c14bazAAR & oxcAAR

Two R packages for the collection, calibration and modelling of 14C dates

---

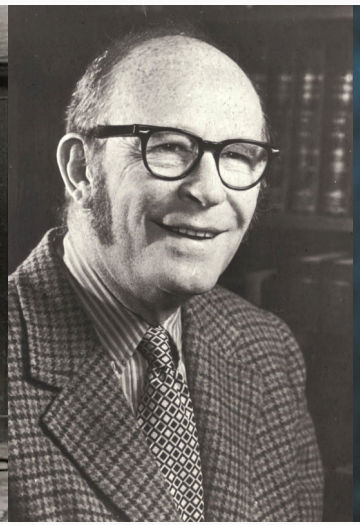
Clemens Schmid, Martin Hinz

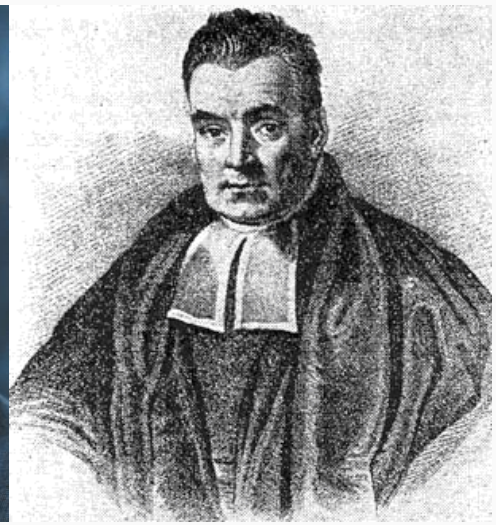
March 2018

Institute of Pre- and Protohistoric Archaeology, Kiel University

^{14}C: The backbone of absolute chronology

- ^{14}C dating with calibration has **revolutionized** the previous typology-based **chronology** in many branches of archaeology
- thanks to **Bayesian statistics**, **Monte-Carlo simulation** and **seriation based models**, **accuracy** now reaches the **level of generations**
- ^{14}C data are **gold standard** today for much of the (pre-)history: almost no archaeologist can work meaningfully at present without using ^{14}C data



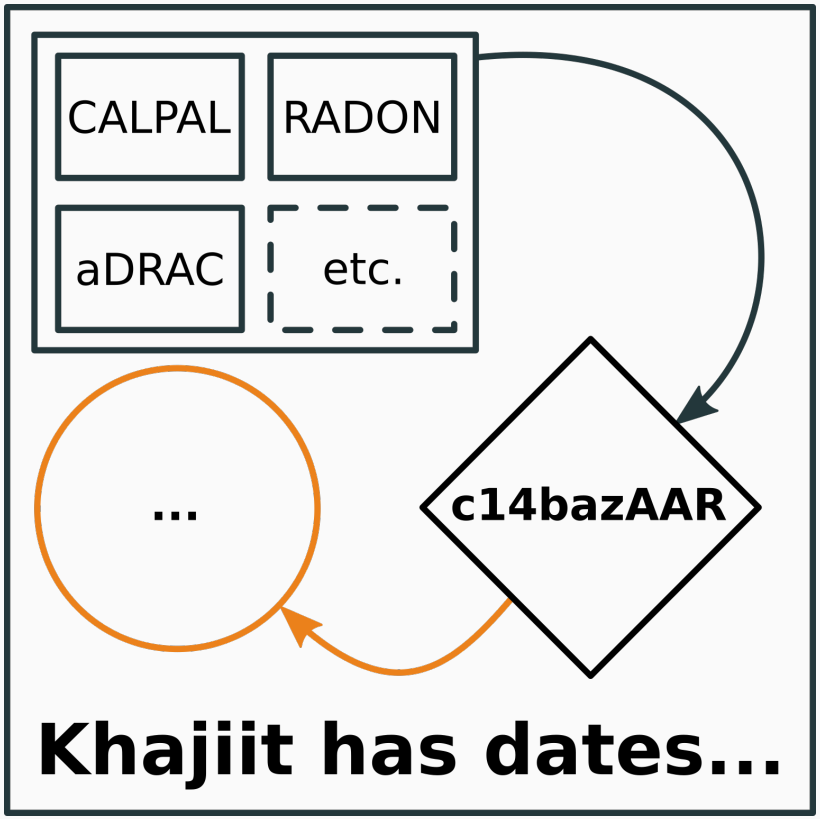
c14bazAAR & oxcAAR

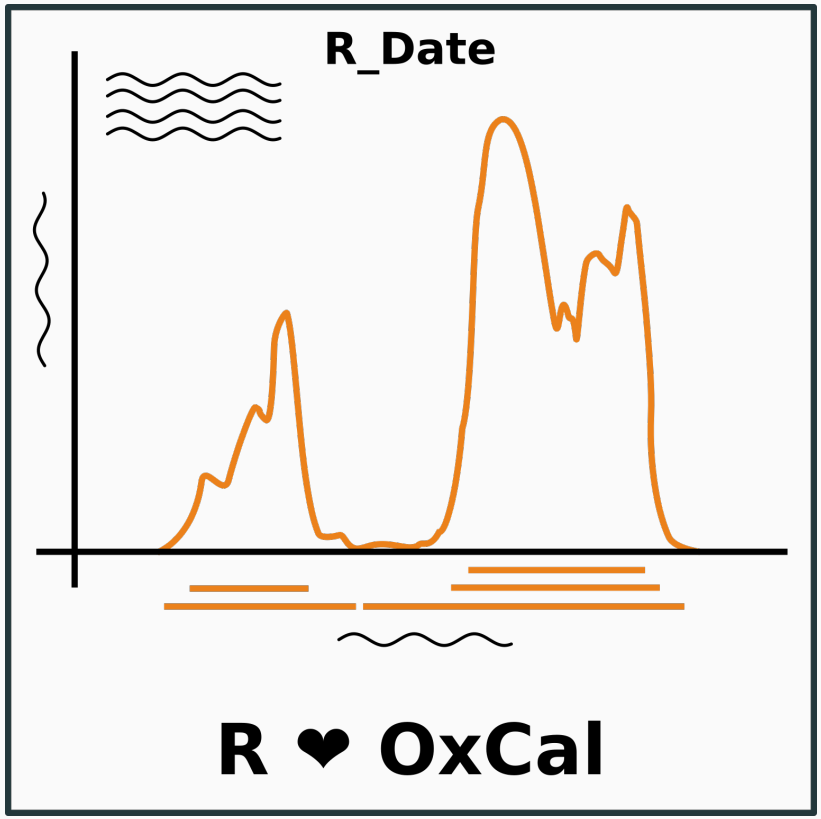
c14bazAAR

R package for **download** and **preparation** of ^{14}C dates from different source databases

oxcAAR

R package API to OxCal for reproducible ^{14}C **calibration**, **sequencing** and **simulation**





c14bazAAR

A thousand databases

14SEA: 14C database for Southeast Europe and Anatolia (10,000–3000 calBC)
aDRAC: Archives des datations radiocarbone d'Afrique centrale
AustArch: Database of 14C and Luminescence Ages from Archaeological Sites in Australia.
Banadora: Banque nationale de données radiocarbone.
CalPal: Radiocarbon Database of the CalPal software package.
CARD: Canadian Archaeological Radiocarbon Database.
CONTEXT: Collection of radiocarbon dates from sites in the Near East and neighboring regions (20.000 - 5.000 calBC).
Euroevol: Cultural Evolution of Neolithic Europe Dataset.
INQUA: Radiocarbon Palaeolithic Europe Database.
Ibercrono: Cronometrías Para la Historia de la Península Ibérica.
KITE East Africa: Radiocarbon dates from eastern Africa in the CARD2.0 format.
PPND: The platform for neolithic radiocarbon dates.
RADON: Central European and Scandinavian database of ^{14}C dates for the Neolithic and Early Bronze Age.
...

