogen delta values in precipitation water (aggregated per location)"
## [12,] "OceanMask"       "Mask of world oceans"
## [13,] "PrecipBrickDE"   "The precipitation monthly amounts for Germany"
## [14,] "isopalette1"     "Colour palettes for plotting"
## [15,] "isopalette2"     "Colour palettes for plotting"
```

And of course you can use your own data, that is the whole point!

A project that aims at becoming a collaborative platform:

`www.github.com/courtiol/IsoriX`

Table of contents

Where can you learn about IsoriX?

- There is a bookdown

  `https://bookdown.org/content/782/`

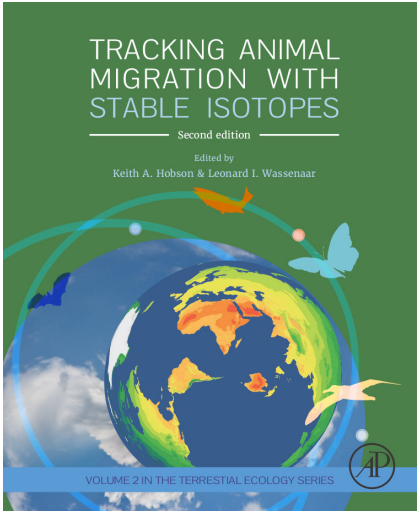
- There is the package documentation

```
help(package = 'IsoriX')
```

- There is a mailing list

  `https://groups.google.com/g/IsoriX`

- There is a book chapter
  (in the 2nd edition of the book)



Is IsoriX better than alternatives?

- All methods differ in some details
  (e.g. IsoriX does smoothing, not interpolation)
- All methods differ in their set of assumptions
  (some uncertainties are considered, some are discarded)

Is IsoriX better than alternatives?

- All methods differ in some details
  (e.g. IsoriX does smoothing, not interpolation)
- All methods differ in their set of assumptions
  (some uncertainties are considered, some are discarded)

Conclusion:

- A single method is unlikely to be best across all situations
- Run several methods on your data!

Is IsoriX better than alternatives?

- All methods differ in some details
  (e.g. IsoriX does smoothing, not interpolation)
- All methods differ in their set of assumptions
  (some uncertainties are considered, some are discarded)

Conclusion:

- A single method is unlikely to be best across all situations
- Run several methods on your data!

What to do?

- if all methods give you the same answer, you have a robust result $\rightarrow$ great!
- if they don't, try to figure out which assumption(s) introduce the difference and figure out in which respect your data contradict the identified assumption(s)
  $\rightarrow$ you will better understand the biology/geology/... in your system $\rightarrow$ great!

Table of contents

The standard workflow for an assignment involves 6 main steps and thus 6 functions

1. Prepare the source data with `prepsources()`
2. Fit the Isoscape model with `isofit()`
3. Prepare the structural raster with `prepraster()`
4. Predict the isoscape with `isoscape()`
5. Fit the calibration function with `calibfit()`
6. Perform the assignment with `isofind()`

A complete assignment can be done in 8 lines of code proof , but it is important to understand well and check the details of each step to get the best out of your data. So we will add a few extra steps.

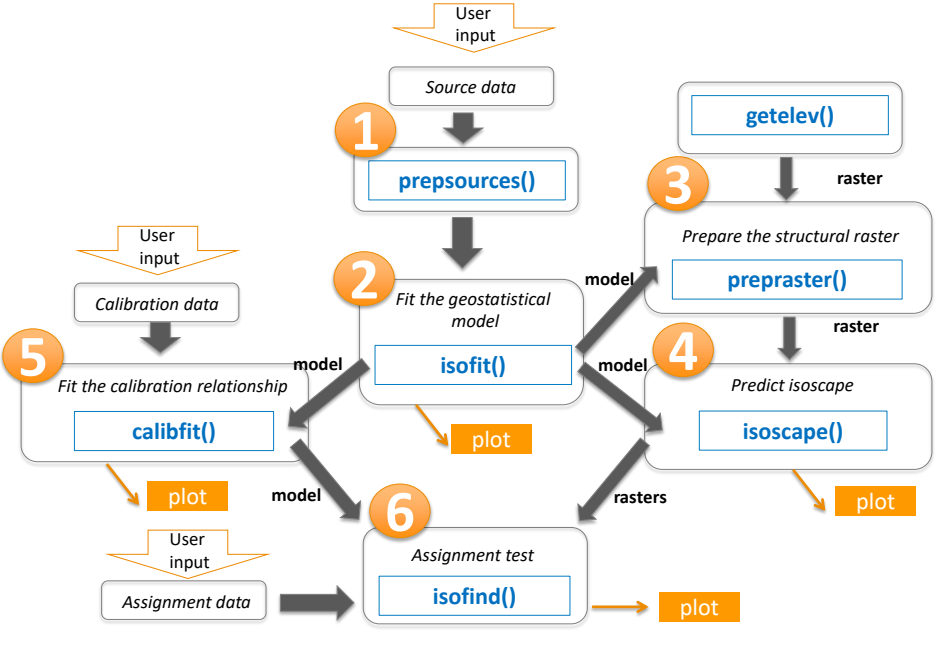
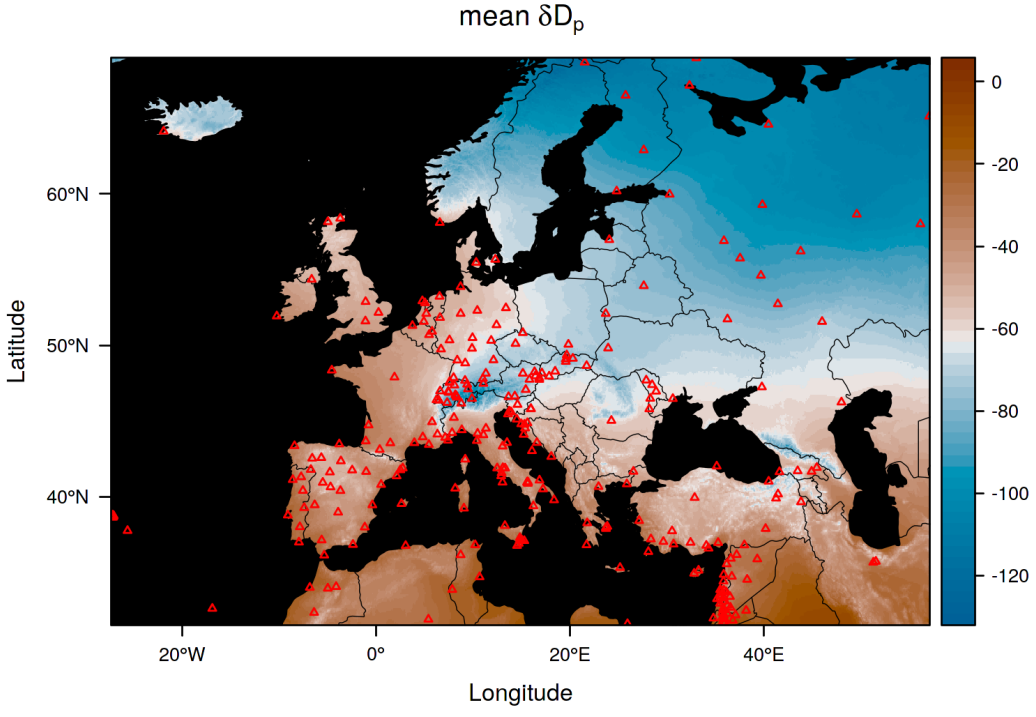

User input
Source data
1
prepsources()
getelev()
raster
3
Prepare the structural raster
prepraster()
raster
User input
Calibration data
2
Fit the geostatistical model
isofit()
model
model
5
Fit the calibration relationship
calibfit()
model
model
4
Predict isoscape
isoscape()
plot
plot
plot
model
6
rasters
User input
Assignment data
Assignment test
isofind()
plot

Table of contents


mean δD$_p$
60°N
50°N
40°N
Latitude
20°W
0°
20°E
40°E
Longitude
0
-20
-40
-60
-80
-100
-120

Step 1. Prepare the source data

1a. Retrieve (e.g. from GNIP)/Collect data and shape them as the toy dataset `GNIPDataDE`:

```
head(GNIPDataDE)
```

```
##   source_ID   lat long elev year month source_value
## 1 SCHLESWIG 54.52 9.54   43 1997     6        -59.0
## 2 SCHLESWIG 54.52 9.54   43 1997     7        -56.0
## 3 SCHLESWIG 54.52 9.54   43 1997     8        -60.8
## 4 SCHLESWIG 54.52 9.54   43 1997     9        -51.0
## 5 SCHLESWIG 54.52 9.54   43 1997    10        -58.7
## 6 SCHLESWIG 54.52 9.54   43 1997    11        -74.6
```

```
tail(GNIPDataDE)
```

```
##                   source_ID   lat  long elev year month source_value
## 8586 GARMISCH-PARTENKIRCHEN 47.48 11.06  719 2013     7       -46.42
## 8587 GARMISCH-PARTENKIRCHEN 47.48 11.06  719 2013     8       -48.07
## 8588 GARMISCH-PARTENKIRCHEN 47.48 11.06  719 2013     9       -63.75
## 8589 GARMISCH-PARTENKIRCHEN 47.48 11.06  719 2013    10      -100.23
## 8590 GARMISCH-PARTENKIRCHEN 47.48 11.06  719 2013    11      -109.23
## 8591 GARMISCH-PARTENKIRCHEN 47.48 11.06  719 2013    12      -105.82
```

Step 1. Prepare the source data

1b. Use `prepsources()` to aggregate (and filter) the data:

```
prepsources(
  data,
  month = 1:12, year,
  long_min, long_max, lat_min, lat_max,
  split_by = NULL,
  prop_random = 0, random_level = "source",
  col_source_value = "source_value",
  col_source_ID = "source_ID", col_lat = "lat", col_long = "long",
  col_elev = "elev", col_month = "month", col_year = "year"
  )
```

1b. Use `prepsources()` to aggregate (and filter) the data:

```
prepsources(
  data,
  month = 1:12, year,
  long_min, long_max, lat_min, lat_max,
  split_by = NULL,
  prop_random = 0, random_level = "source",
  col_source_value = "source_value",
  col_source_ID = "source_ID", col_lat = "lat", col_long = "long",
  col_elev = "elev", col_month = "month", col_year = "year"
  )
```

Example:

```
GNIPDataDEagg <- prepsources(data = GNIPDataDE)
```

```
head(GNIPDataDEagg)
```

```
##       source_ID mean_source_value var_source_value n_source_value   lat  long elev
## 1        ARKONA         -60.99231         247.8767            134 54.67 13.43   42
## 2        ARTERN         -61.00653         510.5720            199 51.37 11.29  164
## 3 BAD SALZUFLEN         -53.80770         310.1046            408 52.10  8.75  135
## 4        BERLIN         -57.53477         395.6484            419 52.46 13.40   48
## 5  BRAUNSCHWEIG         -52.73350         339.4752            420 52.29 10.44   81
## 6      CUXHAVEN         -49.25271         221.6594            420 53.87  8.70    5
```

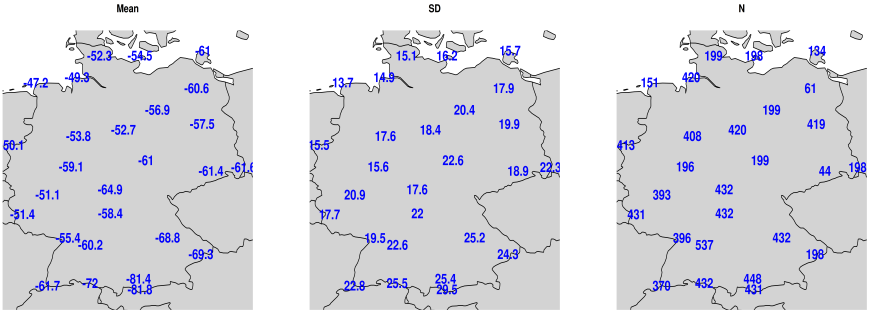
1c. Check your data (it is always a good idea to plot your data!)

