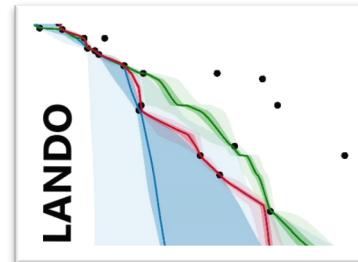


Using LANDO as a universal wrapper for applying multiple age-depth modeling systems for sediment records from Arctic lake systems

Gregor Pfalz, B. Diekmann, J.-C. Freytag, L. Syrykh, D. A. Subetto, and B. K. Biskaborn



The Research Question

How can we reduce the effort to apply all age-depth modelling methods and make results comparable?

LANDO: Linked Age And Depth Modelling

Reference:

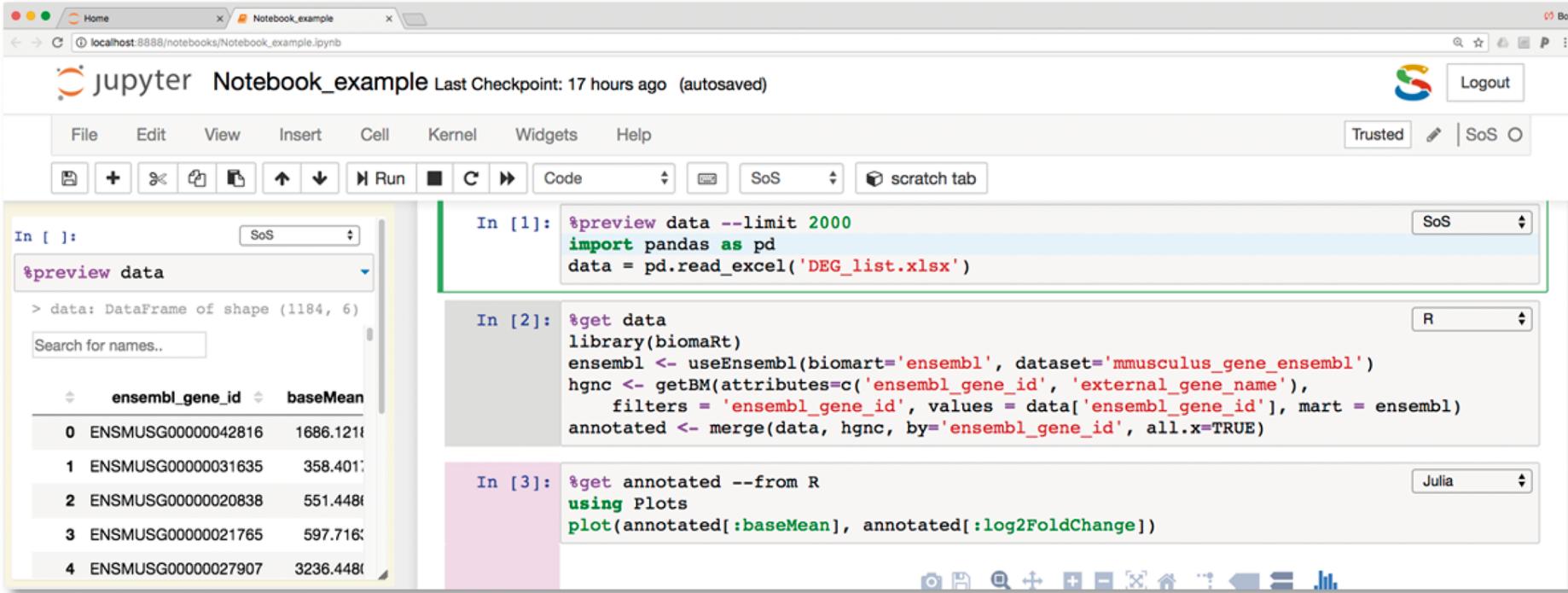
Pfalz, G., Diekmann, B., Freytag, J.-C., Syrykh, L., Subetto, D. A., and Biskaborn, B. K.: Improving age–depth relationships by using the LANDO (“Linked age and depth modeling”) model ensemble, Geochronology, 4, 269–295, <https://doi.org/10.5194/gchron-4-269-2022>, 2022.



The Backbone



ScriptOfScript notebook



The screenshot shows a Jupyter Notebook interface titled "Notebook_example". The top bar includes the Jupyter logo, the notebook name, and a "Logout" button. The menu bar has options: File, Edit, View, Insert, Cell, Kernel, Widgets, Help. Below the menu is a toolbar with icons for file operations like New, Open, Save, Run, Cell, Kernel, Help, and SoS. A "Trusted" button is also present.