Idea

c14bazAAR is a R package to **download**, **merge** and **prepare** ^{14}C dates from **different source databases**

```
# basic workflow
get_dates() %>%              # get dates
  ...() %>%
  calibrate() %>%            # apply some data preparation tools
  ...() %>%
  ...() -> list_of_dates     # get a nice date list for your research
```

- User perspective
  - access many highly different databases with **one interface**
  - **reproducibility** with scripted data selection
  - standard data structure for direct access to powerful R tools (**tidyverse**)
- Developer perspective
  - **Open Source**: examine & improve the implementation and adjust everything for your needs
  - simple **parser** development framework to add further databases
  - embed bulk c14 dates into **your own application**

c14bazAAR provides a ^{14}C dates search engine: **www.neolithicRC.de**

Common data structure: The c14_date_list S3 class

```
c14bazAAR::get_aDRAC()
```

```
##  Radiocarbon date list
##  dates       1258
##  sites       478
##  countries   12
##  uncalBP     46500 - 0
##
## # A tibble: 1,258 x 12
##    sourcedb labnr    c14age c14std c13val site   feature material country
##    <chr>    <chr>     <int>  <int>  <dbl> <chr>  <chr>   <chr>    <chr>
##  1 aDRAC    AA-78447   2362     39     0. Mbaere <NA>    <NA>     CAF
##  2 aDRAC    AA-78448   2171     37     0. Mbaere <NA>    <NA>     CAF
##  3 aDRAC    AA-78449    834     35     0. Mbaere <NA>    <NA>     CAF
##  4 aDRAC    AA-94529    215     34     0. Ngotto BB01-29 <NA>     CAF
##  5 aDRAC    AA-94530    168     35     0. Ngotto BB01-30 <NA>     CAF
##  6 aDRAC    AA-94531    207     35     0. Ngotto BB01-31 <NA>     CAF
##  7 aDRAC    AA-94532    148     34     0. Ngotto BB01-36 <NA>     CAF
##  8 aDRAC    AA-94533    231     34     0. Ngotto BB05-32 <NA>     CAF
##  9 aDRAC    AA-94534    187     34     0. Ngotto BB05-41 <NA>     CAF
## 10 aDRAC    AA-94537    160     35     0. Ngotto ND01-03 <NA>     CAF
## # ... with 1,248 more rows, and 3 more variables: lat <dbl>, lon <dbl>,
## #   shortref <chr>
```

Common data structure: The c14_date_list S3 class

c14_date_list . . .

- . . . is a modified **tibble**/data.frame
- . . . has a set of **predefined variables** – essence of source databases

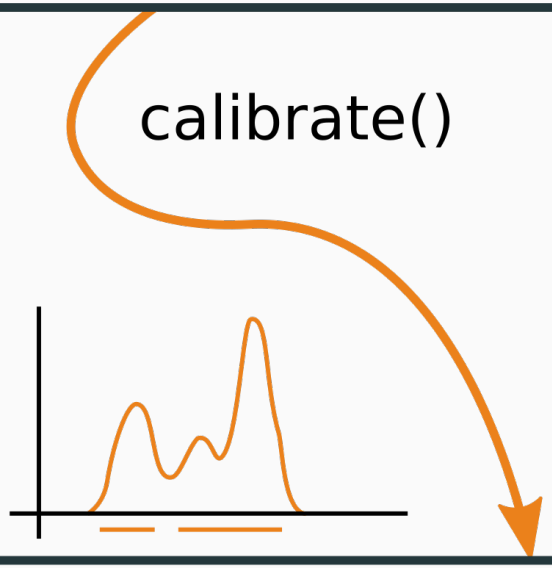
general information	location	archaeological context	sample analysis
labnr	region	period	**c14age**
sourcedb	country	culture	**c14std**
method	lat	sitetype	c13val
shortref	lon	feature	material
comment	site		species

See `c14bazAAR::variable_reference` for the documentation how variables from different databases are related.

- . . . provides **own class methods**

Features – calibration

labnr	c14age	c14std	site	material	country	lat	lon	etc.
AB-123	5530	30	Ingen 5	human	Görmany	48.385	9.121	
CD-456	5340	50	Dorf 4	Charcool	DE	50.778	8.945	
CD-456	5340	50	Dorf IV	glass	FRG	50.778	8.945	



calrange[[1]]

dens	from	to
42.3	6286	6354
51.1	6363	6396

calprobdistr[[1]]

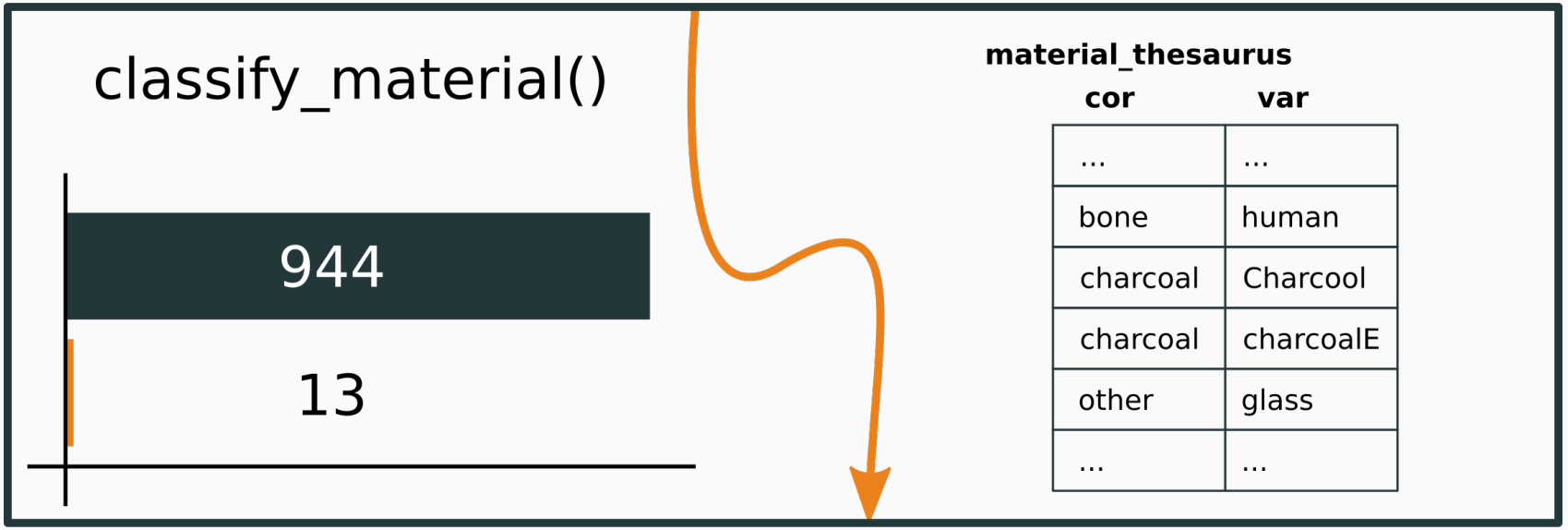
calage	density
...	...
6378	0.01016
6379	0.01036
6380	0.01049
...	...

