User Manual

BayClump v. 0.9

Updated June 11th, 2021

Table of Contents

Contents

Publication Information

C. Román-Palacios, H. M. Carroll, A. J. Arnold, R. Ulrich, R. J. Flores, D. Brown, K. A. McKinnon, and A. Tripati. Bayclump: Bayesian Calibration Models of the 'Clumped Isotopes' Paleothermometer. In preparation for submission to *Paleoceanography and Paleoclimatology*.
(Hyperlink to the paper here)

Contact Information

Developers

Hannah M. Carroll hcarroll@epss.ucla.edu
Cristian Román-Palacios cromanpa94@gmail.com

Senior author

Aradhna Tripati atripati@g.ucla.edu

GitHub repo

https://github.com/Tripati-Lab/BayClump

Bug reports and issues

https://github.com/Tripati-Lab/BayClump/issues

BayClump Basics

BayClump version 0.9 was developed in R version 4.1.0 and RStudio version 1.4.1106. This app accompanies Román-Palacios et al. (xxxx): BayClump: Bayesian Calibration Models of the 'Clumped Isotope' Paleothermometer.

BayClump is separated into calibration and reconstruction workflows. Any calibration models selected on the Calibrations tab are automatically made available for use in the Reconstructions tab. At this time, only Δ_{47}-based calibrations and reconstructions are supported.

Default calibration data

Two calibration datasets are included in BayClump by default. Model 1 is based on the reprocessed datasets of Petersen et al. (2019). Model 2 is based on the reprocessed and original datasets of Anderson et al. (2021). The datasets from Petersen et al. (2019) have been reprojected into the I-CDES reference frame after Bernasconi et al. (2021) at a growth temperature of 90°C, so as to be compatible with the datasets of Anderson et al. (2021). Throughout, we refer to this as I-CDES 90.

The default calibration datasets may be used individually or in combination to create calibration models, and may also be combined with the user's own calibration data in the I-CDES 90 reference frame. The user may also choose to work exclusively with their own calibration data in I-CDES 90 or another reference frame if desired. At this time, BayClump does not provide the option to reproject the default calibration datasets into a different reference frame.

Uploading calibration data

BayClump provides a template for users to upload calibration data. It can be downloaded from the Calibrations tab. Sample data are shown in the first ten rows of the template and should be overwritten. **The template must be left in .csv format, although the user is free to change the name of the spreadsheet as needed. Do not change column headings or orders.** Do not enter extra columns - these will not be used by the app and may create problems.

	A	B	C	D	E	F	G	H
1	Sample Name	Mineralogy	Material	N	D47	D47_SD	Temperature	Temp_Error
2	Leeds 41	Dolomite	Synthetic	6	0.348	0.025	220	2
3	Leeds 84	Dolomite	Synthetic	12	0.311	0.018	220	2
4	Leeds 85	Dolomite	Synthetic	16	0.316	0.028	220	2
5	Leeds 86	Dolomite	Synthetic	11	0.32	0.017	220	2
6	Ab-1	Aragonite	Natural	3	0.7679185	0.01041691	-1	2
7	Ac-1_B-12	Aragonite	Natural	6	0.77348394	0.00278529	-1	2
8	Ac-1_B-14	Aragonite	Natural	3	0.77808716	0.00535896	2	2
9	Ac-2	Aragonite	Natural	3	0.74983746	0.0111864	2	2
10	Brach-m1	Calcite	Natural	3	0.75385961	0.0024896	5.5	1
11	Brach-m2	Calcite	Natural	3	0.76021853	0.01489061	5.5	1
12								
13								

Figure 1: The BayClump calibration data template

Sample Name: This should be a unique identifier for each sample. Combinations of letters, numbers, spaces, and special characters may be used. It is best to avoid number signs (#) as this may cause the code behind BayClump to malfunction.