The notebook contains three code cells:

- In [1]:**

```
%preview data --limit 2000
import pandas as pd
data = pd.read_excel('DEG_list.xlsx')
```
- In [2]:**

```
%get data
library(biomart)
ensembl <- useEnsembl(biomart='ensembl', dataset='mmusculus_gene_ensembl')
hgnc <- getBM(attributes=c('ensembl_gene_id', 'external_gene_name'),
               filters = 'ensembl_gene_id', values = data['ensembl_gene_id'], mart = ensembl)
annotated <- merge(data, hgnc, by='ensembl_gene_id', all.x=TRUE)
```
- In [3]:**

```
%get annotated --from R
using Plots
plot(annotated[:baseMean], annotated[:log2FoldChange])
```

To the left of the cells, there is a sidebar with a search bar and a table preview for the first cell's output:

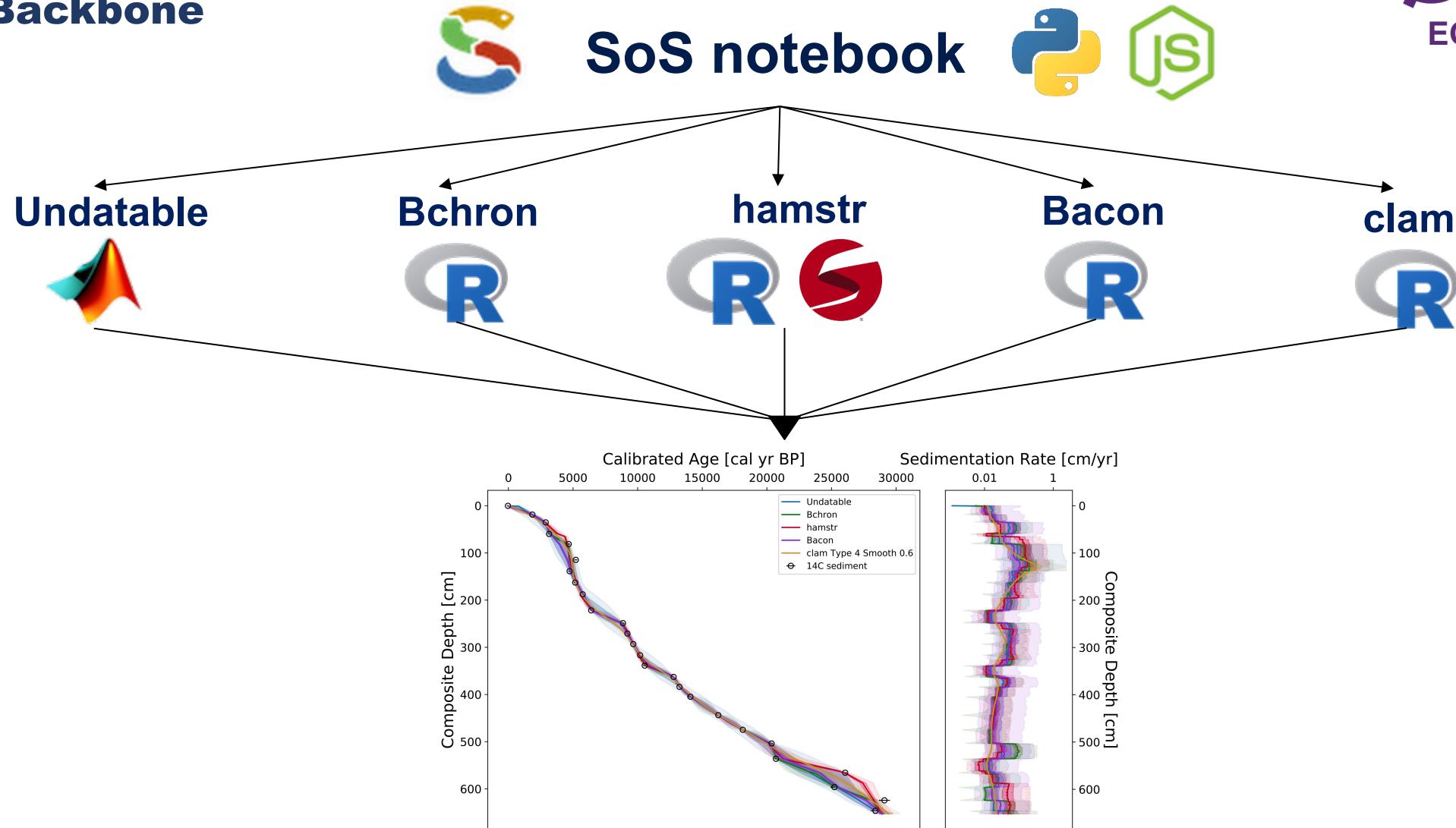
	ensembl_gene_id	baseMean
0	ENSMUSG00000042816	1686.121
1	ENSMUSG00000031635	358.401
2	ENSMUSG00000020838	551.448
3	ENSMUSG00000021765	597.716
4	ENSMUSG00000027907	3236.448