labnr	c14age	c14std	calprobdistr	calrange	sigma	site	material	etc.
AB-123	5530	30	<data.frame>	<data.frame>	2	Ingen 5	human	
CD-456	5340	50	<data.frame>	<data.frame>	2	Dorf 4	Charcool	
CD-456	5340	50	<data.frame>	<data.frame>	2	Dorf IV	glass	

Figure 1: Bulk calibration with calibrate()

Features – material classification

labnr	c14age	c14std	site	material	country	lat	lon	etc.
AB-123	5530	30	Ingen 5	human	Görmany	48.385	9.121	
CD-456	5340	50	Dorf 4	Charcool	DE	50.778	8.945	
CD-456	5340	50	Dorf IV	glass	FRG	50.778	8.945	



labnr	c14age	c14std	site	material	material_thes	country	lat	etc.
AB-123	5530	30	Ingen 5	human	bone	Görmany	48.385	
CD-456	5340	50	Dorf 4	Charcool	charcoal	DE	50.778	
CD-456	5340	50	Dorf IV	glass	other	FRG	50.778	

Figure 2: Simplification of material information with classify_material()

Features – country correction

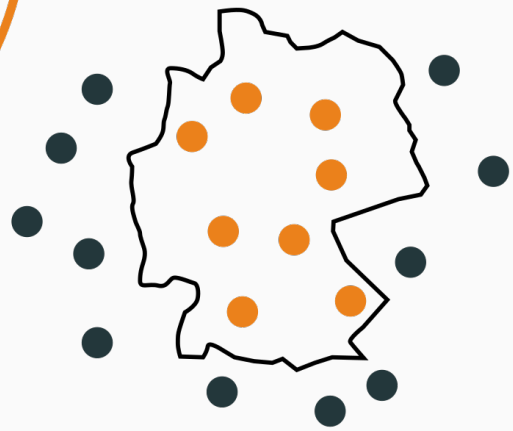
labnr	c14age	c14std	site	material	country	lat	lon	etc.
AB-123	5530	30	Ingen 5	human	Görmany	48.385	9.121	
CD-456	5340	50	Dorf 4	Charcool	DE	50.778	8.945	
CD-456	5340	50	Dorf IV	glass	FRG	50.778	8.945	

countrycode::codelist

iso3c	name.en	name.ar	...
...	...	...	...
GEO	Georgia	جورجيا	...
DEU	Germany	*ألمانيا	...
GHA	Ghana	غانا	...
...	...	...	...

standardize_country_name()

rworldxtra::countriesHigh



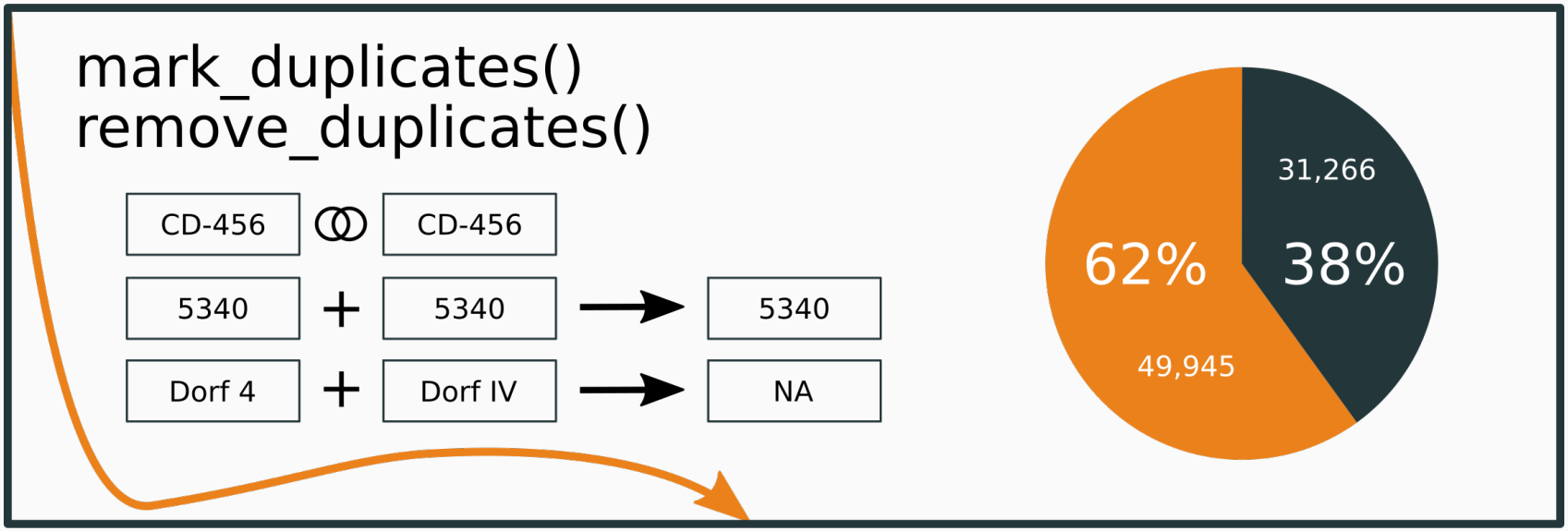
determine_country_by_coordinate()

labnr	c14age	c14std	country	country_coord	country_thes	lat	lon	etc.
AB-123	5530	30	Görmany	Germany	Germany	48.385	9.121	
CD-456	5340	50	DE	Germany	Germany	50.778	8.945	
CD-456	5340	50	FRG	Germany	Germany	50.778	8.945	

Figure 3: Unification of country information with finalize_country()

Features – duplicates

labnr	c14age	c14std	site	material	country	lat	lon	etc.
AB-123	5530	30	Ingen 5	human	Görmany	48.385	9.121	
CD-456	5340	50	Dorf 4	Charcool	DE	50.778	8.945	
CD-456	5340	50	Dorf IV	glass	FRG	50.778	8.945	



labnr	c14age	c14std	site	etc.	duplicate_group	duplicate_remove_log
AB-123	5530	30	Ingen 5		NA	...site: Ingen 5...
CD-456	5340	50	NA		0	...site: Dorf 4 | Dorf IV...
~~CD-456~~	~~5340~~	~~50~~	~~Dorf IV~~		~~0~~	

Figure 4: Finding duplicates with mark_duplicates()