```
library(sp) ## for plotting polygons

plot(CountryBorders, xlim = range(GNIPDataDEagg$long), ylim = range(GNIPDataDEagg$lat), col = "lightgrey", main = "Mean")
text(x = GNIPDataDEagg$long, y = GNIPDataDEagg$lat, labels = round(GNIPDataDEagg$mean_source_value, digits = 1), col = "blue", cex = 1.5, font = 2)

plot(CountryBorders, xlim = range(GNIPDataDEagg$long), ylim = range(GNIPDataDEagg$lat), col = "lightgrey", main = "SD")
text(x = GNIPDataDEagg$long, y = GNIPDataDEagg$lat, labels = round(sqrt(GNIPDataDEagg$var_source_value), digits = 1), col = "blue", cex = 1.5, font = 2)

plot(CountryBorders, xlim = range(GNIPDataDEagg$long), ylim = range(GNIPDataDEagg$lat), col = "lightgrey", main = "N")
text(x = GNIPDataDEagg$long, y = GNIPDataDEagg$lat, labels = GNIPDataDEagg$n_source_value, col = "blue", cex = 1.5, font = 2)
```



Step 1. Prepare the source data

The most important questions you must answer:

- Which years?
- Which months?
- Which geographic area?
- Weighted by precipitation or not? If yes, you have 2 possibilities, before or after fitting:
  - before: do the aggregation yourself (for now)
    NB: don't forget to rise the weights to the square when averaging variances!
  - after: using `split_by` and later on `isomultifit` and `isomultiscape` (but not ready yet for calibration)

Step 1. Prepare the source data

The most important questions you must answer:

- Which years?
- Which months?
- Which geographic area?
- Weighted by precipitation or not? If yes, you have 2 possibilities, before or after fitting:
  - before: do the aggregation yourself (for now)
    NB: don't forget to rise the weights to the square when averaging variances!
  - after: using `split_by` and later on `isomultifit` and `isomultiscape` (but not ready yet for calibration)

Issue:

- More data = more signal or more noise?

Step 1. Prepare the source data

The most important questions you must answer:

- Which years?
- Which months?
- Which geographic area?
- Weighted by precipitation or not? If yes, you have 2 possibilities, before or after fitting:
  - before: do the aggregation yourself (for now)
    NB: don't forget to rise the weights to the square when averaging variances!
  - after: using `split_by` and later on `isomultifit` and `isomultiscape` (but not ready yet for calibration)

Issue:

- More data = more signal or more noise?

Trade-off:

- The isotopic signature in an environment at time t is not solely influenced by the precipitation at time t.
- Too small area leads to extrapolation which is unreliable but too large leads to impose a single model over several regions that may differ in how the environment relates to isotopes.

Step 1. Prepare the source data

The most important questions you must answer:

- Which years?
- Which months?
- Which geographic area?
- Weighted by precipitation or not? If yes, you have 2 possibilities, before or after fitting:
  - before: do the aggregation yourself (for now)
    NB: don't forget to rise the weights to the square when averaging variances!
  - after: using `split_by` and later on `isomultifit` and `isomultiscape` (but not ready yet for calibration)

Issue:

- More data = more signal or more noise?

Trade-off:

- The isotopic signature in an environment at time t is not solely influenced by the precipitation at time t.
- Too small area leads to extrapolation which is unreliable but too large leads to impose a single model over several regions that may differ in how the environment relates to isotopes.

Note:

- In IsoriX it is easy to change a given line of code and re-run your entire workflow to study the impact of your choices.

Step 1. Prepare the source data

Uncertainty check list:

- ✓ variation across space
- ✓ variation across years and months
- × variation within months (includes technical variation)

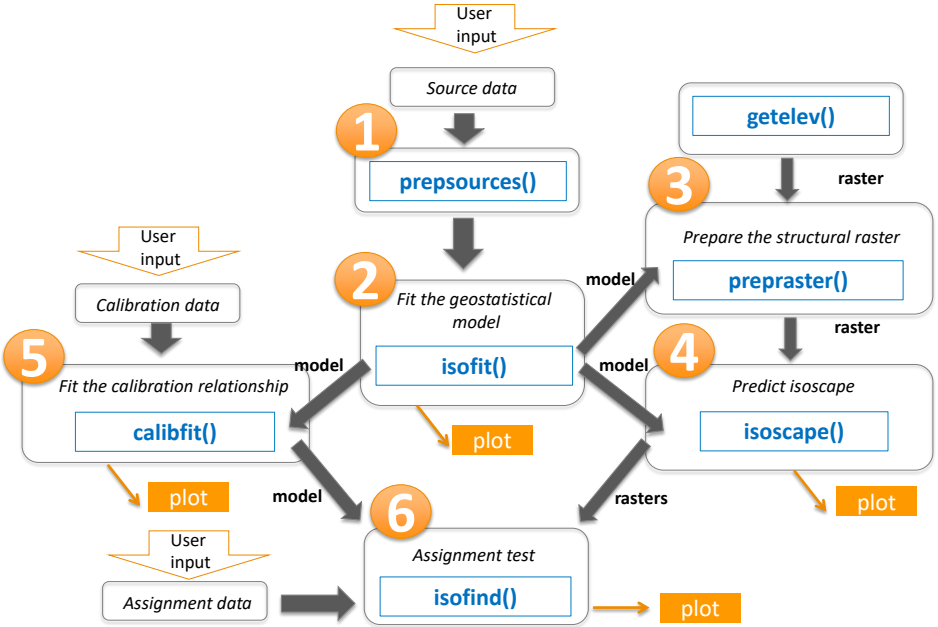
Uncertainty check list:

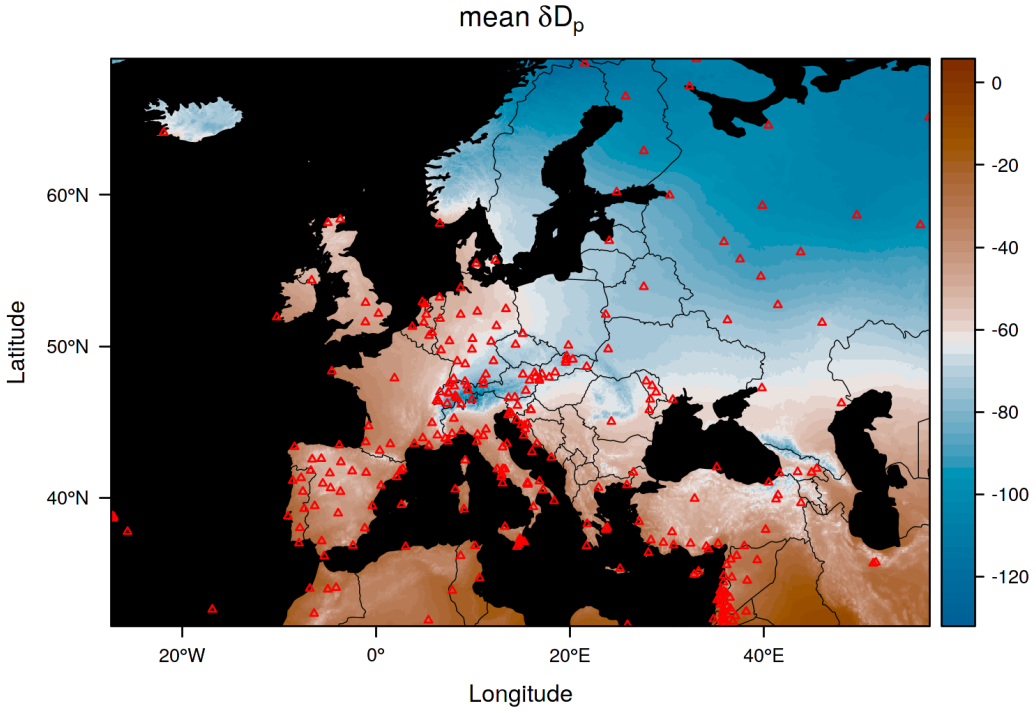
- ✓ variation across space
- ✓ variation across years and months
- × variation within months (includes technical variation)

Remember:

- Always try to be explicit about which sources of uncertainty are accounted for and which are not!
- No work can consider them all, so it is worth recognising and discussing the limitations of any approach.
- Being open source, in IsoriX, with a little effort you can figure out which assumptions are being made.

Table of contents


User input
Source data
1
prepsources()
getelev()
raster
3
Prepare the structural raster
prepraster()
raster
User input
Calibration data
2
Fit the geostatistical model
isofit()
model
model
5
Fit the calibration relationship
calibfit()
model
plot
4
Predict isoscape
isoscape()
plot
plot
model
6
rasters
User input
Assignment data
Assignment test
isofind()
plot


mean δD_p
60°N
50°N
40°N
Latitude
20°W
0°
20°E
40°E
Longitude
0
-20
-40
-60
-80
-100
-120

The statistical model(s) behind the isoscape:

- Mean model (LMM):

$$\texttt{Source}_{gi} = \texttt{Fix}_{g} + \texttt{Random}_{g} + \texttt{Error}_{gi}$$

The statistical model(s) behind the isoscape:

- Mean model (LMM):

$$\texttt{Source}_{gi} = \texttt{Fix}_{g} + \texttt{Random}_{g} + \texttt{Error}_{gi}$$

- where

$$\texttt{Fix}_{g} = \beta_0 + \beta_{\texttt{lat}} \times |\texttt{lat}|_{g} + \beta_{\texttt{elev}} \times \texttt{elev}_{g}$$

$$\texttt{Random}_{g} = \texttt{RandomSpatial}_{g} + \texttt{RandomUncorr}_{g}$$

and

$$\texttt{Correlation}(\texttt{RandomSpatial}_{a}, \texttt{RandomSpatial}_{b}) = \texttt{Matern}(d_{ab}, \rho, \nu)$$

The statistical model(s) behind the isoscape:

- Mean model (LMM):

$$\texttt{Source}_{gi} = \texttt{Fix}_g + \texttt{Random}_g + \texttt{Error}_{gi}$$

- where

$$\texttt{Fix}_g = \beta_0 + \beta_{\texttt{lat}} \times |\texttt{lat}|_g + \beta_{\texttt{elev}} \times \texttt{elev}_g$$

$$\texttt{Random}_g = \texttt{RandomSpatial}_g + \texttt{RandomUncorr}_g$$

  and

$$\texttt{Correlation}(\texttt{RandomSpatial}_a, \texttt{RandomSpatial}_b) = \texttt{Matern}(d_{ab}, \rho, \nu)$$

- Residual dispersion model (disp model; Γ GLMM):

$$\mathrm{E}\left(\texttt{Error}_{gi}^2\right) = \exp\left(\texttt{Fix}_g^{\mathrm{D}} + \texttt{RandomSpatial}_g^{\mathrm{D}} + \texttt{RandomUncorr}_g^{\mathrm{D}}\right)$$

The statistical model(s) behind the isoscape:

- Mean model (LMM):

$$\texttt{Source}_{gi} = \texttt{Fix}_g + \texttt{Random}_g + \texttt{Error}_{gi}$$

- where

$$\texttt{Fix}_g = \beta_0 + \beta_{\texttt{lat}} \times |\texttt{lat}|_g + \beta_{\texttt{elev}} \times \texttt{elev}_g$$

$$\texttt{Random}_g = \texttt{RandomSpatial}_g + \texttt{RandomUncorr}_g$$

  and

$$\texttt{Correlation}(\texttt{RandomSpatial}_a, \texttt{RandomSpatial}_b) = \texttt{Matern}(d_{ab}, \rho, \nu)$$

- Residual dispersion model (disp model; Γ GLMM):

$$\mathrm{E}\left(\texttt{Error}_{gi}^2\right) = \exp\left(\texttt{Fix}_g^{\mathrm{D}} + \texttt{RandomSpatial}_g^{\mathrm{D}} + \texttt{RandomUncorr}_g^{\mathrm{D}}\right)$$

Step 2. Fit the isoscape

The statistical model(s) behind the isoscape:

- Mean model (LMM):

$$\texttt{Source}_{gi} = \texttt{Fix}_g + \texttt{Random}_g + \texttt{Error}_{gi}$$

- where

$$\texttt{Fix}_g = \beta_0 + \beta_{\texttt{lat}} \times |\texttt{lat}|_g + \beta_{\texttt{elev}} \times \texttt{elev}_g$$

$$\texttt{Random}_g = \texttt{RandomSpatial}_g + \texttt{RandomUncorr}_g$$

  and

$$\texttt{Correlation}(\texttt{RandomSpatial}_a, \texttt{RandomSpatial}_b) = \texttt{Matern}(d_{ab}, \rho, \nu)$$

- Residual dispersion model (disp model; Γ GLMM):

$$\mathrm{E}\left(\texttt{Error}_{gi}^2\right) = \exp\left(\texttt{Fix}_g^{\mathrm{D}} + \texttt{RandomSpatial}_g^{\mathrm{D}} + \texttt{RandomUncorr}_g^{\mathrm{D}}\right)$$

Note:
d_{ab} – the distance between 2 locations – is (by default) measured accounting for the curvature of the Earth.

Step 2. Fit the isoscape

2a. Use `isofit()` to fit the geostatistical model(s):

```
isofit(
  data,
  mean_model_fix = list(elev = FALSE, lat_abs = FALSE, lat_2 = FALSE, long = FALSE, long_2 = FALSE),
  disp_model_fix = list(elev = FALSE, lat_abs = FALSE, lat_2 = FALSE, long = FALSE, long_2 = FALSE),
  mean_model_rand = list(uncorr = TRUE, spatial = TRUE),
  disp_model_rand = list(uncorr = TRUE, spatial = TRUE),
  uncorr_terms = list(mean_model = "lambda", disp_model = "lambda"),
  spaMM_method = list(mean_model = "fitme", disp_model = "fitme"),
  dist_method = "Earth",
  control_mean = list(),
  control_disp = list(),
  verbose = interactive()
  )
```

Step 2. Fit the isoscape

2a. Use `isofit()` to fit the geostatistical model(s):

```
isofit(
  data,
  mean_model_fix = list(elev = FALSE, lat_abs = FALSE, lat_2 = FALSE, long = FALSE, long_2 = FALSE),
  disp_model_fix = list(elev = FALSE, lat_abs = FALSE, lat_2 = FALSE, long = FALSE, long_2 = FALSE),
  mean_model_rand = list(uncorr = TRUE, spatial = TRUE),
  disp_model_rand = list(uncorr = TRUE, spatial = TRUE),
  uncorr_terms = list(mean_model = "lambda", disp_model = "lambda"),
  spaMM_method = list(mean_model = "fitme", disp_model = "fitme"),
  dist_method = "Earth",
  control_mean = list(),
  control_disp = list(),
  verbose = interactive()
  )
```

Example:

```
GermanFit <- isofit(data = GNIPDataDEagg,
                    mean_model_fix = list(elev = TRUE, lat_abs = TRUE))
```

2b. Check the fitted geostatistical model(s):

```
names(GermanFit)
```

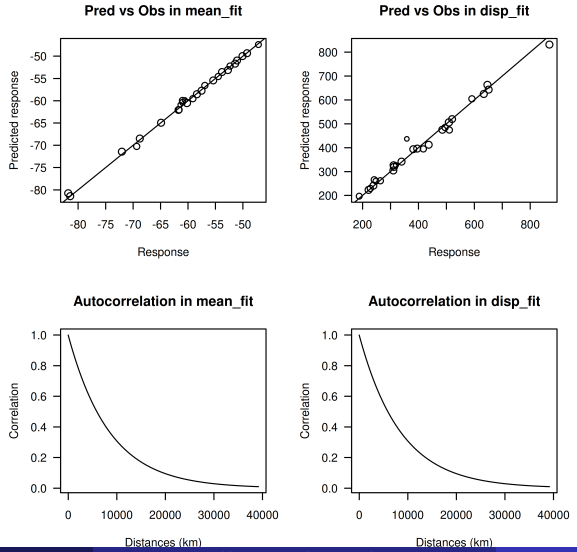
```
## [1] "mean_fit" "disp_fit" "info_fit"
```

```
GermanFit$mean_fit
```

```
## formula: mean_source_value ~ 1 + elev + lat_abs + (1 | source_ID) + Matern(1 |
##     long + lat)
## REML: Estimation of lambda and corrPars by REML.
##       Estimation of fixed effects by ML.
## family: gaussian( link = identity )
##  ------------ Fixed effects (beta) ------------
##               Estimate  Cond. SE t-value
## (Intercept) -145.78874 65.386452  -2.230
## elev          -0.01163  0.002449  -4.750
## lat_abs        1.68318  1.210011   1.391
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
## Distance(s): Earth
##                     --- Correlation parameters:
##        2.nu        2.rho
## 0.5000000000 0.0001176492
##            --- Variance parameters ('lambda'):
## lambda = var(u) for u ~ Gaussian;
##    source_ID  :  1e-08
##    long + lat  :  482
##             --- Coefficients for log(lambda):
##       Group        Term Estimate Cond.SE
##   source_ID (Intercept)   -18.42    7999
##  long + lat (Intercept)    6.178  0.3232
## # of obs: 27; # of groups: source_ID, 27; long + lat, 27
##  -------------- Residual variance  ------------
## Prior weights: 134 199 408 419 420 ...
## phi was fixed [through  "phi" ~ 0 + offset(pred_disp) ] to 260.009 473.857 304.444 396.195 341.633 ...
##  ------------- Likelihood values  -------------
##                          logLik
## logL       (p_v(h)): -71.72788
```

2b. Check the fitted geostatistical model(s):

```
plot(GermanFit)
```



2c. You can compare the predictive power of different fitted model(s):

```
GermanFit2 <- isofit(data = GNIPDataDEagg,
                     mean_model_fix = list(elev = TRUE, lat_abs = FALSE))
```

```
AIC(GermanFit$mean_fit)
```

```
##        marginal AIC:     conditional AIC:      dispersion AIC:        effective df:
##           157.455760           105.958819           151.485989             4.858705
```

```
AIC(GermanFit2$mean_fit)
```

```
##        marginal AIC:     conditional AIC:      dispersion AIC:        effective df:
##           157.452991           106.076810           155.619924             4.733703
```

2c. You can compare the predictive power of different fitted model(s):

```
GermanFit2 <- isofit(data = GNIPDataDEagg,
                     mean_model_fix = list(elev = TRUE, lat_abs = FALSE))
```

```
AIC(GermanFit$mean_fit)
```

```
##        marginal AIC:     conditional AIC:      dispersion AIC:        effective df:
##           157.455760           105.958819           151.485989             4.858705
```

```
AIC(GermanFit2$mean_fit)
```

```
##        marginal AIC:     conditional AIC:      dispersion AIC:        effective df:
##           157.452991           106.076810           155.619924             4.733703
```

Note:

- Use the conditional AIC (cAIC, which is different from AICc!) as it takes into account of the realization of the random effects.

2c. You can compare the predictive power of different fitted model(s):

```
GermanFit2 <- isofit(data = GNIPDataDEagg,
                     mean_model_fix = list(elev = TRUE, lat_abs = FALSE))
```

```
AIC(GermanFit$mean_fit)
```

```
##        marginal AIC:     conditional AIC:      dispersion AIC:        effective df:
##           157.455760           105.958819           151.485989             4.858705
```

```
AIC(GermanFit2$mean_fit)
```

```
##        marginal AIC:     conditional AIC:      dispersion AIC:        effective df:
##           157.452991           106.076810           155.619924             4.733703
```

Note:

- Use the conditional AIC (cAIC, which is different from AICc!) as it takes into account of the realization of the random effects.
- If you compare models with different structures for the residual dispersion model, (for the time being) you must correct the cAIC of the mean model by adding to it twice the number of parameters of the residual dispersion model. Here this latter model has 5 parameters (1 intercept, 2 variances of random effects, 2 correlation parameters for the Matern). [if residual dispersion models are the same then the correction won't change the ranking of the models]

Step 2. Fit the isoscape

The most important questions you must answer:

- Did the models fit properly your data?
- Which model structure?

Step 2. Fit the isoscape

The most important questions you must answer:

- Did the models fit properly your data?
- Which model structure?

Issue:

- More parameters = more signal or more noise?

Step 2. Fit the isoscape

The most important questions you must answer:

- Did the models fit properly your data?
- Which model structure?

Issue:

- More parameters = more signal or more noise?

Note:

- The optimal model depends on your data!
- As before, in IsoriX it is easy to change a given line of code and re-run your entire workflow to study the impact of your choices.

Step 2. Fit the isoscape

Uncertainty check list:

- mean fit
  - ✓ uncertainty in fixed-effect estimates
  - ✓ uncertainty in random-effect estimates
  - × uncertainty in correlation parameters
  - ✓/× residual variation

Step 2. Fit the isoscape

Uncertainty check list:

- mean fit
  - ✓ uncertainty in fixed-effect estimates
  - ✓ uncertainty in random-effect estimates
  - × uncertainty in correlation parameters
  - ✓/× residual variation

- residual dispersion fit
  - ✓ mean residual variation at a given location
  - × uncertainty in this residual variation

Step 2. Fit the isoscape

Uncertainty check list:

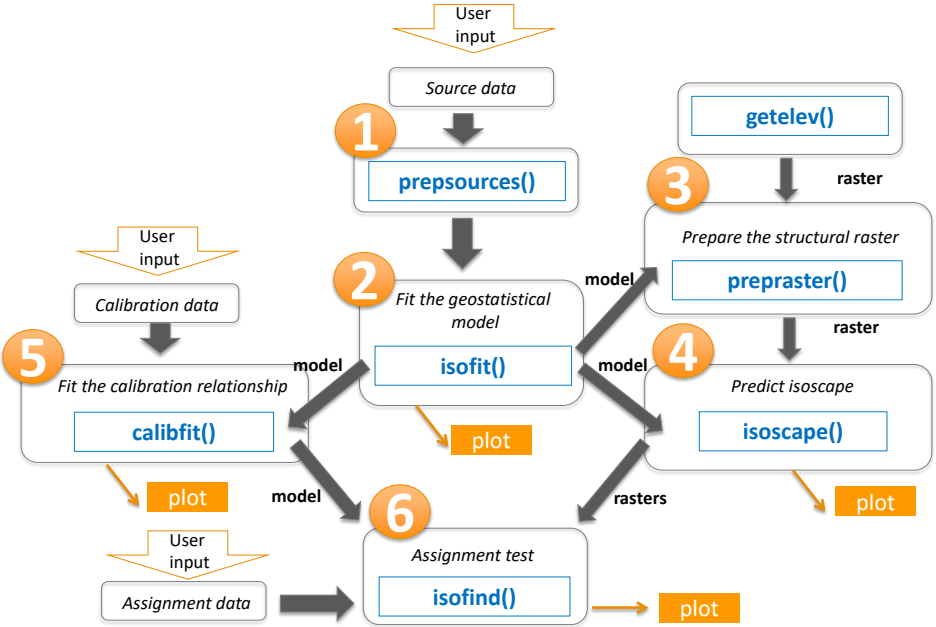
- mean fit
  - ✓ uncertainty in fixed-effect estimates
  - ✓ uncertainty in random-effect estimates
  - × uncertainty in correlation parameters
  - ✓/× residual variation
- residual dispersion fit
  - ✓ mean residual variation at a given location
  - × uncertainty in this residual variation

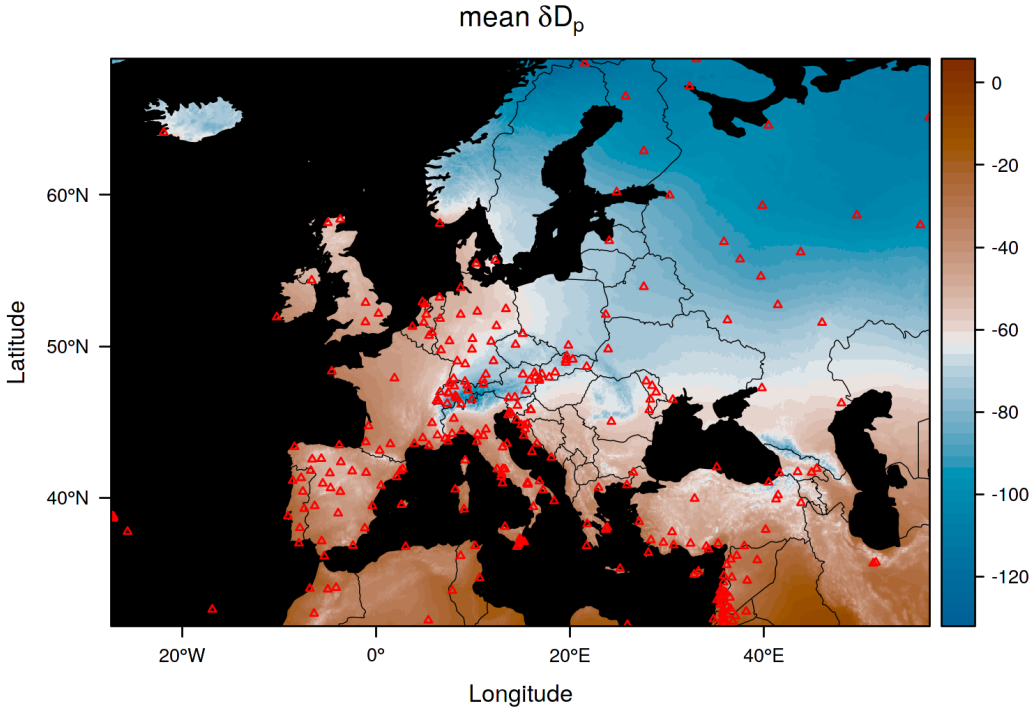
Note:

- Most of these uncertainties are neglected from alternative approaches!
- Some of these uncertainties are huge!!

⏩ Back to the future

Table of contents


User input
Source data
1
prepsources()
getelev()
raster
3
Prepare the structural raster
prepraster()
raster
User input
Calibration data
2
Fit the geostatistical model
isofit()
model
model
5
Fit the calibration relationship
calibfit()
model
plot
4
Predict isoscape
isoscape()
plot
plot
model
6
rasters
User input
Assignment data
Assignment test
isofind()
plot


mean δD_p
60°N
50°N
40°N
Latitude
20°W
0°
20°E
40°E
Longitude
0
-20
-40
-60
-80
-100
-120

Step 3. Prepare the structural raster

3a. Use `getelev()` to prepare the structural raster providing the locations and covariates for predictions (elevation/bathymetry, latitude, longitude):

```
getelev(
  file = "~/elevation_world_z5.tif",
  z = 5,
  long_min = -180, long_max = 180, lat_min = -90, lat_max = 90,
  margin_pct = 5,
  override_size_check = FALSE,
  overwrite = FALSE,
  Ncpu = getOption_IsoriX("Ncpu"),
  verbose = interactive(),
  ...
  )
```

Step 3. Prepare the structural raster

3a. Use `getelev()` to prepare the structural raster providing the locations and covariates for predictions (elevation/bathymetry, latitude, longitude):

```
getelev(
  file = "~/elevation_world_z5.tif",
  z = 5,
  long_min = -180, long_max = 180, lat_min = -90, lat_max = 90,
  margin_pct = 5,
  override_size_check = FALSE,
  overwrite = FALSE,
  Ncpu = getOption_IsoriX("Ncpu"),
  verbose = interactive(),
  ...
  )
```

Example:

```
getelev(file = "elevation_DE_z5.tif",
        long_min = 5.86, long_max = 15.02, lat_min = 47.23, lat_max = 54.90)
```

Note: you can also just use the function without arguments, but that would download the whole world!