Peng, Bo et al. (2018). SoS notebook: An interactive multi-language data analysis environment. *Bioinformatics*, 34(21), 3768–3770. <https://doi.org/10.1093/bioinformatics/bty405>



The Backbone



The Components

Input & Preparation

Execution & Result Aggregation

- Input via spreadsheet or PostgreSQL database
 - must follow predetermined format of necessary attributes
- LANDO automatically sets calibration curve based on selected material category – users can change to different calibration curve (IntCal20, Marine20, SHCal20)

- LANDO can approximate the reservoir effect from existing radiocarbon data by using hamstr package
- LANDO automatically prepares dataset for each modeling software



The Components

Input & Preparation

- Input via spreadsheet or PostgreSQL database
 - must follow predetermined format of necessary attributes
- LANDO automatically sets calibration curve based on selected material category – users can change to different calibration curve (IntCal20, Marine20, SHCal20)

Execution & Result Aggregation

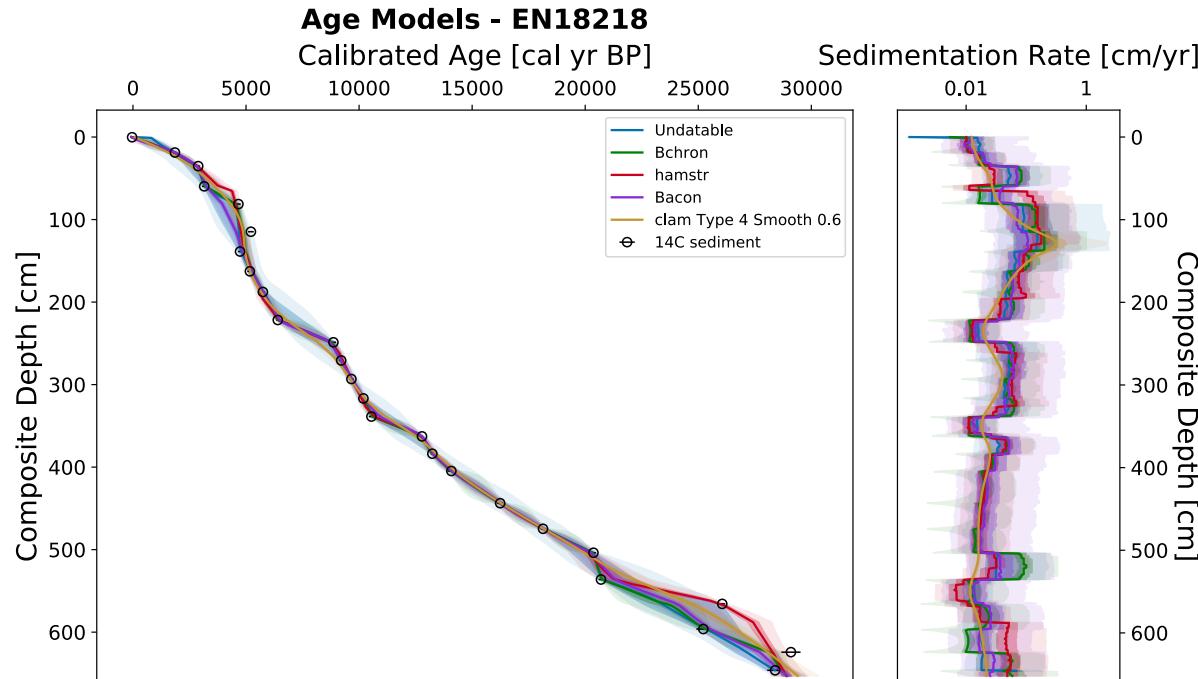
- LANDO executes multiple sediment cores in parallel using R and Python parallelization libraries
- LANDO uses default setting of every modeling software, which users can adapt parameters before modeling process
- LANDO calculates 1σ and 2σ age ranges and sedimentation rates for every modeling software from 10,000 converged age estimates

- LANDO can approximate the reservoir effect from existing radiocarbon data by using hamstr package
- LANDO automatically prepares dataset for each modeling software