Demotime

```
library(magrittr)

c14bazAAR::get_all_dates() %>%
  # dplyr::sample_n(500) %>%
  # c14bazAAR::as.c14_date_list() %>%
  c14bazAAR::calibrate(
    choices = c("calprobdistr", "calrange")
  ) %>%
  c14bazAAR::classify_material() %>%
  c14bazAAR::finalize_country_name() %>%
  c14bazAAR::mark_duplicates() %>%
  dplyr::arrange(dplyr::desc(c14age))
```

oxcAAR

Several possibilities for calibration

outside of R

- CalPal, BCal, CALIB, Fairbanks calibration, OxCal, iosacal, MatCal
- calibrator
- ChronoModel

within R

-  Andrew Parnell: **Bchron**

```
Bchron::BchronCalibrate(5000, 30, calCurves = "intcal13")
```

-  Andrew Bevan: **rcarbon**

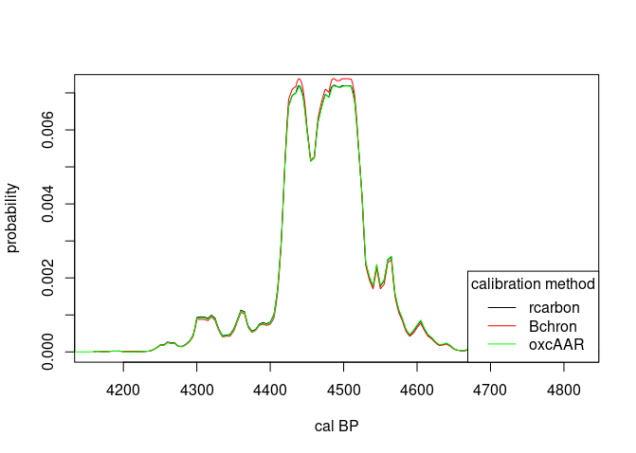
```
rcarbon::calibrate(5000, 30)
```

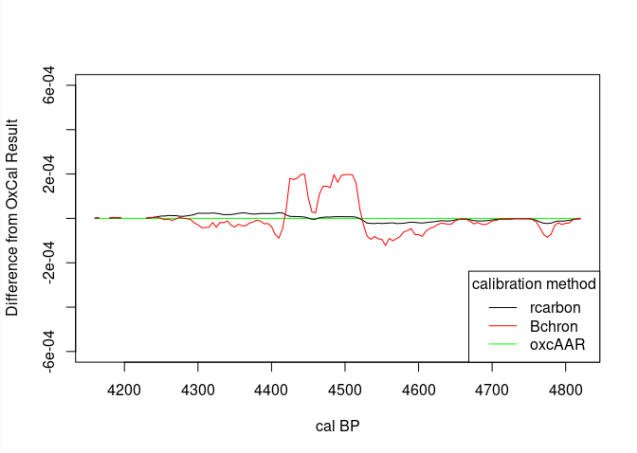
-  ISAAK: **oxcAAR**

```
oxcAAR::oxcalCalibrate(5000, 30)
```

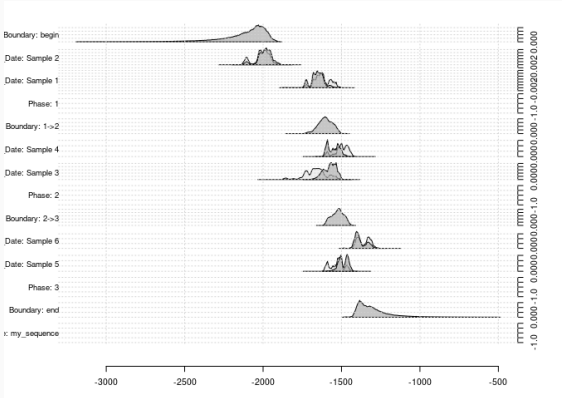
Motivation: Why another calibration package?

1. Every package produces **slightly different results**





2. **Sequential calibration** with other tools (currently) not available



Idea

oxcAAR is a R package to **calibrate**, **simulate** and **sum** ^{14}C dates using **OxCal** as calibration engine

```
quickSetupOxcal()               # download Oxcal to tempory folder
```

```
## NULL
```

```
oxcalCalibrate( bp=5000, std=20) # do the calibration
```

```
##
## ============================
##  R_Date: 1
## ============================
##
##
## BP = 5000, std = 20
##
##
## unmodelled:
##
## one sigma
## 3794 BC - 3759 BC (43.8%)
## 3740 BC - 3712 BC (24.4%)
##
## two sigma
## 3910 BC - 3876 BC (12.1%)
## 3802 BC - 3706 BC (83.3%)
##
## three sigma
## 3939 BC - 3858 BC (15.5%)
## 3813 BC - 3698 BC (84.2%)
##
## Calibrated with:
##    IntCal13 atmospheric curve (Reimer et al 2013)
```

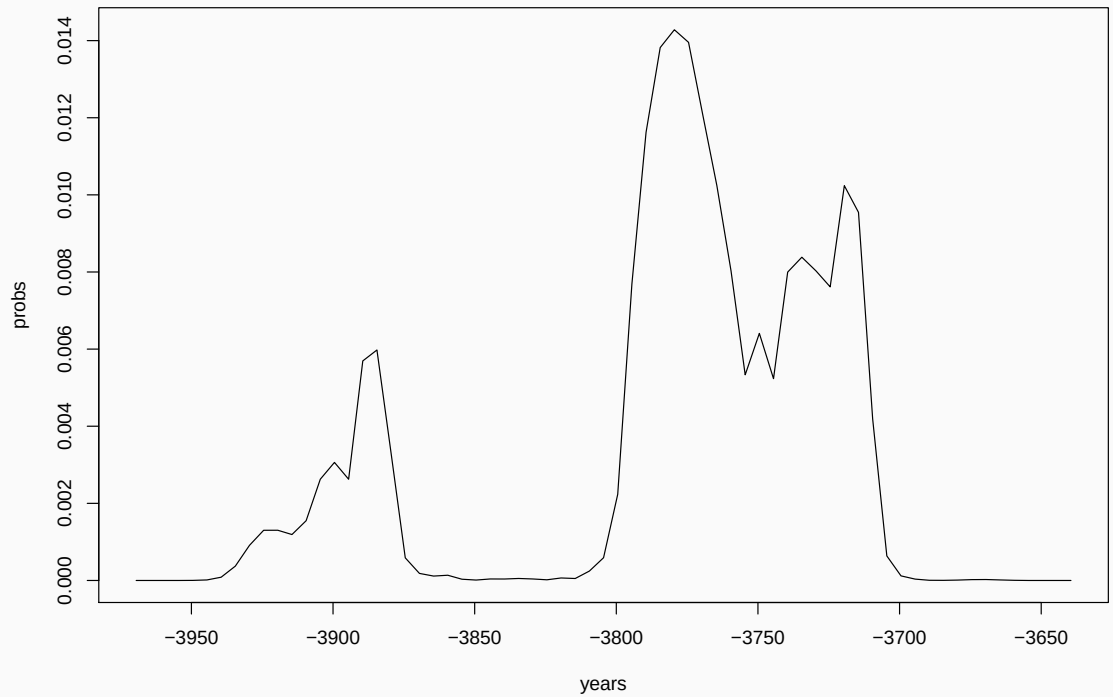
Accessing the calibration result - structure

```
my_cal_date <- oxcalCalibrate( bp=5000, std=20)
str(my_cal_date, max.level = 3)
```

```
## List of 1
##  $ 1:List of 9
##   ..$ name                    : chr "1"
##   ..$ type                    : chr "R_Date"
##   ..$ bp                      : int 5000
##   ..$ std                     : int 20
##   ..$ cal_curve               :List of 5
##   .. ..$ name      : chr " IntCal13 atmospheric curve (Reimer et al 2013)"
##   .. ..$ resolution: num 5
##   .. ..$ bp        : num [1:10001] 46401 46396 46391 46386 46381 ...
##   .. ..$ bc        : num [1:10001] -48050 -48044 -48040 -48034 -48030 ...
##   .. ..$ sigma     : num [1:10001] 274 274 274 273 273 ...
##   ..$ sigma_ranges            :List of 3
##   .. ..$ one_sigma  :'data.frame':   2 obs. of  3 variables:
##   .. ..$ two_sigma  :'data.frame':   2 obs. of  3 variables:
##   .. ..$ three_sigma:'data.frame':   2 obs. of  3 variables:
##   ..$ raw_probabilities       :'data.frame':  67 obs. of  2 variables:
##   .. ..$ dates        : num [1:67] -3970 -3964 -3960 -3954 -3950 ...
##   .. ..$ probabilities: num [1:67] 0.00 0.00 0.00 1.43e-07 1.94e-06 ...
##   ..$ posterior_sigma_ranges :List of 3
##   .. ..$ one_sigma  : logi NA
##   .. ..$ two_sigma  : logi NA
##   .. ..$ three_sigma: logi NA
##   ..$ posterior_probabilities: logi NA
##   ..- attr(*, "class")= chr "oxcAARCalibratedDate"
##  - attr(*, "class")= chr [1:2] "list" "oxcAARCalibratedDatesList"
```