```
getelev()
```

3b. Load the structural raster:

```
ElevDE <- raster("elevation_DE_z5.tif")  ## turn the tif into raster
```

```
ElevDE
## class      : RasterLayer
## dimensions : 451, 542, 244442  (nrow, ncol, ncell)
## resolution : 0.01859565, 0.01859565  (x, y)
## extent     : 5.392739, 15.47158, 46.90645, 55.29309  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : elevation_DE_z5.tif
## names      : elevation_DE_z5
## values     : -92, 3158  (min, max)
```

The reasons why `getelev()` does not load the raster is that we want to discourage multiple downloads of the same file and we also want people to be able to use their own structural rasters (i.e. by-passing `getelev()`).

3c. Use `prepraster()` to remove values below (or above) the surface, aggregate and resize the structural raster:

```
prepraster(
  raster,
  isofit = NULL,
  margin_pct = 5,
  aggregation_factor = 0L,
  aggregation_fn = mean,
  manual_crop = NULL,
  values_to_zero = c(-Inf, 0),
  verbose = interactive()
  )
```

3c. Use `prepraster()` to remove values below (or above) the surface, aggregate and resize the structural raster:

```
prepraster(
  raster,
  isofit = NULL,
  margin_pct = 5,
  aggregation_factor = 0L,
  aggregation_fn = mean,
  manual_crop = NULL,
  values_to_zero = c(-Inf, 0),
  verbose = interactive()
  )
```

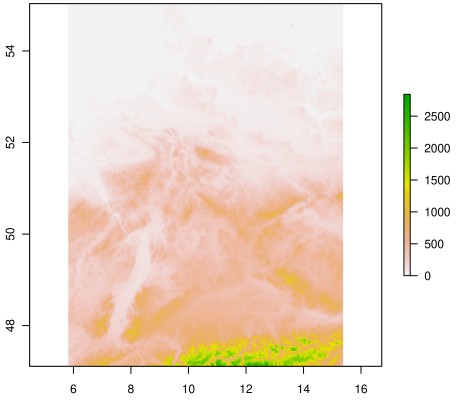
Example:

```
ElevRasterDE <- prepraster(ElevDE, isofit = GermanFit, aggregation_factor = 2)
```

```
## class      : RasterLayer
## dimensions : 213, 257, 54741  (nrow, ncol, ncell)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      : layer
## values     : 0, 2843.5  (min, max)
```

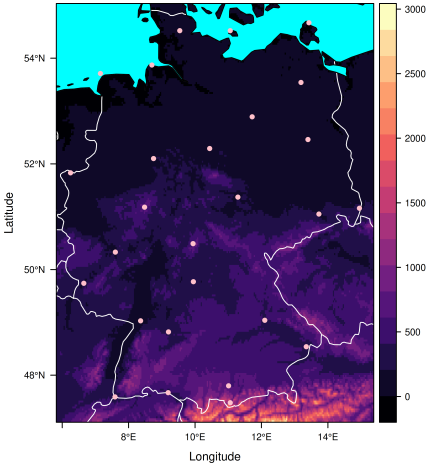
3d. Check the obtained the structural raster:

```
plot(ElevRasterDE)
```



3c. Check the obtained the structural raster:

```
levelplot(ElevRasterDE, margin = FALSE) +
  layer(sp.polygons(CountryBorders, col = "white")) +
  layer(sp.polygons(OceanMask, fill = "cyan")) +
  xyplot(GNIPDataDEagg$lat ~ GNIPDataDEagg$long, col = "pink")
```



Step 3. Prepare the structural raster

The most important questions you must answer:

- Over which geographic area you want to draw an isoscape?
- At which resolution?
- Using which aggegation function?

Step 3. Prepare the structural raster

The most important questions you must answer:

- Over which geographic area you want to draw an isoscape?
- At which resolution?
- Using which aggegation function?

Issue:

- Higher resolution = more precise, but requires more computer memory and computation time
- For the aggregation function, the mean is not always the best (it depends on the question!)

Step 3. Prepare the structural raster

The most important questions you must answer:

- Over which geographic area you want to draw an isoscape?
- At which resolution?
- Using which aggegation function?

Issue:

- Higher resolution = more precise, but requires more computer memory and computation time
- For the aggregation function, the mean is not always the best (it depends on the question!)

Recommendations:

- Start coarse
- Then, increase the resolution as high as it remains reasonable (below the pixel level makes little sense)

Step 3. Prepare the structural raster

Uncertainty check list:

- × uncertainty behind the raw raster
  (structural rasters are the results of some smoothing or interpolation procedure which contains uncertainty)
- ✓ variation between the cells of the processed raster
- × variation within the cells of the processed raster

Step 3. Prepare the structural raster

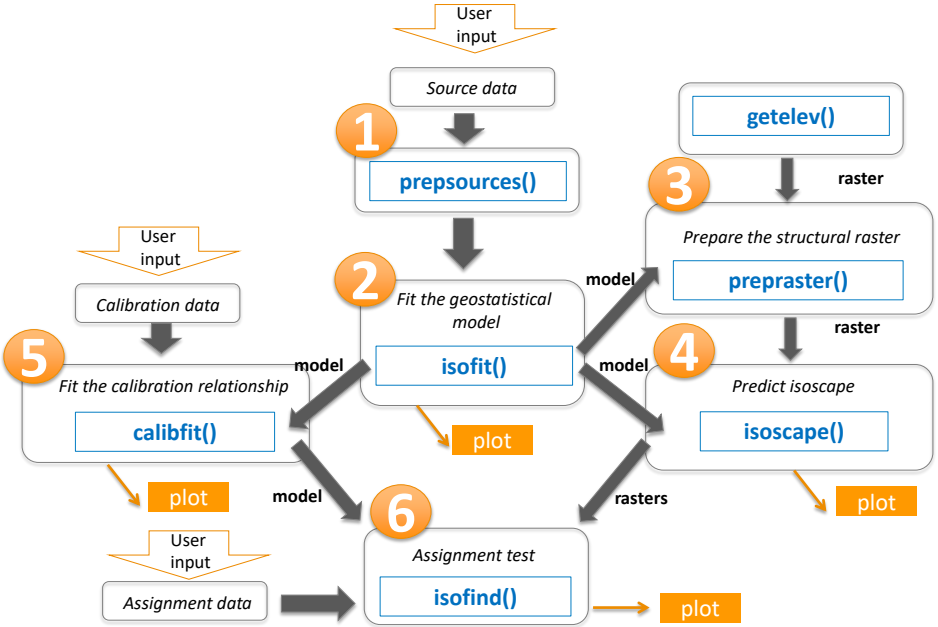
Uncertainty check list:

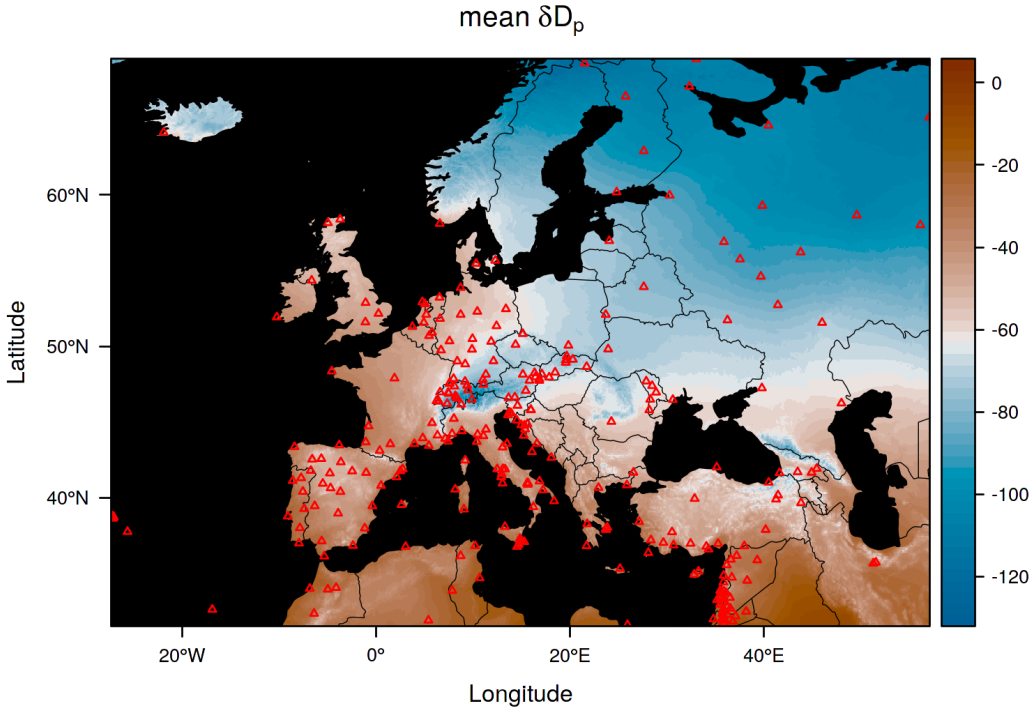
- × uncertainty behind the raw raster
  (structural rasters are the results of some smoothing or interpolation procedure which contains uncertainty)
- ✓ variation between the cells of the processed raster
- × variation within the cells of the processed raster

Recommendations:

- Get to know how the raster you use has been created:
  (`getelev()` uses this: `https://github.com/tilezen/joerd/blob/master/docs/data-sources.md`)
  - Which uncertainty? (not just the average one)
  - Which smoothing or aggregation function?
- Increasing the resolution reduces the variance not considered!

Table of contents


User input
Source data
1
prepsources()
getelev()
raster
3
Prepare the structural raster
prepraster()
raster
User input
Calibration data
2
Fit the geostatistical model
isofit()
model
model
5
Fit the calibration relationship
calibfit()
model
model
4
Predict isoscape
isoscape()
plot
plot
plot
model
6
rasters
User input
Assignment data
Assignment test
isofind()
plot


mean δD_p
60°N
50°N
40°N
Latitude
20°W
0°
20°E
40°E
Longitude
0
-20
-40
-60
-80
-100
-120

Step 4. Predict the isoscape

4a. Use `isoscape()` predicts the isoscape:

```
isoscape(
  raster,
  isofit,
  verbose = interactive()
  )
```

Step 4. Predict the isoscape

4a. Use `isoscape()` predicts the isoscape:

```
isoscape(
  raster,
  isofit,
  verbose = interactive()
  )
```

Example:

```
GermanScape <- isoscape(raster = ElevRasterDE, isofit = GermanFit)
```

```
names(GermanScape)
## [1] "isoscapes" "sp_points"
```

Step 4. Predict the isoscape

There are 8 rasters behind the isoscape:

```
GermanScape$isoscapes
## class      : RasterBrick
## dimensions : 213, 257, 54741, 8  (nrow, ncol, ncell, nlayers)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      :          mean,  mean_predVar, mean_residVar,  mean_respVar,          disp,  disp_predVar, disp_residVar,  disp_respVar
## min values : -1.047537e+02,  5.903244e-01,  1.953206e+02,  2.032967e+02,  1.953206e+02,  1.405766e-02,  7.630030e+04,  7.894167e+04
## max values : -4.672956e+01,  3.747816e+01,  7.748391e+02,  7.972232e+02,  7.748391e+02,  8.226663e-02,  1.200751e+06,  1.214789e+06
```

Step 4. Predict the isoscape

There are 8 rasters behind the isoscape:

```
GermanScape$isoscapes
## class      : RasterBrick
## dimensions : 213, 257, 54741, 8  (nrow, ncol, ncell, nlayers)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      :          mean,  mean_predVar, mean_residVar,  mean_respVar,          disp,  disp_predVar, disp_residVar,  disp_respVar
## min values : -1.047537e+02,  5.903244e-01,  1.953206e+02,  2.032967e+02,  1.953206e+02,  1.405766e-02,  7.630030e+04,  7.894167e+04
## max values : -4.672956e+01,  3.747816e+01,  7.748391e+02,  7.972232e+02,  7.748391e+02,  8.226663e-02,  1.200751e+06,  1.214789e+06
```

- `mean` = prediction of the mean isotopic value
- `mean_predVar` = quantify the uncertainty associated to `mean`
- `mean_residVar` = prediction of the variance between observations

Step 4. Predict the isoscape

There are 8 rasters behind the isoscape:

```
GermanScape$isoscapes
## class      : RasterBrick
## dimensions : 213, 257, 54741, 8  (nrow, ncol, ncell, nlayers)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      :          mean,  mean_predVar, mean_residVar,  mean_respVar,          disp,  disp_predVar, disp_residVar,  disp_respVar
## min values : -1.047537e+02,  5.903244e-01,  1.953206e+02,  2.032967e+02,  1.953206e+02,  1.405766e-02,  7.630030e+04,  7.894167e+04
## max values : -4.672956e+01,  3.747816e+01,  7.748391e+02,  7.972232e+02,  7.748391e+02,  8.226663e-02,  1.200751e+06,  1.214789e+06
```

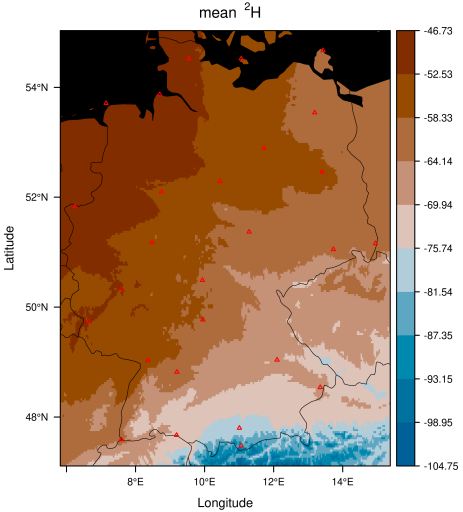
- `mean` = prediction of the mean isotopic value
- `mean_predVar` = quantify the uncertainty associated to `mean`
- `mean_residVar` = prediction of the variance between observations
- `mean_respVar` = `mean_predVar` + `mean_residVar`
- `disp` = `mean_residVar`
- `disp_predVar` = quantify the uncertainty associated to `disp`
- `disp_residVar` = 2
- `disp_respVar` = exp(`disp_predVar` + `disp_residVar`)

4b. Plot the rasters with `plot()`:

```
plot.ISOSCAPE(
  x,
  which   = "mean",
  y_title = list(which = TRUE, title = bquote(delta^2 * H)),
  sources = list(draw = TRUE, cex = 0.5, pch = 2, lwd = 1, col = "red"),
  borders = list(borders = NA, lwd = 0.5, col = "black"),
  mask    = list(mask = NA, lwd = 0, col = "black", fill = "black"),
  palette = list(step = NA, range = c(NA, NA), n_labels = 11, digits = 2, fn = NA),
  plot    = TRUE,
  sphere  = list(build = FALSE, keep_image = FALSE),
  ...
  )
```

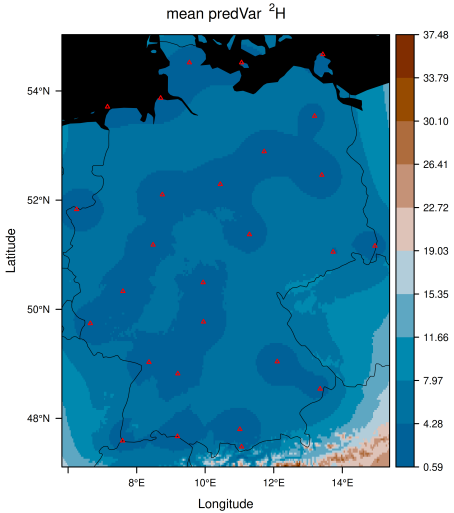
4b. Plot the rasters with `plot()`:

```
plot(GermanScape)
```



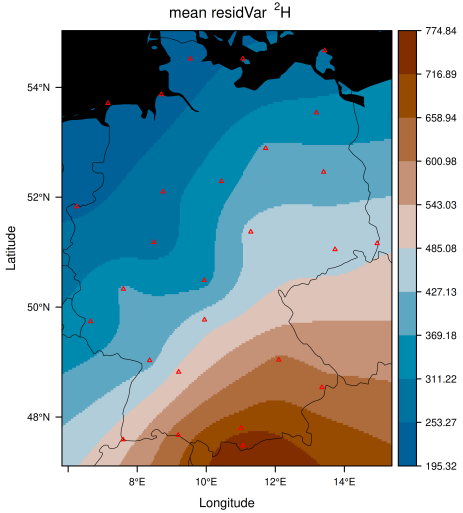
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     which = "mean_predVar")
```



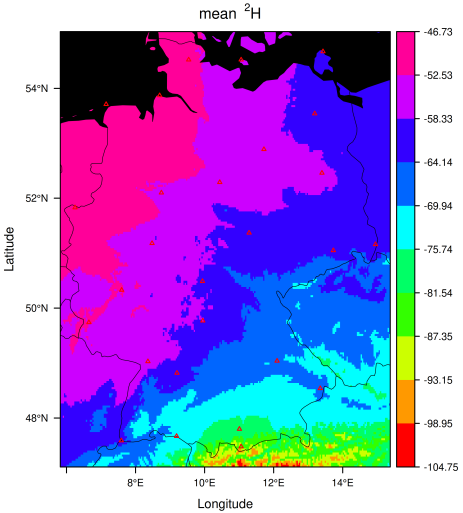
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     which = "mean_residVar")
```



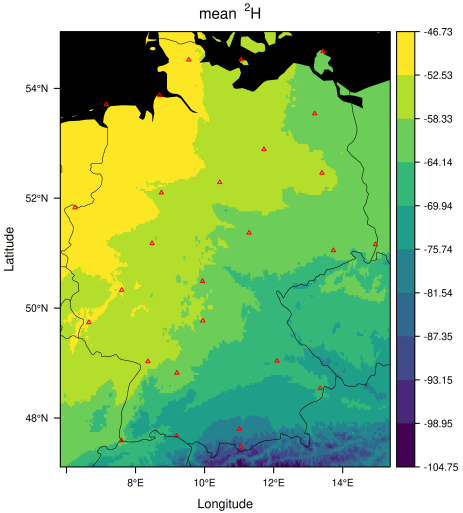
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     palette = list(fn = rainbow))
```



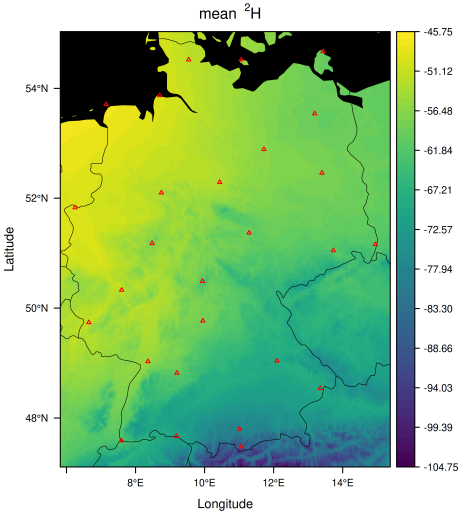
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     palette = list(fn = NULL))
```



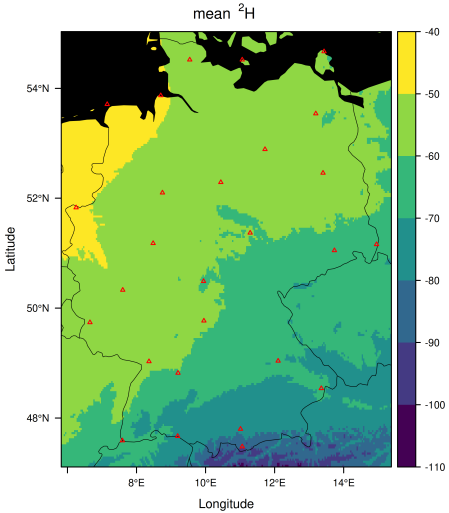
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     palette = list(fn = NULL, step = 1))
```



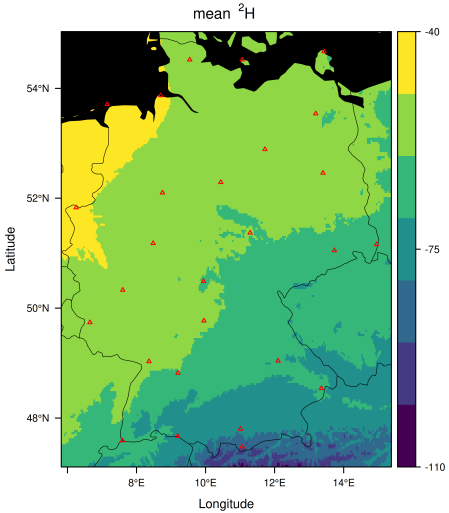
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     palette = list(fn = NULL, step = 10, range = c(-110, -40)))
```



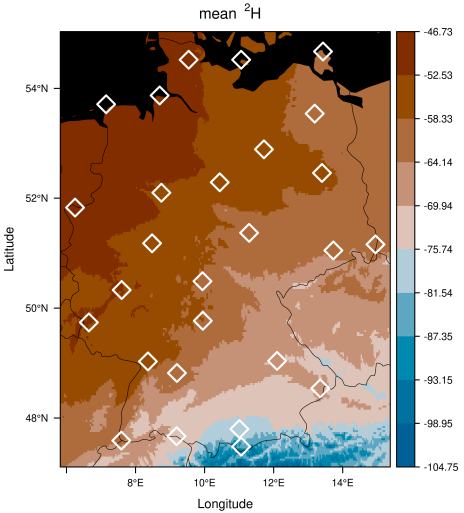
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     palette = list(fn = NULL, step = 10, range = c(-110, -40), n_labels = 3))
```



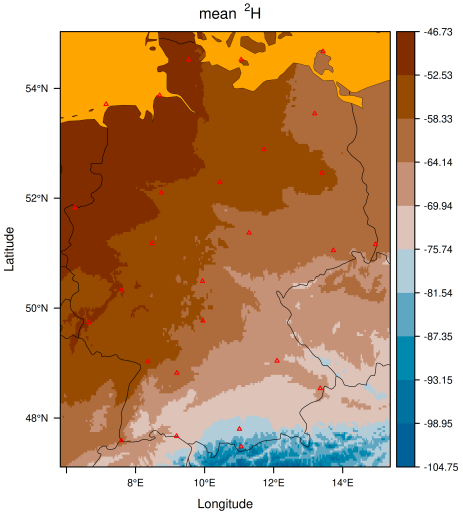
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     sources = list(draw = TRUE, cex = 2, pch = 5, lwd = 2, col = "white"))
```



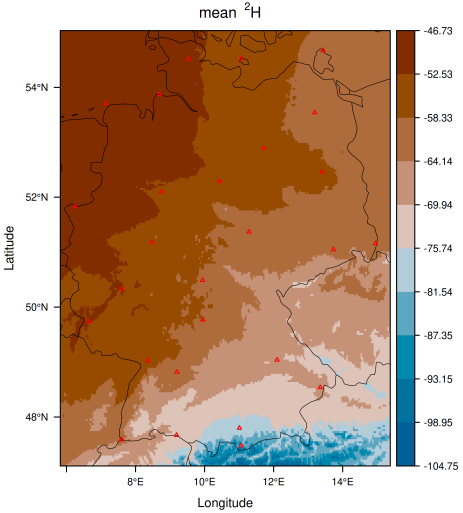
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     mask = list(fill = "orange"))
```



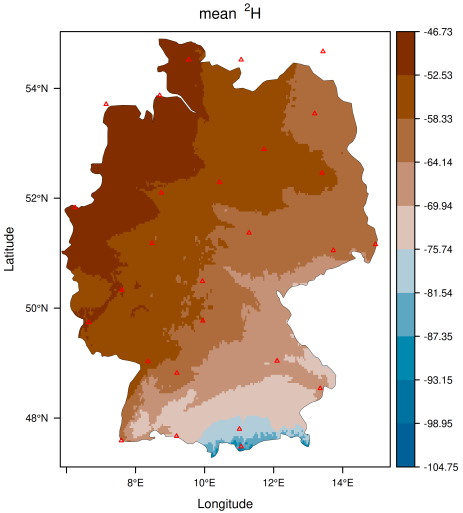
4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
     mask = list(mask = NULL))
```



4b. Plot the rasters with `plot()`:

```
plot(GermanScape,
mask = list(mask = rbind(OceanMask, CountryBorders[names(CountryBorders) != "Germany"]), fill = "white"))
```



Step 4. Predict the isoscape

If the plotting possibilities of IsoriX are not good enough for you, you can always export the raster for other GIS software (e.g. QGIS).

Example:

```
library(raster)

writeRaster(GermanScape$isoscapes$mean,
            filename = "GermanScape.asc",
            format = "ascii",
            overwrite = TRUE,
            NAflag = -9999)

writeRaster(GermanScape$isoscapes$mean,
            filename = "GermanScape.tif",
            format = "GTiff",
            overwrite = TRUE,
            NAflag = -9999)
```

Note: you can also export other spatial objects included or created with IsoriX.

Step 4. Predict the isoscape

Uncertainty check list → same as for `isofit` ⏩ Step 2.

Step 4. Predict the isoscape

Uncertainty check list → same as for `isofit` ⏩ Step 2.

The most important questions you must answer:

- How many colours?
- Which colours?

Step 4. Predict the isoscape

Uncertainty check list → same as for `isofit` ⏩ Step 2.

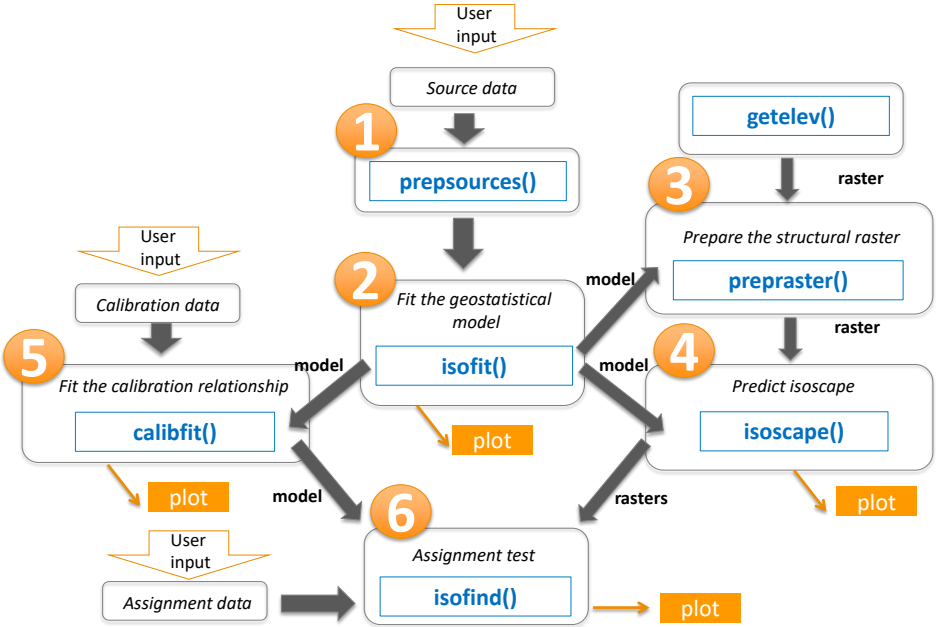
The most important questions you must answer:

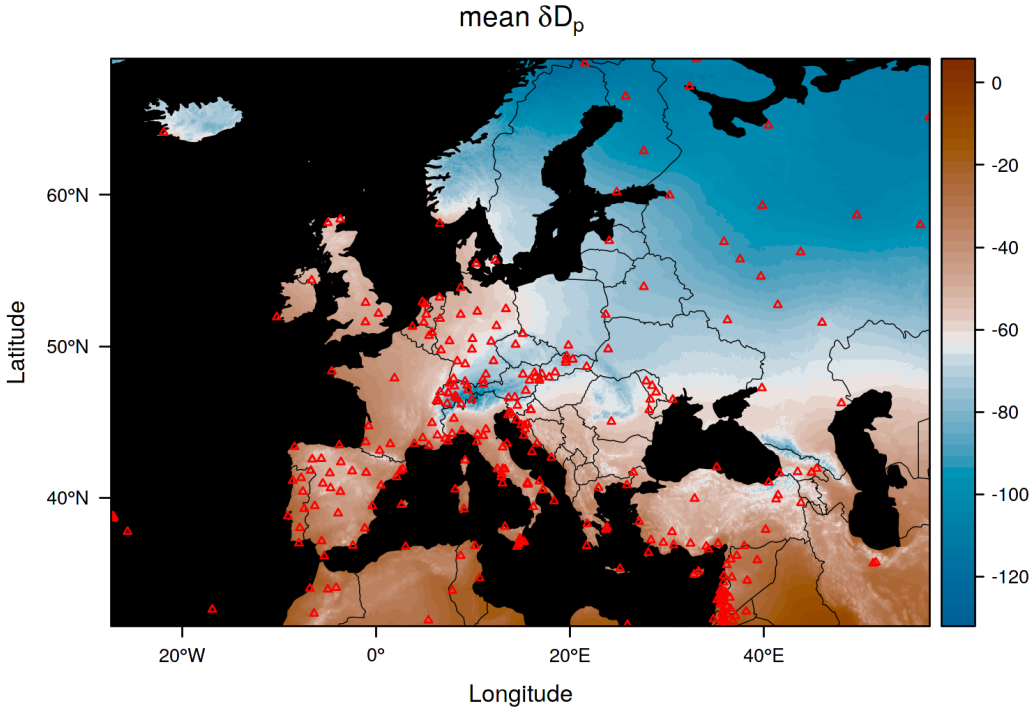
- How many colours?
- Which colours?

Recommendation:

- Perceptually uniform palettes are best
- Think of B&W print, think of colour-blind people

Table of contents


User input
Source data
1
prepsources()
getelev()
raster
3
Prepare the structural raster
prepraster()
raster
User input
Calibration data
2
Fit the geostatistical model
isofit()
model
model
5
Fit the calibration relationship
calibfit()
model
4
Predict isoscape
isoscape()
plot
plot
plot
model
6
rasters
User input
Assignment data
Assignment test
isofind()
plot


mean δD_p
60°N
50°N
40°N
Latitude
20°W
0°
20°E
40°E
Longitude
0
-20
-40
-60
-80
-100
-120

Step 5. Fit the calibration function

A calibration procedure is needed whenever the isoscape is not based on the same material as the one you want to assign (e.g. precipitation water vs bats, plancton vs penguins, ...).

The calibration function relates the isotopic composition of samples of known origin to the isotopic composition of the environment at these origin.

It is required to predict what would be the isotopic composition of animals originating from any given position in an isoscape.

Step 5. Fit the calibration function

A calibration procedure is needed whenever the isoscape is not based on the same material as the one you want to assign (e.g. precipitation water vs bats, plancton vs penguins, ...).

The calibration function relates the isotopic composition of samples of known origin to the isotopic composition of the environment at these origin.

It is required to predict what would be the isotopic composition of animals originating from any given position in an isoscape.

Isorix can handle four different calibration methods (here we will use the `"wild"` method, the default):

Method	**Isotopic composition of the calibration samples**	**Isotopic composition of the environment associated with the calibration samples**
`"wild"`	known	unknown, estimated from the isoscapes
`"lab"`	known	known, or assumed to be known
`"desk"`	unknown, estimated from `lm(sample_value ~ source_value)`	unknown, estimated from `lm(sample_value ~ source_value)`
`"desk_inverse"`	unknown, estimated from `lm(source_value ~ sample_value)`	unknown, estimated from `lm(source_value ~ sample_value)`

Step 5. Fit the calibration function

5a. Prepare your calibration data:

```
head(CalibDataBat)
```

```
##   site_ID     long      lat elev sample_ID species sample_value
## 1 site_41 18.87101 46.19747   90         1       4        -71.8
## 2 site_41 18.87101 46.19747   90         2       4       -103.4
## 3 site_41 18.87101 46.19747   90         3       4        -68.9
## 4 site_41 18.87101 46.19747   90         4       4       -103.3
## 5 site_67 19.10489 48.53405  522         5       4        -92.3
## 6 site_67 19.10489 48.53405  522         6       4       -102.0
```

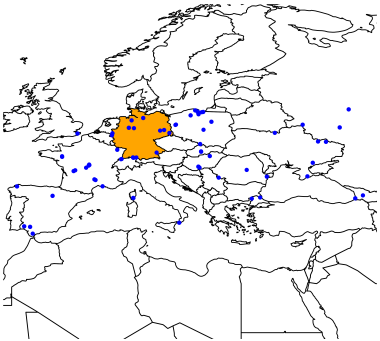
```
str(CalibDataBat)
```

```
## 'data.frame': 335 obs. of  7 variables:
##  $ site_ID     : Factor w/ 71 levels "site_10","site_11",..: 33 33 33 33 60 60 60 11 11 11 ...
##  $ long        : num  18.9 18.9 18.9 18.9 19.1 ...
##  $ lat         : num  46.2 46.2 46.2 46.2 48.5 ...
##  $ elev        : int  90 90 90 90 522 522 522 20 20 20 ...
##  $ sample_ID   : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ species     : Factor w/ 6 levels "1","2","3","4",..: 4 4 4 4 4 4 4 4 4 4 ...
##  $ sample_value: num  -71.8 -103.4 -68.9 -103.3 -92.3 ...
```

Note: you need several replicates per location!

5b. Check your calibration data:

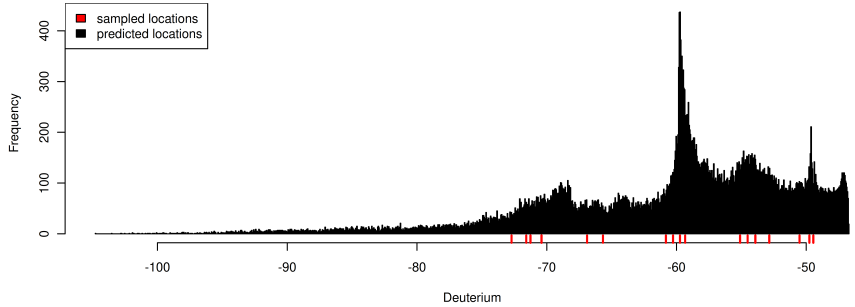
```
plot(CountryBorders, xlim = range(CalibDataBat$long), ylim = range(CalibDataBat$lat))
plot(CountryBorders["Germany"], col = "orange", add = TRUE)
plot(SpatialPoints(coords = CalibDataBat[, c("long", "lat")]), col = "blue", pch = 20, add = TRUE)
```



Note: this is NOT good! (but we will do with that for this tutorial)

5b. Check your calibration data:

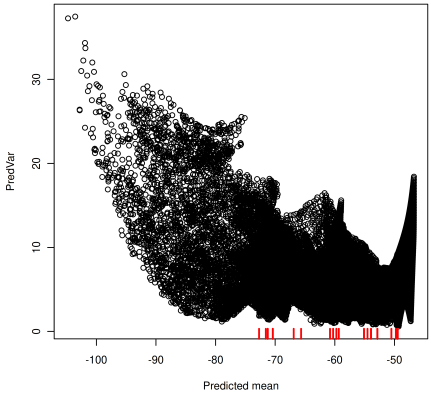
```
IsoPredicted <- values(GermanScape$isoscapes$mean)
IsoAtBats <- extract(GermanScape$isoscapes$mean, SpatialPoints(coords = CalibDataBat[, c("long", "lat")]))
hist(IsoPredicted, nclass = 1000, main = "", xlab = "Deuterium")
rug(IsoAtBats, col = "red", lwd = 2)
legend("topleft", fill = c("red", "black"), legend = c("sampled locations", "predicted locations"))
```



Note: this is quite good but very few points are considered in this plot (only the German ones).

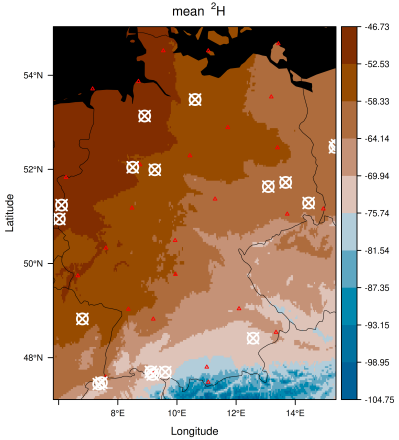
5b. Check your calibration data:

```
IsoPredictedPredVar <- values(GermanScape$isoscapes$mean_predVar)
plot(x = IsoPredicted, y = IsoPredictedPredVar, xlab = "Predicted mean", ylab = "PredVar")
rug(IsoAtBats, col = "red", lwd = 2)
```



5b. Check your calibration data:

```
plot(GermanScape, plot = FALSE) +
  xyplot(CalibDataBat$lat ~ CalibDataBat$long,
         pch = 13, col = "white", cex = 2, lwd = 2, panel = panel.points)
```



The coverage within Germany is quite good but all the location outside Germany will be considered too.

Step 5. Fit the calibration function

5c. Use `calibfit()` to fit the calibration function:

```
calibfit(
  data,
  isofit = NULL,
  method = c("wild", "lab", "desk", "desk_inverse"),
  verbose = interactive(),
  control_optim = list()
  )
```

Step 5. Fit the calibration function

5c. Use `calibfit()` to fit the calibration function:

```
calibfit(
  data,
  isofit = NULL,
  method = c("wild", "lab", "desk", "desk_inverse"),
  verbose = interactive(),
  control_optim = list()
  )
```

Example:

```
CalibBats <- calibfit(data = CalibDataBat, isofit = GermanFit)
```

```
## Note:  extrapolation issues
## Out of your 335 calibration samples,
##  * 272 correspond to locations outside the area covered by the measurements you used to build your isoscape.
##  * 1 are associated to predicted values more extreme than the ones present in the isoscape.
## -> These cases correspond to extrapolation during the calibration step, which could imped the reliability of your assignments.
## If the proportion of problematic samples is large, you should perhaps rethink the design of your isoscape and/or collect more callibration data within the expected range to avoid any problem.
```

Step 5. Fit the calibration function

5d. Check the fitted the calibration model:

```
CalibBats
```

```
##
## Fixed effect estimates of the calibration fit
## sample_value = intercept + slope * mean_source_value + corrMatrix(1|site_ID) + slope^2 * (1|site_ID) + Error
##
##            intercept (+/- SE) = -35.44 +/- 8.56
##            slope     (+/- SE) = 0.87 +/- 0.15
##
## [for more information, use summary()]
```

5d. Check the fitted the calibration model:

```
CalibBats
```

```
##
## Fixed effect estimates of the calibration fit
## sample_value = intercept + slope * mean_source_value + corrMatrix(1|site_ID) + slope^2 * (1|site_ID) + Error
##
##            intercept (+/- SE) = -35.44 +/- 8.56
##            slope     (+/- SE) = 0.87 +/- 0.15
##
## [for more information, use summary()]
```

The model is quite complex...

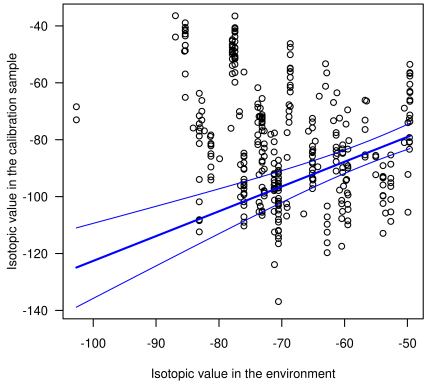
We would like to fit the model $\text{calib}_{gi} = \beta_0^C + \beta_\delta^C \delta_g + \epsilon_{gi}^C$.

However we often don't have δ_g, but only $\hat{\delta}_g$ and you should not fit $\text{calib}_{gi} = \beta_0^C + \beta_\delta^C \hat{\delta}_g + \varepsilon_{gi}^C$, because there is error on y (usual) but also on x (violating exogeneity assumption).

Which leads us to $\text{calib}_{gi} = \beta_0^C + \beta_\delta^C \hat{\delta}_g + \beta_\delta^C \left(\delta_g - \hat{\delta}_g\right) + \epsilon_{gi}^C$, where difference $\delta_g - \hat{\delta}_g$ represents the prediction error of the mean source value at the sampling location g by the mean fit. We deduce the prediction error from the prediction covariance matrix (hence the square above).

5d. Check the fitted the calibration model:

```
plot(CalibBats)
```



Note: this looks bad, but in fact it is incredibly good! Despite huge extrapolation, the fit is very close to a calibration fitted on an European isoscape (see `https://bookdown.org/content/782/calibration.html`). The reason why it works so well is that the uncertainty in x is accounted for.

Step 5. Fit the calibration function

The most important questions you must answer:

- How to get the best possible calibration data?
- Do they sample the full range of predicted source isoscape values?
- Are the samples of similar type than the ones you want to assign?
  - same species?
  - same reproductive state?
  - same age?
  - same diet?
  - ...

Step 5. Fit the calibration function

The most important questions you must answer:

- How to get the best possible calibration data?
- Do they sample the full range of predicted source isoscape values?
- Are the samples of similar type than the ones you want to assign?
  - same species?
  - same reproductive state?
  - same age?
  - same diet?
  - ...

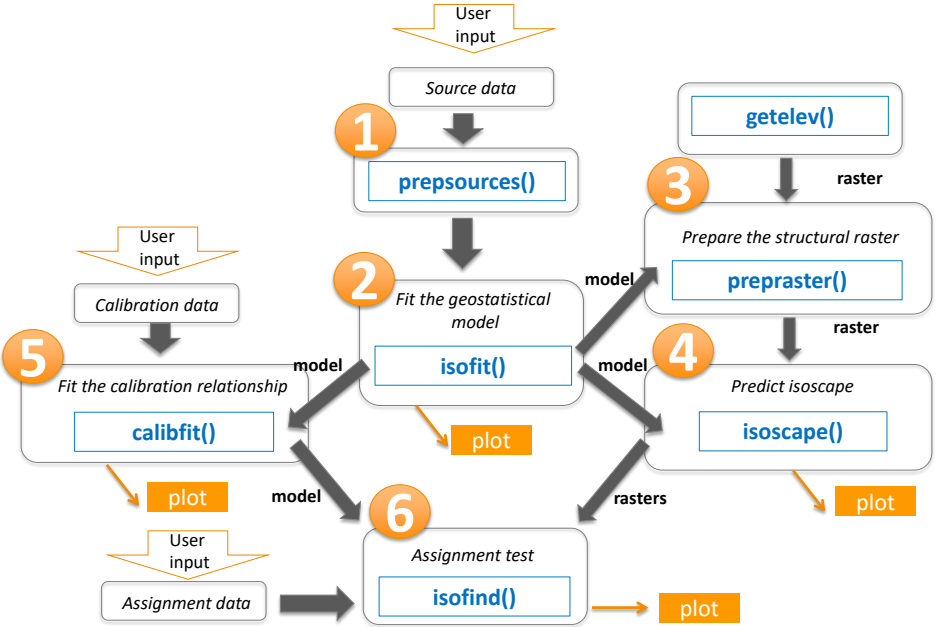
Recommendation:

- do your very best, this is a critical step.

Uncertainty check list:

- ✓ uncertainty in the isoscape
- ✓ uncertainty in the intercept and slope of the calibration function
- ✓ residual uncertainty

Table of contents


User input
Source data
1
prepsources()
getelev()
raster
3
Prepare the structural raster
prepraster()
raster
User input
Calibration data
2
Fit the geostatistical model
isofit()
model
model
5
Fit the calibration relationship
calibfit()
model
model
4
Predict isoscape
isoscape()
plot
plot
plot
model
rasters
6
User input
Assignment data
Assignment test
isofind()
plot

Step 6. Perform the assignment

6a. Prepare your assignment data:

```
AssignDataBat
##    sample_ID      lat     long sample_value
## 1     Nnoc_1 51.72323 13.67630    -86.58500
## 2     Nnoc_2 51.72323 13.67630    -93.96700
## 3     Nnoc_3 51.72323 13.67630    -88.71600
## 4     Nnoc_4 51.69520 14.64945    -84.85700
## 5     Nnoc_5 51.70366 14.63915   -104.21900
## 6     Nnoc_7 51.69520 14.64945    -85.58300
## 7     Nnoc_8 51.70536 14.63289    -89.99500
## 8     Nnoc_9 51.70536 14.63289    -92.62700
## 9    Nnoc_10 51.89348 13.76488    -85.55300
## 10   Nnoc_11 51.84883 13.64525    -87.05000
## 11   Nnoc_12 51.60535 11.92700    -89.34916
## 12   Nnoc_13 51.96695 12.08462   -103.65232
## 13   Nnoc_15 51.96695 12.08462   -121.18926
## 14   Nnoc_16 51.45457 11.50689    -98.49742
```

Our question: among the bats found dead at the bottom of wind turbines, could some be migratory bats?

Step 6. Perform the assignment

6b. Use `isofind()` to perform the assignment tests:

```
isofind(
  data,
  isoscape,
  calibfit = NULL,
  mask = NA,
  neglect_covPredCalib = TRUE,
  verbose = interactive()
  )
```

Step 6. Perform the assignment

6b. Use `isofind()` to perform the assignment tests:

```
isofind(
  data,
  isoscape,
  calibfit = NULL,
  mask = NA,
  neglect_covPredCalib = TRUE,
  verbose = interactive()
  )
```

Example:

```
AssignedBats <- isofind(data = AssignDataBat, isoscape = GermanScape, calibfit = CalibBats)

names(AssignedBats)
## [1] "sample"    "group"     "sp_points"
```

Step 6. Perform the assignment

`isofind()` returns the result of the assignment test for each sample:

```
AssignedBats$sample
```

```
## $stat
## class      : RasterBrick
## dimensions : 213, 257, 54741, 14  (nrow, ncol, ncell, nlayers)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      :     Nnoc_1,     Nnoc_2,     Nnoc_3,     Nnoc_4,     Nnoc_5,     Nnoc_7,     Nnoc_8,     Nnoc_9,    Nnoc_10,    Nnoc_11,    Nnoc_12,    Nnoc_13,    Nnoc_
## min values : -11.508848, -19.980661, -13.954450,  -9.525742, -31.746175, -10.358922, -15.422270, -18.442835, -10.324493, -12.042497, -14.681080, -31.095837, -51.2217
## max values :  46.059437,  37.587624,  43.613835,  48.042543,  25.822110,  47.209363,  42.146015,  39.125450,  47.243792,  45.525788,  42.887205,  26.472447,   6.3465
## 
## 
## $stat_var
## class      : RasterBrick
## dimensions : 213, 257, 54741, 14  (nrow, ncol, ncell, nlayers)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      :   Nnoc_1,   Nnoc_2,   Nnoc_3,   Nnoc_4,   Nnoc_5,   Nnoc_7,   Nnoc_8,   Nnoc_9,  Nnoc_10,  Nnoc_11,  Nnoc_12,  Nnoc_13,  Nnoc_15,  Nnoc_16 
## min values :       0,       0,       0,       0,       0,       0,       0,       0,       0,       0,       0,       0,       0,       0 
## max values : 161.3832, 164.5269, 161.8698, 161.2395, 175.6926, 161.2725, 162.3259, 163.6517, 161.2704, 161.4603, 162.0802, 174.8689, 211.5515, 168.4861 
## 
## 
## $pv
## class      : RasterBrick
## dimensions : 213, 257, 54741, 14  (nrow, ncol, ncell, nlayers)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      :    Nnoc_1,    Nnoc_2,    Nnoc_3,    Nnoc_4,    Nnoc_5,    Nnoc_7,    Nnoc_8,    Nnoc_9,   Nnoc_10,   Nnoc_11,   Nnoc_12,   Nnoc_13,   Nnoc_15,   Nnoc_16 
## min values :         0,         0,         0,         0,         0,         0,         0,         0,         0,         0,         0,         0,         0,         0 
## max values : 0.9999941, 0.9999987, 0.9999991, 0.9999842, 0.9999896, 0.9999940, 0.9999849, 0.9999993, 0.9999864, 0.9999899, 0.9999755, 0.9998770, 0.9997543, 0.9999976
```

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).
- The null hypothesis of the test is that the sample comes from the cell being examined.

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).
- The null hypothesis of the test is that the sample comes from the cell being examined.
- A large p-value (i.e. non significant) indicates support for the null hypothesis.

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).
- The null hypothesis of the test is that the sample comes from the cell being examined.
- A large p-value (i.e. non significant) indicates support for the null hypothesis.
- A small p-value (i.e. significant) indicates the rejection of the null hypothesis.

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).
- The null hypothesis of the test is that the sample comes from the cell being examined.
- A large p-value (i.e. non significant) indicates support for the null hypothesis.
- A small p-value (i.e. significant) indicates the rejection of the null hypothesis.

Note:

- A very high p-value does not imply certainty of the location of origin, only that the isotopic signature between the location of origin and the location under evaluation are similar.

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).
- The null hypothesis of the test is that the sample comes from the cell being examined.
- A large p-value (i.e. non significant) indicates support for the null hypothesis.
- A small p-value (i.e. significant) indicates the rejection of the null hypothesis.

Note:

- A very high p-value does not imply certainty of the location of origin, only that the isotopic signature between the location of origin and the location under evaluation are similar.
- It is possible that all candidate locations are rejected! (unlike relative Bayesian probabilities)

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).
- The null hypothesis of the test is that the sample comes from the cell being examined.
- A large p-value (i.e. non significant) indicates support for the null hypothesis.
- A small p-value (i.e. significant) indicates the rejection of the null hypothesis.

Note:

- A very high p-value does not imply certainty of the location of origin, only that the isotopic signature between the location of origin and the location under evaluation are similar.
- It is possible that all candidate locations are rejected! (unlike relative Bayesian probabilities)
- A lack of rejection may stem from a good match, a large uncertainty, or both.

Step 6. Perform the assignment

Some important facts about the assignment test:

- The test is performed at the level of each sample and each cell defined by the structural raster.
- The test statistics is the difference between the predicted isoscape value at the origin location (deduced from the calibration fit) and the predicted isoscape value at the location under consideration (read from the isoscape).
- The null hypothesis of the test is that the sample comes from the cell being examined.
- A large p-value (i.e. non significant) indicates support for the null hypothesis.
- A small p-value (i.e. significant) indicates the rejection of the null hypothesis.

Note:

- A very high p-value does not imply certainty of the location of origin, only that the isotopic signature between the location of origin and the location under evaluation are similar.
- It is possible that all candidate locations are rejected! (unlike relative Bayesian probabilities)
- A lack of rejection may stem from a good match, a large uncertainty, or both.
- The statistics behind the test are quite complex; see Appendix of book chapter.

Step 6. Perform the assignment

You can easily extract the outcome of the assignment test in the cell corresponding to a given location!

Example 1: may the first bat (`Nnoc_1`) come from where it has been found?

```
extract(AssignedBats$sample$stat[[1]], cbind(AssignDataBat$long[1], AssignDataBat$lat[1]))
## [1] 1.038767
extract(AssignedBats$sample$pv[[1]], cbind(AssignDataBat$long[1], AssignDataBat$lat[1]))
## [1] 0.9271638
```

Step 6. Perform the assignment

You can easily extract the outcome of the assignment test in the cell corresponding to a given location!

Example 1: may the first bat (`Nnoc_1`) come from where it has been found?

```
extract(AssignedBats$sample$stat[[1]], cbind(AssignDataBat$long[1], AssignDataBat$lat[1]))
## [1] 1.038767
extract(AssignedBats$sample$pv[[1]], cbind(AssignDataBat$long[1], AssignDataBat$lat[1]))
## [1] 0.9271638
```

Answer: it is possible (we cannot reject the proposition).

Step 6. Perform the assignment

You can easily extract the outcome of the assignment test in the cell corresponding to a given location!

Example 1: may the first bat (`Nnoc_1`) come from where it has been found?

```
extract(AssignedBats$sample$stat[[1]], cbind(AssignDataBat$long[1], AssignDataBat$lat[1]))
## [1] 1.038767
extract(AssignedBats$sample$pv[[1]], cbind(AssignDataBat$long[1], AssignDataBat$lat[1]))
## [1] 0.9271638
```

Answer: it is possible (we cannot reject the proposition).

Example 2: may the bat 13 (`Nnoc_15`) come from where it has been found?