Mineralogy: This should be the full, capitalized name, i.e., Aragonite, Calcite, Dolomite, Mixed, Unknown, etc.

Material: Natural or Synthetic

N: The number of replicates measured to produce the final value. This must be numeric - no letters or special characters are accepted.

D47: The final Δ_{47} value of the sample after completing all appropriate quality assurance steps. This must be numeric - no letters or special characters except for periods are accepted. Provide as many decimal places as possible to improve calculations.

D47_SD: The final standard deviation around the Δ_{47} value after completing all appropriate quality assurance steps. This must be numeric - no letters or special characters except for periods are accepted. Provide as many decimal places as possible to improve calculations.

Temperature: Growth temperature in degrees Celsius (°C). This must be numeric - no letters or special characters except for periods are accepted.

Temp_Error: Absolute error of the growth temperature in degrees Celsius (°C). This must be numeric - no letters or special characters except for periods are accepted.

Selecting calibration options

Scale data
Daeron
Uncertainties

Calibration Model Setup and Selection

BayClump offers five calibration model options. Bootstrapping is used to produce robust estimates of regression coefficients. The user is able to select 50, 100, 500, or 1,000 bootstrap replicates. We recommend that the number of bootstrap replicates chosen be at least 10 times the number of unique samples, with a minimum of 50 and a maximum of 1,000 offered. For example, if you have ten unique samples you wish to use in calibration models, you should choose at least 100 bootstrap replicates (10 x 10). If you have 3 or fewer samples, do not exceed 50 bootstrap replicates.

Calibration data should be entered in degrees Celsius. BayClump will convert the temperature to Kelvin and perform analyses using $10^6/T^2$.

Calibration model descriptions

Note: The descriptions provided below are taken verbatim from the publication this app accompanies. For additional information, please see Román-Palacios et al. (xxxx).

Simple linear model

> The simple linear regression does not explicitly account for measurement error in $10^6/T^2$ or Δ_{47}. Therefore, this model assumes that measurements in both $10^6/T^2$ or Δ_{47} are not imperfect. We fit the simple linear regression model using the lm function in the stats R package…under default parameters. Note that the approach implemented in the lm function in R is based on least squares. Under least squares, the resulting regression model is selected as the line that minimizes the sum of squared error (*i.e.*, sum over the squared differences between the points and the line) in the relationship between $10^6/T^2$ or Δ_{47}.

Inverse weighted linear model

[The inverse weighted linear model is] a simple linear model, also under least squares, with observations being weighted based on the inverse of their squared uncertainty in the measured Δ_{47}. This model is hereafter referred to as the weighted simple linear model. Note that, although this approach indirectly accounts for uncertainty in Δ_{47}, weighted models analyzed in this study do not explicitly account for measurement error in $10^6/T^2$. The weighted simple linear regression was fit using the lm function in the stats R package. The weights argument was set to the $1/(standarderror)^2$ of each observation. Therefore, observations with higher uncertainty (standard error) had less importance at estimating the error of alternative proposed lines during the least square optimization of the model.

York regression

The York model [is fit] using the york function in the IsoplotR R package version 3.4 (Vermeesch 2018). Note that this approach is strongly based on the initial ideas behind the Deming regression model... However, under the York model, the ratio of the weights (related to their uncertainty) in Δ_{47} and $10^6/T^2$ varies over data points (Martin 2000). Therefore, in our paper, the York model is the second most simple model to explicitly account for measurement error in both Δ_{47} and $10^6/T^2$. We specified observations in Δ_{47} and $10^6/T^2$, along with the corresponding standard error in each observation, when fitting this regression model.

Deming regression

...A Deming regression [is fit] using the deming R package version 1.4 (Therneau 2018). In our paper, the Deming regression model is the simplest model that explicitly accounts for measurement error in both Δ_{47} and $10^6/T^2$. Under the Deming regression, the ratio of the weights (related to their uncertainty) in Δ_{47} and $10^6/T^2$ is constant over all data points (Martin 2000). To fit this model, we specified observations in Δ_{47} and $10^6/T^2$, along with the corresponding standard error in each observation. Note that although the Deming model also fits under least squares, the optimization steps while fitting the straight line follows the algorithm of Deming (1943).