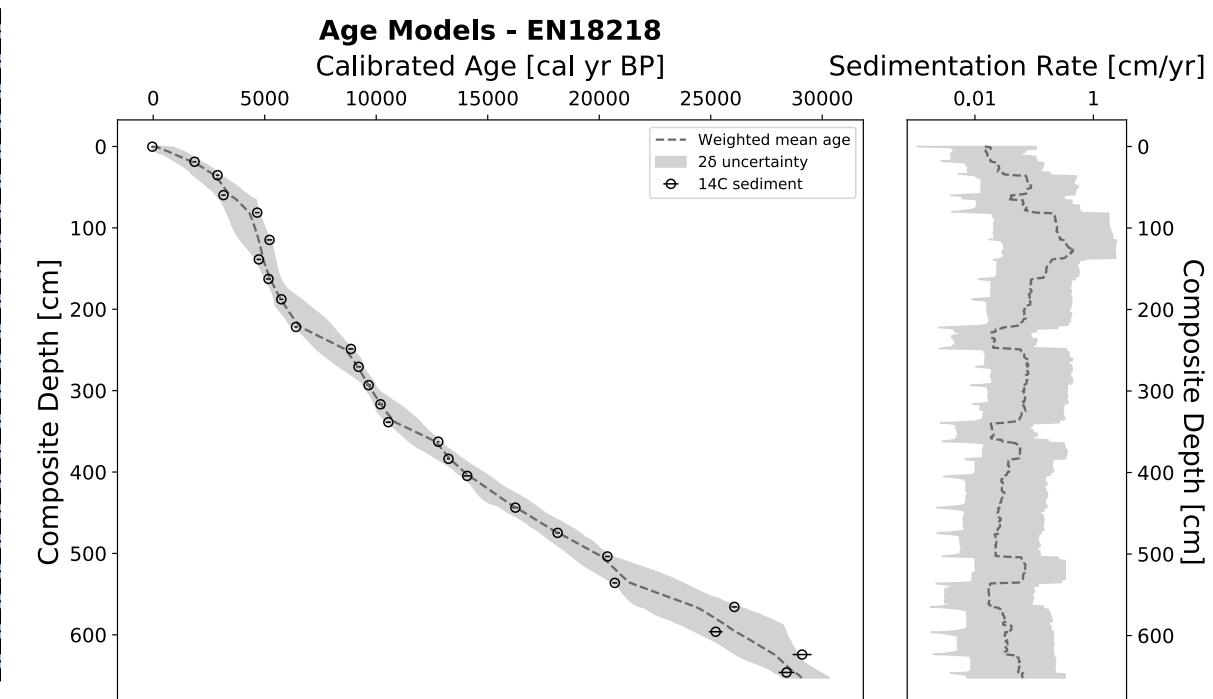


The Use Cases

1st use case - continuously deposited sequence:



Computes five age-depth models with 1σ or 2σ range for one lake

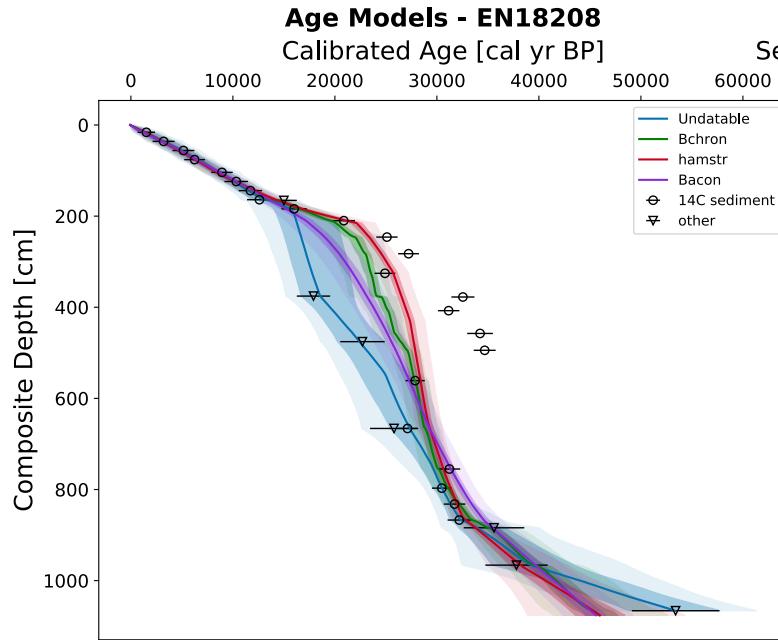


Computes combined age-depth model for one lake

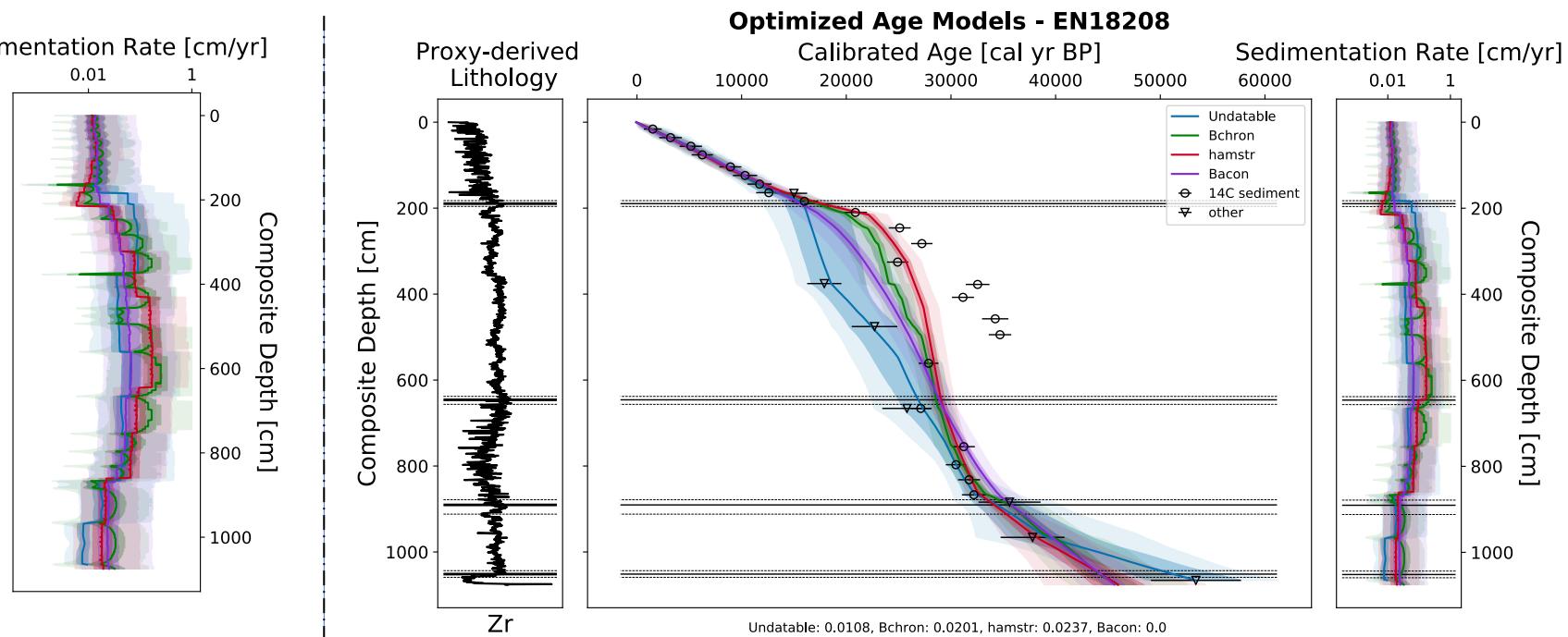


The Use Cases

2nd use case - inconsistent sequence:



Scatter age determination data leads to disagreeing models



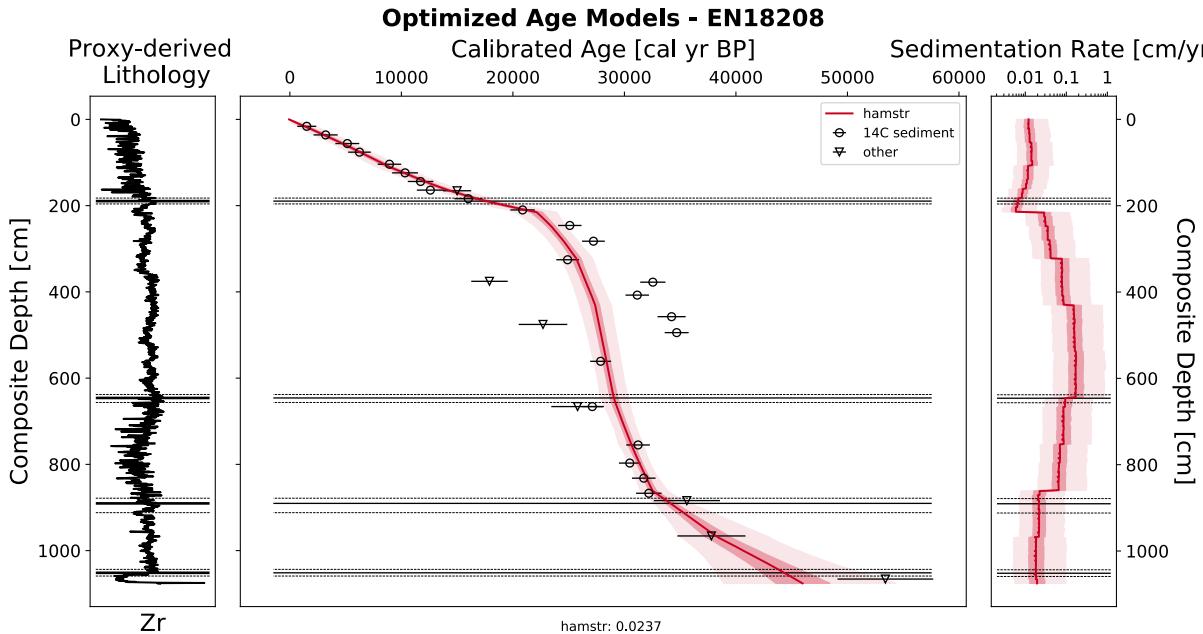
Using independent proxy data indicating lithological changes help us to select models which represent these changes

Holloway, M. J., Henrys, P. A., Killick, R., Leeson, A., and Watkins, J.: Evaluating the ability of numerical models to capture important shifts in environmental time series: A fuzzy change point approach, Environmental Modeling Software, 139, 104993, <https://doi.org/10.1016/j.envsoft.2021.104993>, 2021.