```
extract(AssignedBats$sample$stat[[13]], cbind(AssignDataBat$long[13], AssignDataBat$lat[13]))
## [1] -40.48787
extract(AssignedBats$sample$pv[[13]], cbind(AssignDataBat$long[13], AssignDataBat$lat[13]))
## [1] 0.0024878
```

Answer: it is very unlikely (we can reject the proposition).

Step 6. Perform the assignment

You can easily extract the outcome of the assignment test in the cell corresponding to a given location!

Example 3: which bats are unlikely to come from where they have been found?

```
Pvalues <- sapply(1:nrow(AssignDataBat),
                  function(i) extract(AssignedBats$sample$pv[[i]],
                                      cbind(AssignDataBat$long[i],
                                            AssignDataBat$lat[i])))
AssignDataBat$sample_ID[Pvalues <= 0.05]
```

```
## [1] Nnoc_15
## Levels: Nnoc_1 Nnoc_10 Nnoc_11 Nnoc_12 Nnoc_13 Nnoc_15 Nnoc_16 Nnoc_2 Nnoc_3 Nnoc_4 Nnoc_5 Nnoc_7 Nnoc
```

Answer: only `Nnoc_15` seem to be a migrant.

Step 6. Perform the assignment

`isofind()` also returns the result of the assignment test for the whole group of samples:

```
AssignedBats$group
```

```
## $pv
## class      : RasterLayer
## dimensions : 213, 257, 54741  (nrow, ncol, ncell)
## resolution : 0.03719131, 0.03719131  (x, y)
## extent     : 5.820439, 15.3786, 47.111, 55.03275  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +no_defs
## source     : memory
## names      : layer
## values     : 0, 0.7532592  (min, max)
```

Note:
The null hypothesis is here that all samples come from the location being examined.
(so it can often be rejected anywhere)

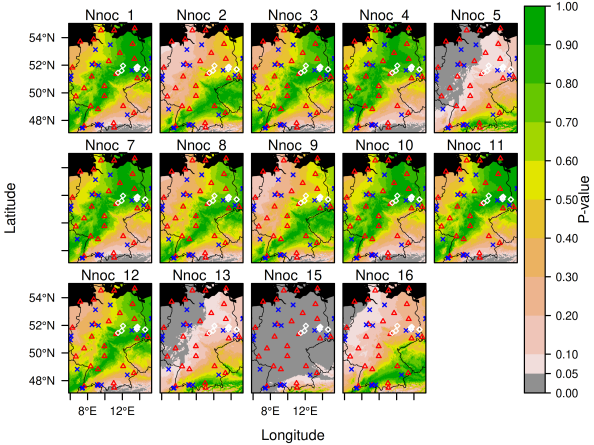
6c. Plot the assignment tests with `plot()`:

```
plot(
  x,
  who      = "group",
  cutoff   = list(draw = TRUE, level = 0.05, col = "#909090"),
  sources  = list(draw = TRUE, cex = 0.5, pch = 2, lwd = 1, col = "red"),
  calibs   = list(draw = TRUE, cex = 0.5, pch = 4, lwd = 1, col = "blue"),
  assigns  = list(draw = TRUE, cex = 0.5, pch = 5, lwd = 1, col = "white"),
  borders  = list(borders = NA, lwd = 0.5, col = "black"),
  mask     = list(mask = NA, lwd = 0, col = "black", fill = "black"),
  mask2    = list(mask = NA, lwd = 0, col = "purple", fill = "purple"),
  palette  = list(step = NA, range = c(0, 1), n_labels = 11, digits = 2, fn = NA),
  plot     = TRUE,
  sphere   = list(build = FALSE, keep_image = FALSE),
  ...
  )
```

It is quite similar to the function plotting the isoscapes.

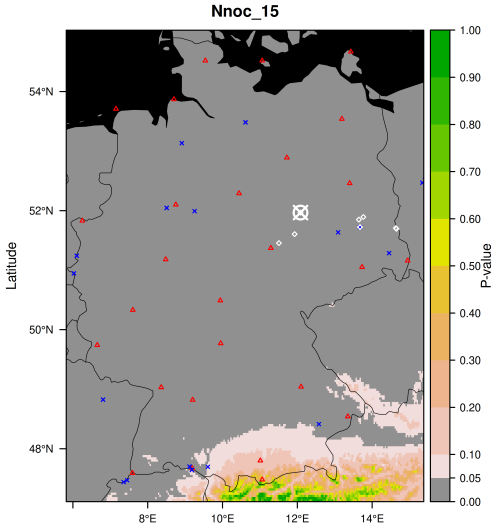
6c. Plot the assignment tests with `plot()`:

```
plot(AssignedBats, who = 1:14)
```



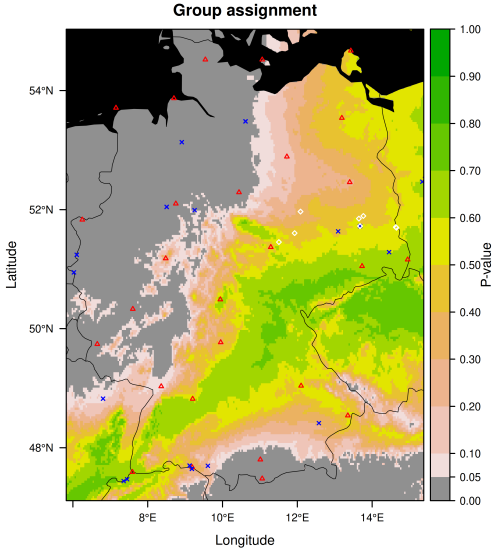
6c. Plot the assignment tests with `plot()`:

```
plot(AssignedBats, who = "Nnoc_15", plot = FALSE) +
  xyplot(AssignDataBat$lat[AssignDataBat$sample_ID == "Nnoc_15"] ~ AssignDataBat$long[AssignDataBat$sample_ID == "Nnoc_15"],
         pch = 13, col = "white", cex = 2, lwd = 2, panel = panel.points)
```



6c. Plot the assignment tests with `plot()`:

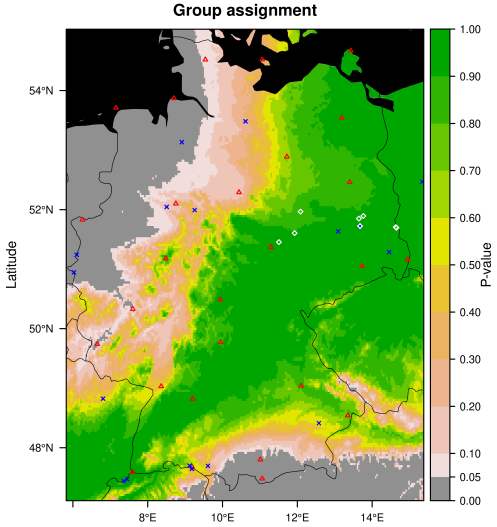
```
plot(AssignedBats) ## plot the group assignment
```



Step 6. Perform the assignment

We can rerun the group assignment without Nnoc_15:

```
AssignDataBat2 <- subset(AssignDataBat, sample_ID != "Nnoc_15")
AssignedBats2 <- isofind(data = AssignDataBat2, isoscape = GermanScape, calibfit = CalibBats)
plot(AssignedBats2)
```



Step 6. Perform the assignment

The most important questions you must answer:

- Are all samples coming from the same location?
- Can the true origin be outside the map?

Step 6. Perform the assignment

The most important questions you must answer:

- Are all samples coming from the same location?
- Can the true origin be outside the map?

Recommendation:

- Always look at plot per sample before looking at group assignment.

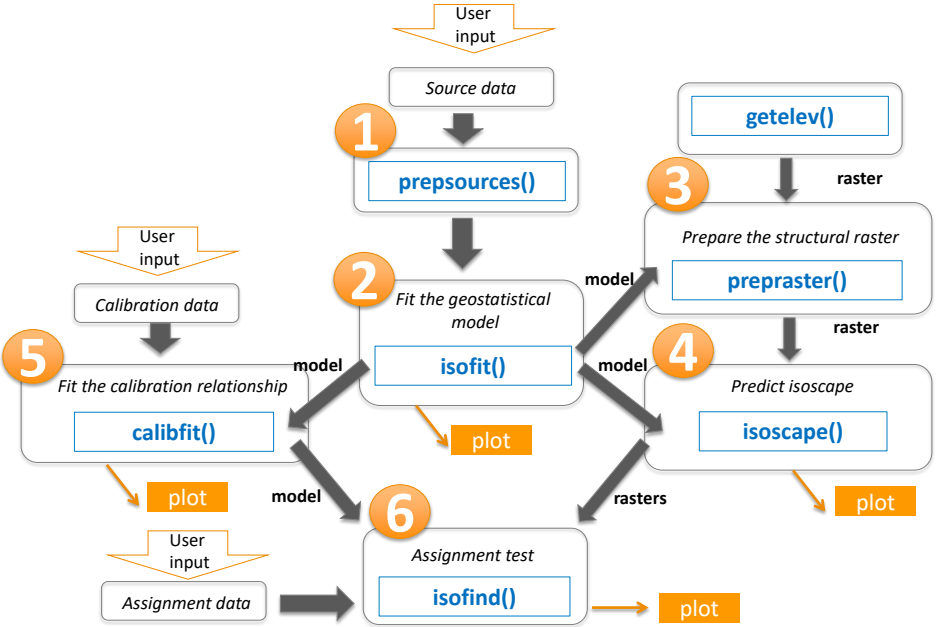
Step 6. Perform the assignment

Uncertainty check list:

- ✓ prediction variance of the fit of the mean model
- ✓ prediction variance of the inverted calibration fit
- ✓ residual variance of the calibration fit
- ×/✓ [1] covariation between predictions from the mean model and the inverted calibration fit

[1] not included as default but included if 'neglect_covPredCalib = FALSE'

Table of contents


User input
Source data
1
prepsources()
getelev()
raster
3
Prepare the structural raster
prepraster()
raster
User input
Calibration data
2
Fit the geostatistical model
isofit()
model
model
5
Fit the calibration relationship
calibfit()
model
plot
4
Predict isoscape
isoscape()
plot
plot
model
6
rasters
User input
Assignment data
Assignment test
isofind()
plot

Where does the bat Nnoc_15 comes from?

An entire workflow can takes as little as 8 lines of code:

```
GNIPDataDEagg <- prepsources(data = GNIPDataDE)

GermanFit <- isofit(data = GNIPDataDEagg, mean_model_fix = list(elev = TRUE, lat_abs = TRUE))

getelev()
ElevDE <- raster("~/elevation_world_z5.tif")
ElevRasterDE <- prepraster(ElevDE, isofit = GermanFit, aggregation_factor = 2)

GermanScape <- isoscape(raster = ElevRasterDE, isofit = GermanFit)

CalibBats <- calibfit(data = CalibDataBat, isofit = GermanFit)

Assigned15 <- isofind(data = subset(AssignDataBat, sample_ID == "Nnoc_15"),
                      isoscape = GermanScape, calibfit = CalibBats)
```

Note: you should always check all intermediary outputs as we did.

Where does the bat Nnoc_15 comes from?

```
plot(Assigned15)
```

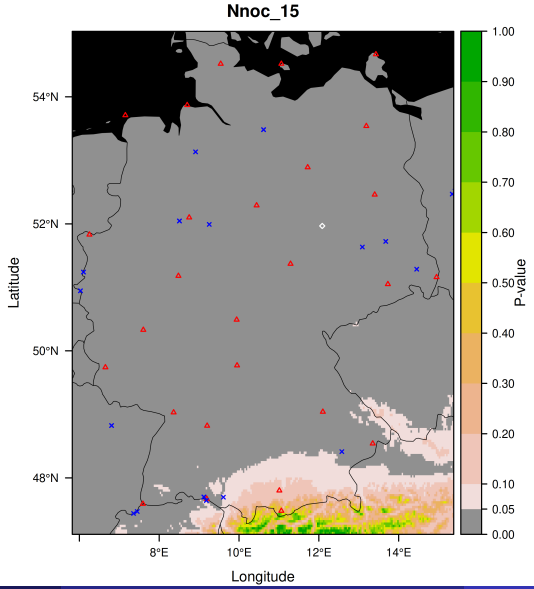


Table of contents

Future features

- More predictors for building isoscapes (e.g. temperature), with data download performed via **R** packages.
- More documentation in the bookdown (`https://bookdown.org/content/782/`).
- Easier functions for weighting by precipitation amount.
- Adding clustering after assignment (with p-value against null model).
- Replacing {`raster`} by {`terra`} or {`stars`}.
- Alternative plots using {`ggplot2`}.
- DHGLM to capture the uncertainty of the residual dispersion model.
- 3D/4D isoscapes?

Future features

- More predictors for building isoscapes (e.g. temperature), with data download performed via **R** packages.
- More documentation in the bookdown (`https://bookdown.org/content/782/`).
- Easier functions for weighting by precipitation amount.
- Adding clustering after assignment (with p-value against null model).
- Replacing {`raster`} by {`terra`} or {`stars`}.
- Alternative plots using {`ggplot2`}.
- DHGLM to capture the uncertainty of the residual dispersion model.
- 3D/4D isoscapes?
- Anything else?

To follow or contribute to the development of IsoriX

- https://github.com/courtiol/IsoriX
- https://groups.google.com/g/IsoriX