Bayesian linear model

...A Bayesian simple linear regression model [is provided]. The Bayesian model is also the simplest Bayesian model fit in the study. This first Bayesian model is equivalent to the simple linear regression model presented above. However, instead of parameter estimates based on least squares, optimization of regression parameters is conducted under a Bayesian framework. Below, we present the mathematical definition of this model.

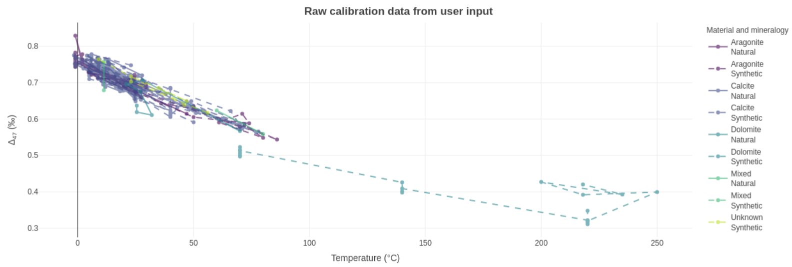
$\Delta_{47_i} \sim Normal(\mu_i, \sigma^2)$
$\mu_i = \alpha + \beta \frac{10^6}{T^2}_i$
$\alpha \sim Normal(0, 10^3)$
$\beta \sim Normal(0, 10^3)$
$\sigma^2 \sim Uniform(0, 10^3)$
$i = 1, ..., N$

Note that we specified uninformative priors above. However, based on published calibrations of the clumped isotope paleothermometer, we also run alternative definitions of the models that assumed normal or uniform priors on:

$\alpha \sim Normal(0.231, 0.065); \beta \sim Normal(0.039, 0.004)$
$\alpha \sim Uniform(0.068, 0.323); \beta \sim Uniform(0.033, 0.046)$

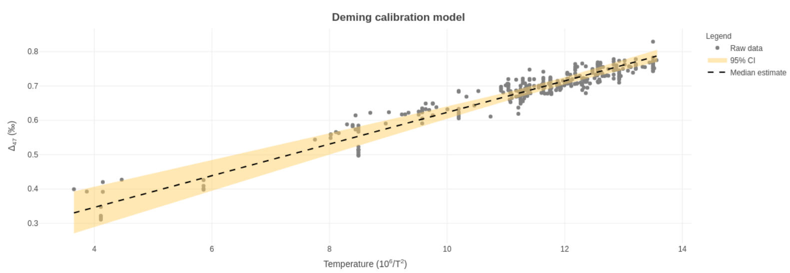
The Calibration Plots Tab

All plots are created in plotly version 4.9.3 (Sievert 2020). They are fully interactive and downloadable. Plots download at web viewing resolution (72 dpi), 1244x400 pixels, and in .png format by default, but may be scaled and resized in external software if desired.



The Calibration Plots tab will display raw calibration data based on user input. The default display is of calibration data for Model 1. It will automatically update to reflect the choice of calibration data made on the Calibrations tab.

Once calibration model runs are complete, the Calibration Plots tab will update to show plots of raw data on the $10^6/T^2$ temperature scale, with overlaid 95% confidence intervals (credible intervals in the case of the Bayesian model), and a median line.



The Reconstructions Tab

Uploading reconstruction data

BayClump provides a template for users to upload reconstruction data. It can be downloaded from the Reconstructions tab. Sample data are shown in the first ten rows of the template and should be overwritten. **The template must be left in .csv format, although the user is free to change the name of the spreadsheet as needed. Do not change column headings or orders.** Do not enter extra columns - these will not be used by the app and may create problems.