Accessing the calibration result - for basic plot

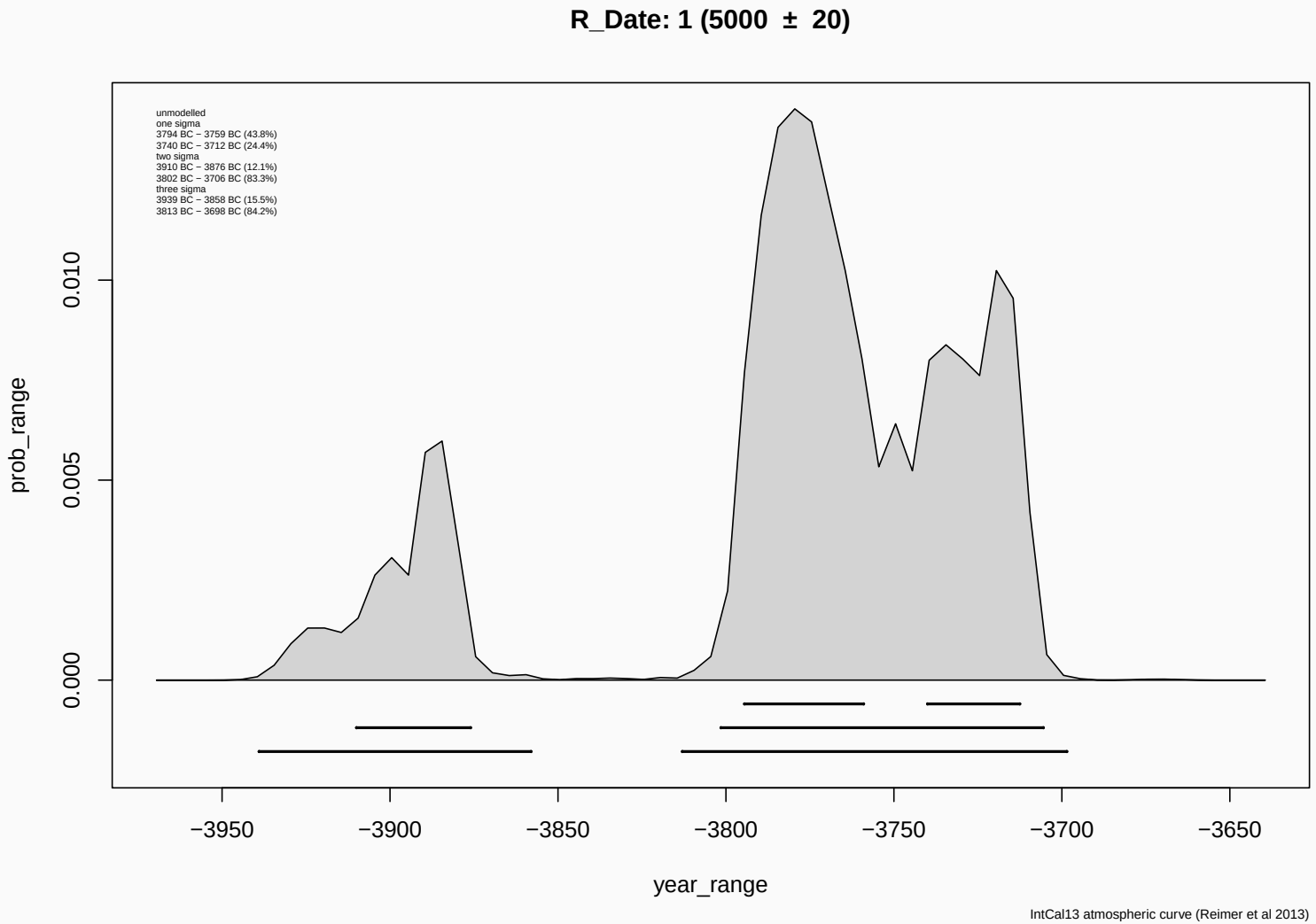
```r
plot(
  my_cal_date$`1`$raw_probabilities$dates,
  my_cal_date$`1`$raw_probabilities$probabilities,
  type = "l", xlab = "years", ylab = "probs"
)
```