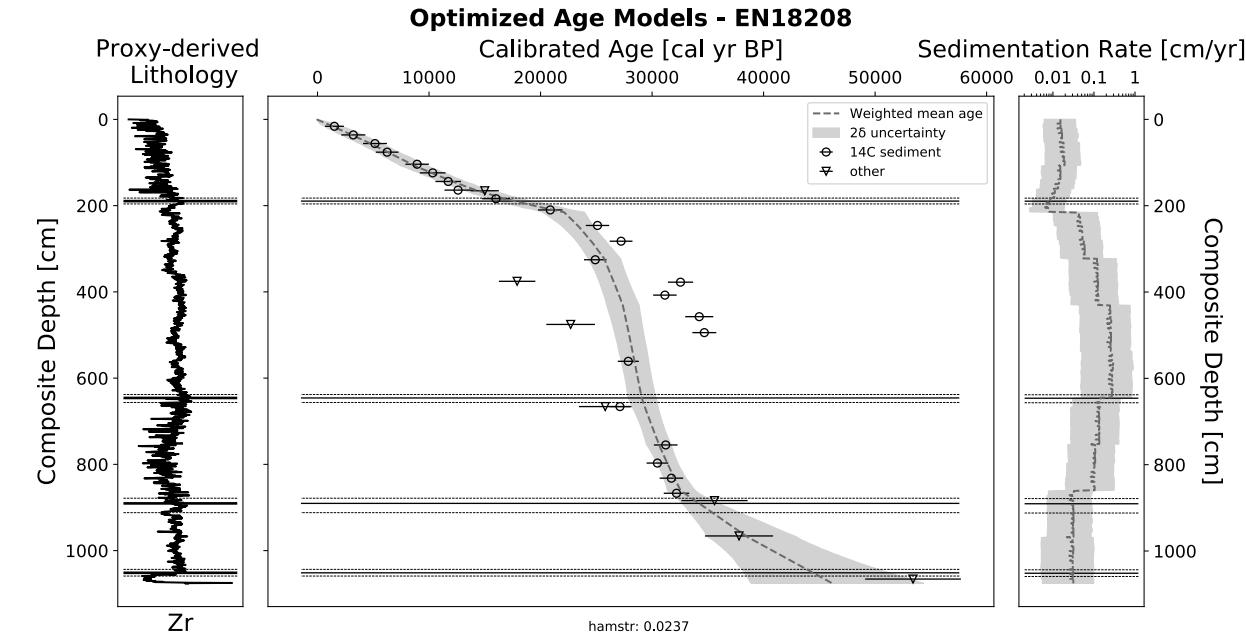


The Use Cases

2nd use case - inconsistent sequence:



Models over a certain threshold represent more likely age-depth models without removing dating points