Features - plotting an individual date

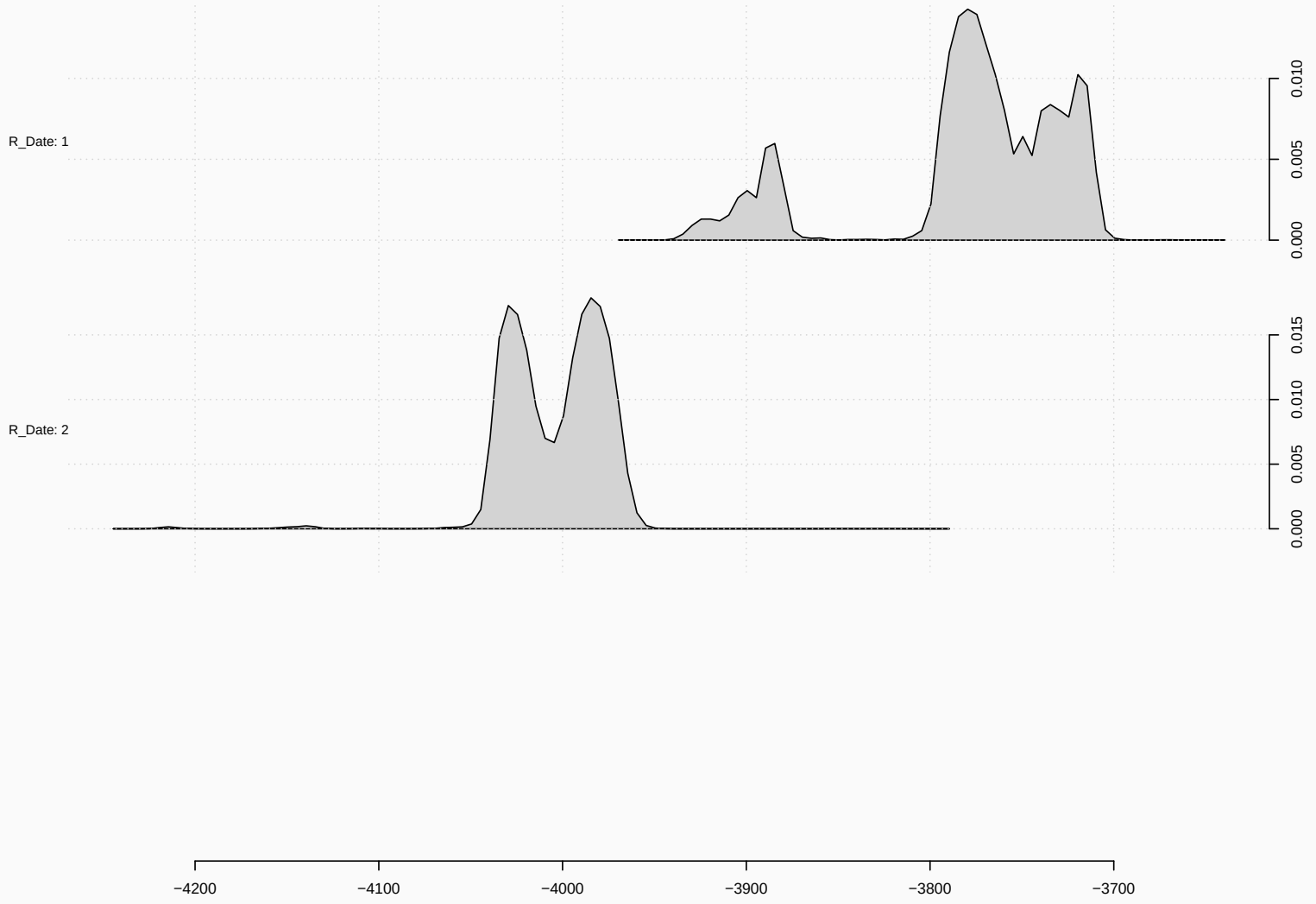
Plotting an individual date

```
oxcalCalibrate( bp=5000, std=20) %>% plot()
```



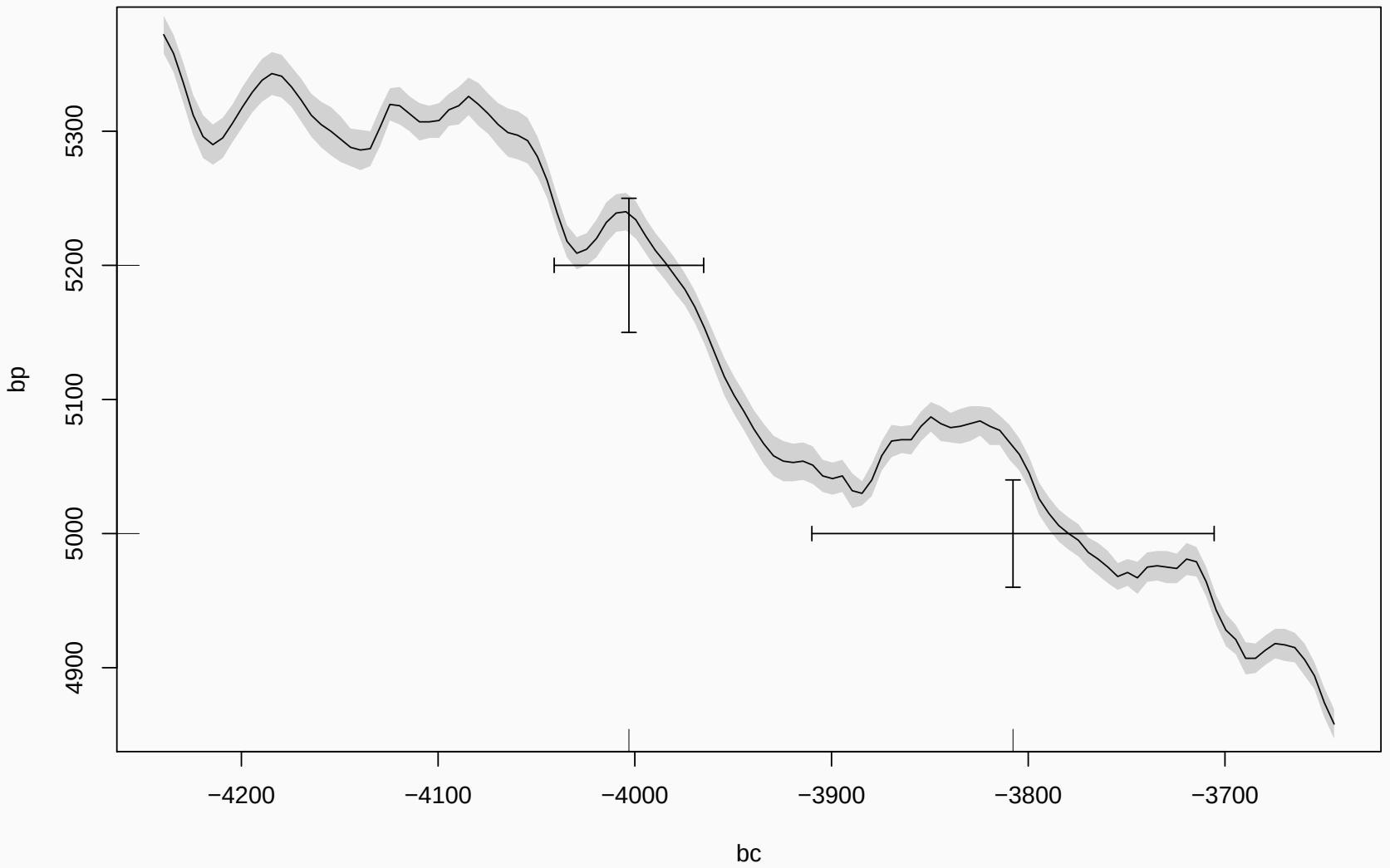
Features - plotting multiple dates

```
oxcalCalibrate( bp=c(5000,5200), std=c(20,25)) %>% plot()
```



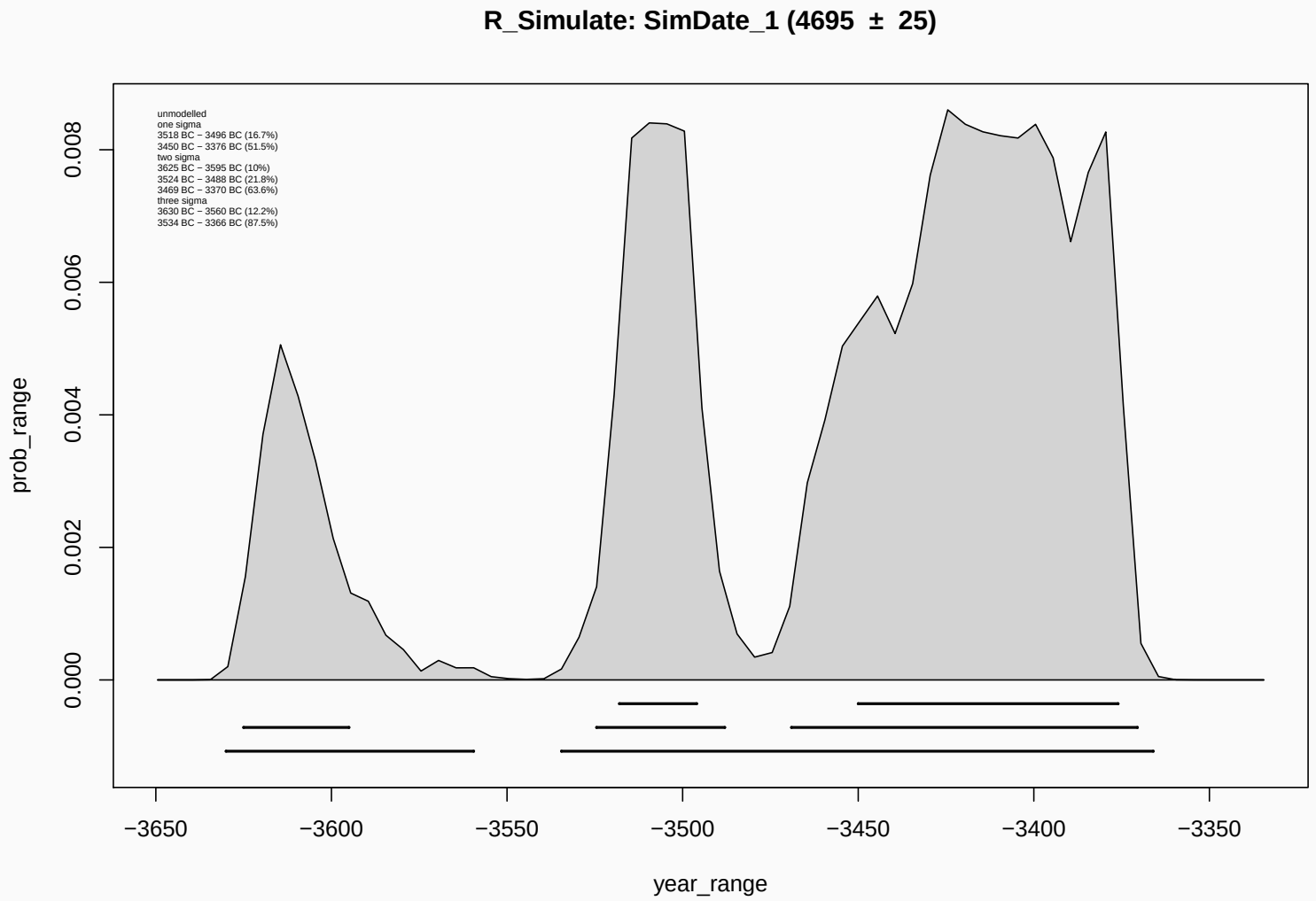
Features - plotting on the calibration curve

```
oxcalCalibrate( bp=c(5000,5200), std=c(20,25)) %>% calcurve_plot()
```



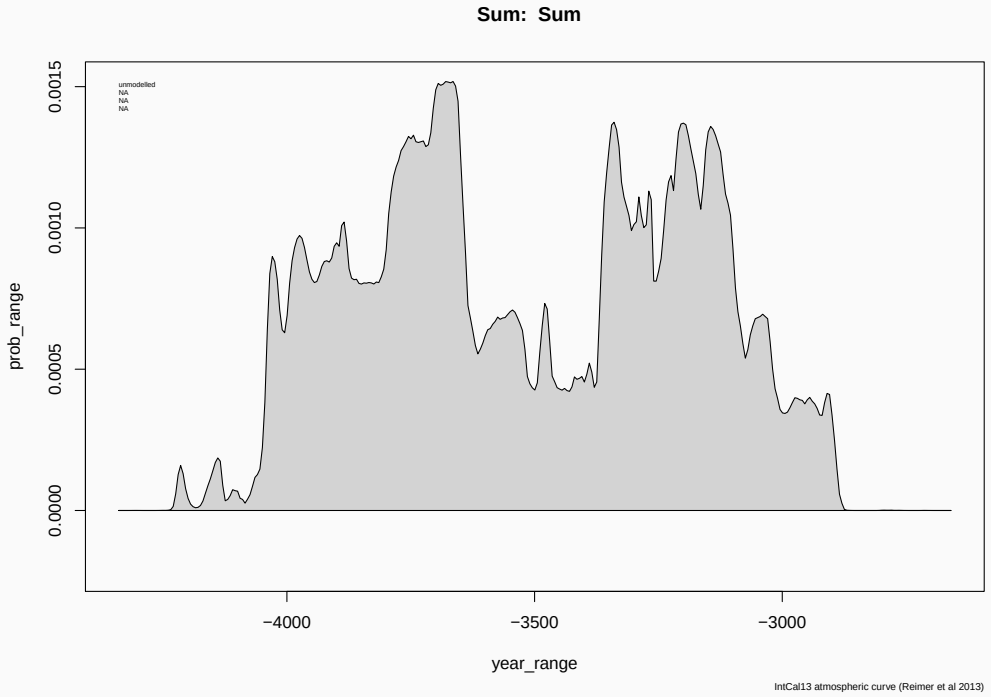
Features - simulating a R_Date

```
oxcalSimulate(-3400, 25, "SimDate_1") %>% plot()
```



Features - simulating a sum calibration

```
oxcalSumSim(
  timeframe_begin = -4000,       # From when
  timeframe_end = -3000,         # To when
  n = 50,                        # Number of dates
  stds = 35,                     # Standard deviation of dates (can also be a vector of length n)
  date_distribution = "uniform" # random uniform; alternatively: equidist for equidistant dates
) %>% plot()
```