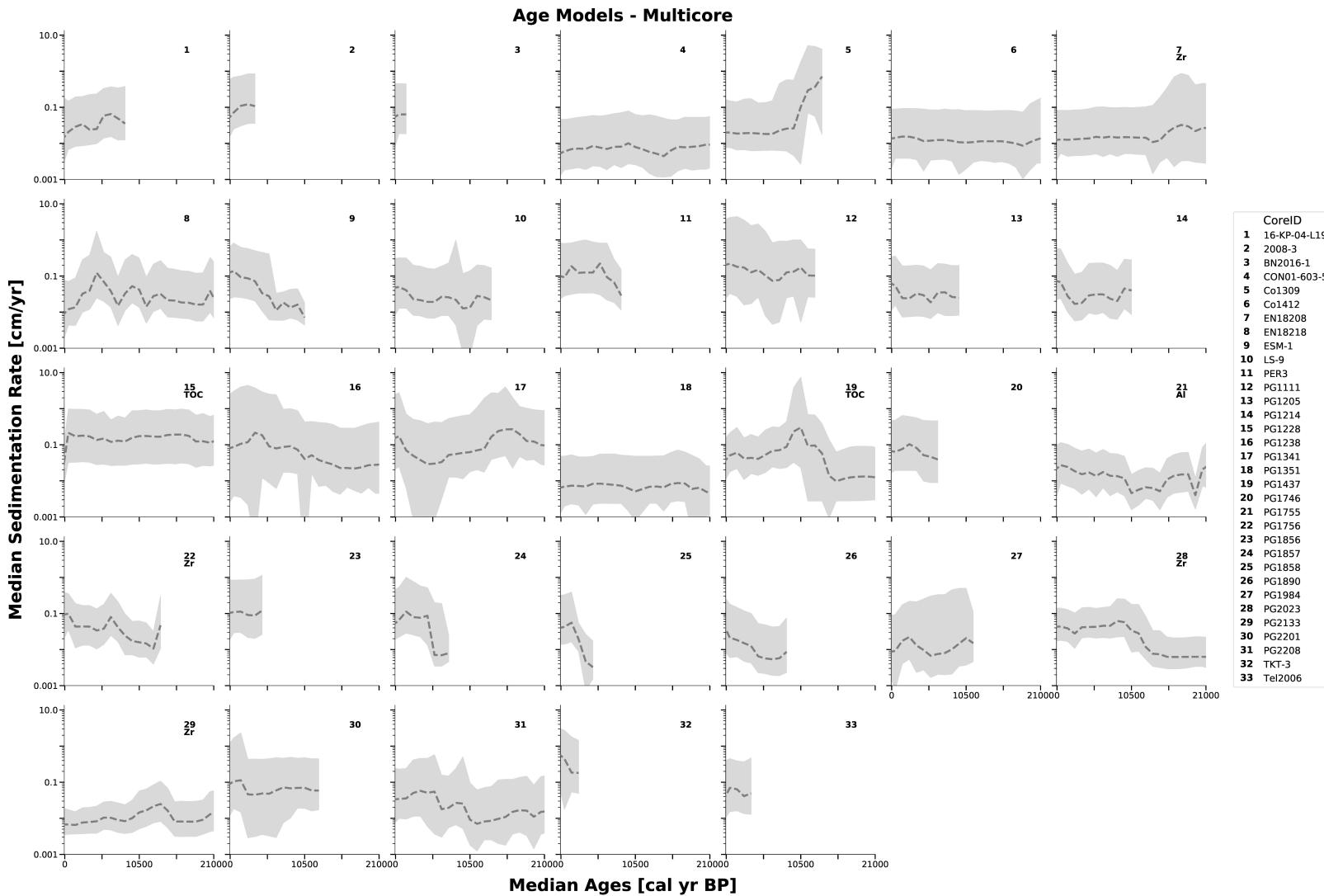


Combined model takes 2-sigma uncertainty from models into account



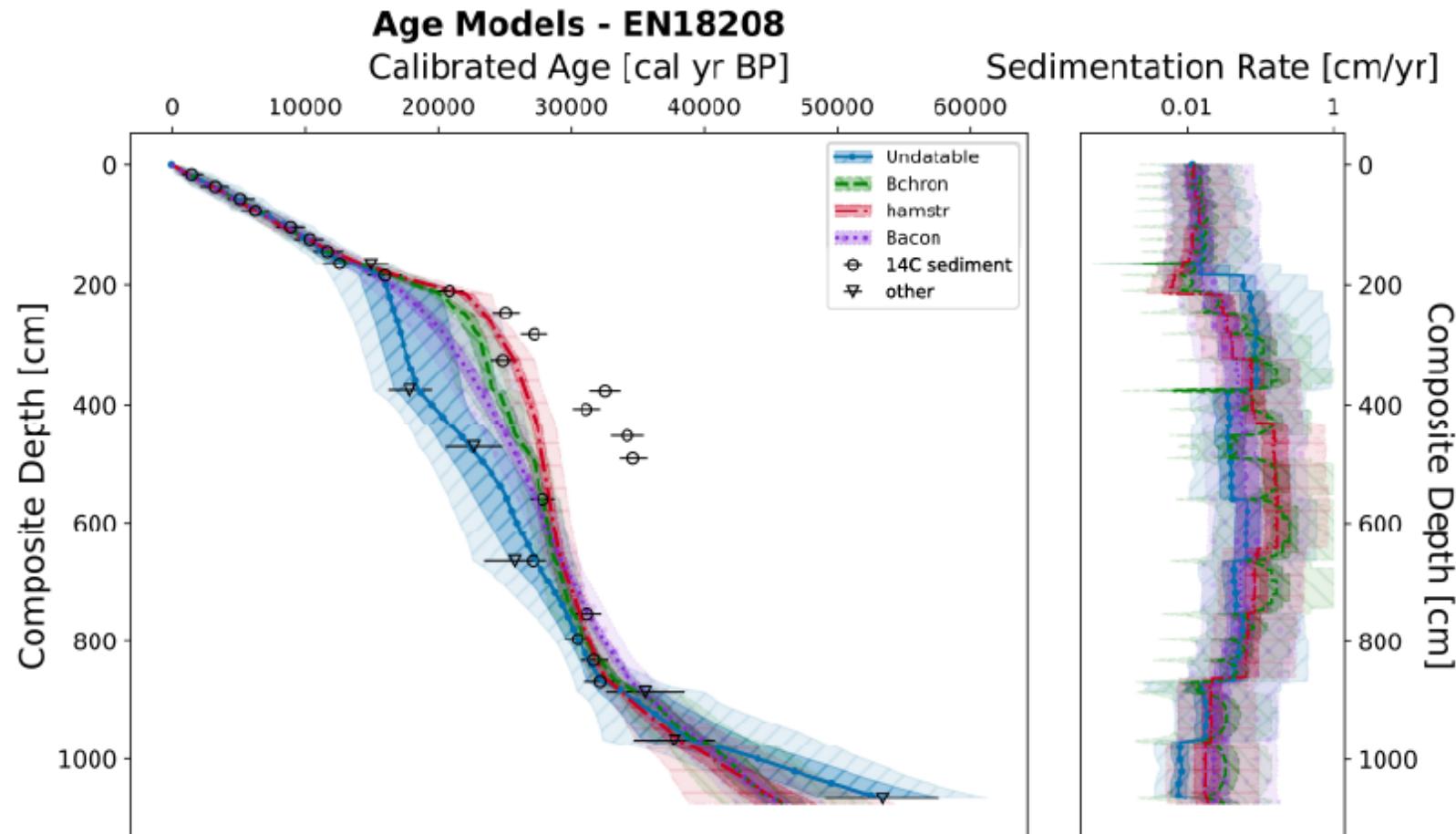
The Use Cases

3rd use case - multiple sediment cores:



The Additional Feature

- For people with color visual deficiency, LANDO creates color-blind friendly output



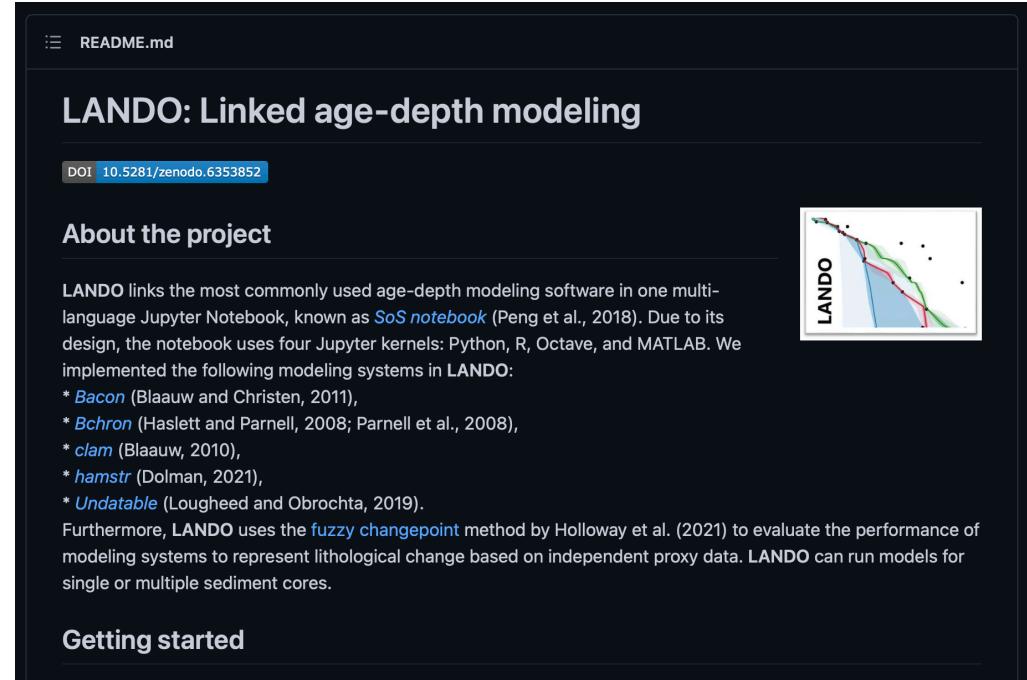
On GitHub

<https://github.com/GPawi/LANDO>



Reference:

Pfalz, G., Diekmann, B., Freytag, J.-C., Syrykh, L., Subetto, D. A., and Biskaborn, B. K.: *Improving age–depth relationships by using the LANDO (“Linked age and depth modeling”) model ensemble*, Geochronology, 4, 269–295, <https://doi.org/10.5194/gchron-4-269-2022>, 2022.



LANDO: Linked age-depth modeling

DOI 10.5281/zenodo.6353852

About the project

LANDO links the most commonly used age-depth modeling software in one multi-language Jupyter Notebook, known as *SoS notebook* (Peng et al., 2018). Due to its design, the notebook uses four Jupyter kernels: Python, R, Octave, and MATLAB. We implemented the following modeling systems in LANDO:

- * *Bacon* (Blaauw and Christen, 2011),
- * *Bchron* (Haslett and Parnell, 2008; Parnell et al., 2008),
- * *clam* (Blaauw, 2010),
- * *hamstr* (Dolman, 2021),
- * *Undatable* (Lougeh and Obrochta, 2019).

Furthermore, LANDO uses the *fuzzy changepoint* method by Holloway et al. (2021) to evaluate the performance of modeling systems to represent lithological change based on independent proxy data. LANDO can run models for single or multiple sediment cores.

Getting started