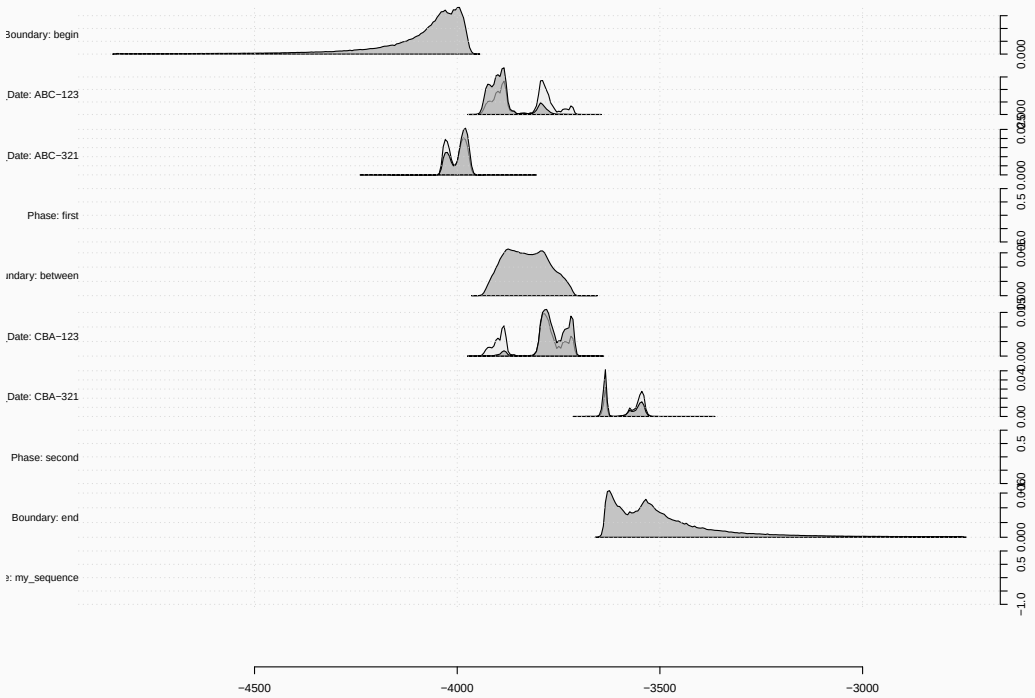
Features - injecting arbitray OxCal code

```
my_sequence <- Sequence(
  c(
    Boundary("begin"),
    Phase(
      R_Date(
        c("ABC-123", "ABC-321"),
        c(5020, 5200),
        20), "first"
    ),
    Boundary("between"),
    Phase(
      R_Date(
        c("CBA-123","CBA-321"),
        c(5010, 4810),
        20), "second"
    ),
    Boundary("end")
  ), "my_sequence"
)
cat(my_sequence)
```

```
## Sequence("my_sequence")
## {Boundary("begin");
## Phase("first"){
## R_Date("ABC-123", 5020, 20);
## R_Date("ABC-321", 5200, 20);};
## Boundary("between");
## Phase("second"){
## R_Date("CBA-123", 5010, 20);
## R_Date("CBA-321", 4810, 20);};
## Boundary("end");};
```

Features - using Bayesian calibration of OxCal

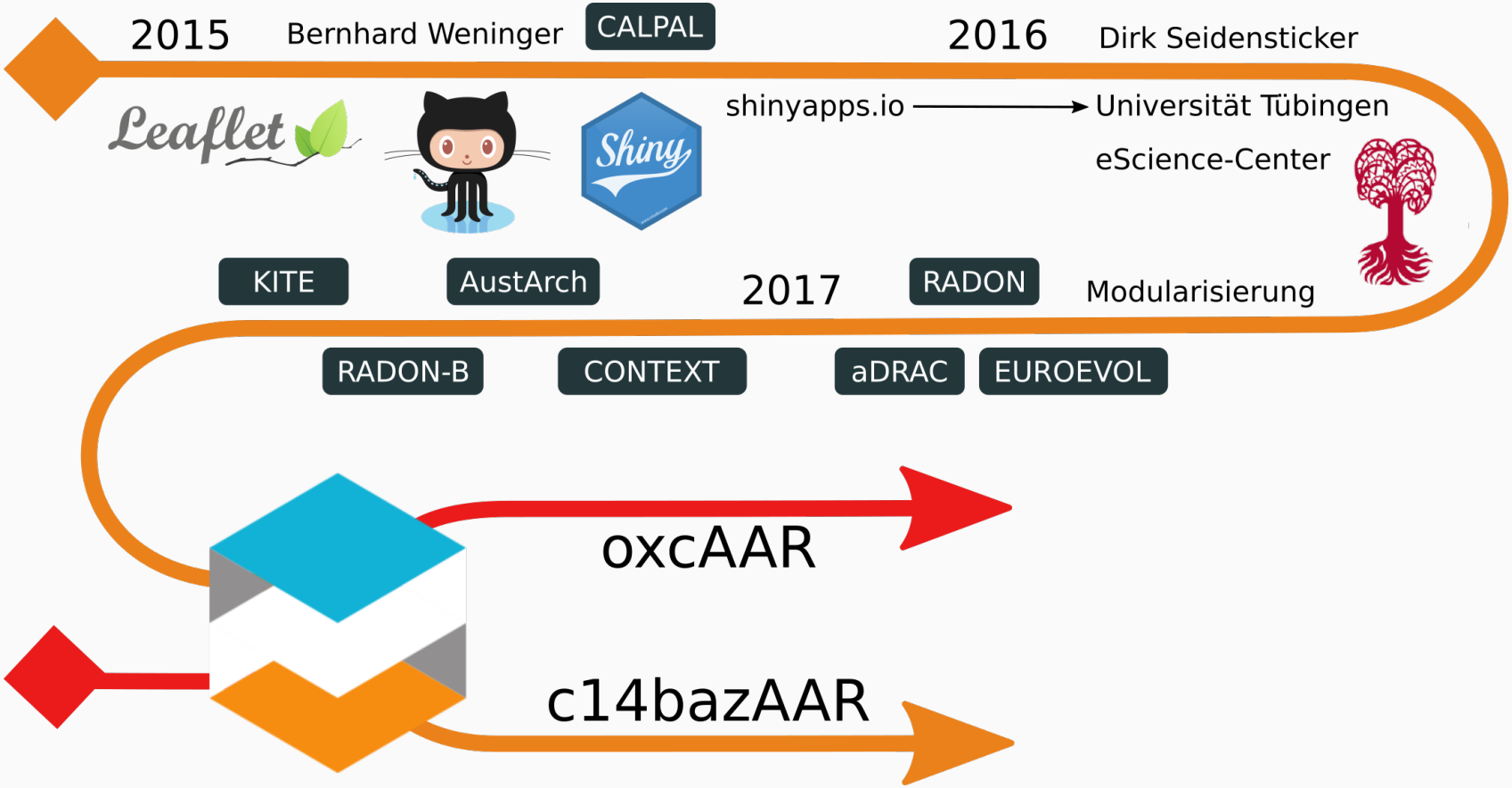
```
my_result_data  <- my_sequence %>%
  executeOxcalScript() %>%
  readOxcalOutput() %>%
  parseOxcalOutput(only.R_Date = F)

plot(my_result_data)
```



State and Roadmap

- c14bazAAR
  - Ready for takeoff at Github
    - https://github.com/ISAAKiel/c14bazAAR
    - CRAN release within this month
- oxcAAR
  - oxcAAR 1.0.0 ‘Nicola’ released at CRAN
  - Bayesian Calibration available at Github
    - https://github.com/ISAAKiel/oxcAAR
  - oxcAAR 1.1.0 ‘Flora’
    - CRAN release within this month

Projekt Timeline



Clemens Schmid
clemens@nevrome.de

Martin Hinz
martin.hinz@ufg.uni-kiel.de