CSciBox: An intelligent assistant for dating ice and sediment cores

Elizabeth Bradley[1], Kathleen Finlinson[2], Tyler Jones[3], Colin Lindsey[3], Brett Israelsen[1]
Kenneth A. Anderson[1], Thomas M. Marchitto[3], Laura Rassbach de Vesine[1], and James W. C. White[3],
[1] Department of Computer Science, University of Colorado, Boulder, Colorado, USA
[2] Department of Applied Mathematics, University of Colorado, Boulder, Colorado, USA
[3] Institute for Alpine and Arctic Research (INSTAAR), University of Colorado, Boulder, Colorado, USA

Kathleen.finlinson@colorado.edu

University of Colorado Boulder

Geoscience | CSciBox | Computer science

INSTAAR University of Colorado Boulder

Reproducibility

Reproducibility is a key component of the scientific method. As **computation becomes more central** to the scientific enterprise, it is urgent to address concerns regarding **reproducibility of computational results**.

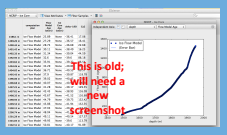
CSciBox addresses reproducibility concerns by:

- Tracks all actions taken with the core, for complete reproducibility
- Includes citation notes every time you use another scientist's tool
- Metadata is always bundled with the core data
- We use Linked Paleoclimate Date format
- It is easy (and mandatory) in CSciBox to save **all the data and metadata that produced an age model** (if you want to save the age model). By metadata, we mean specific parameter choices and such-everything that's required to replicate the computation exactly.
- It is also easy (but optional) to save information about the other **models that the scientist considered but rejected**. <Something about why this is important for science too>
- Hobbes's knowledge base encodes the reasoning behind geoscience research. This reasoning may not always be reported in traditional scientific journaling (?) The process of "algorithmizing" the rules of inference forces scientists to codify and report all of their reasoning. That's cool!
- CSciBox specifically addresses many of the recommendations of the Yale Law School Roundtable on Data and Code Sharing (cite).

The main idea:


Photo: Dorthe Dahl-Jensen

→ CSciBox [Hobbes] →



Geoscientist:
- Has core
- Wants age model
- Brings scientific expertise

CSciBox:
- Helps apply scientific reasoning and algorithms
- Explores alternatives automatically
- Keeps track of all the reasoning involved in the model creation
- Need to explain the relationship between Hobbes and CSciBox

Output:
- One or more proposed age models for the core
- Sophisticated uncertainty estimates
- Summary of the reasoning behind each model

AI-geoscientist interaction

- The scientist is kept "in the loop" at each decision point, and can **guide the process as she sees fit**. At each step, she can compare the results to her intuition or to other age models.
- Hobbes explores alternatives, performs computations, and tracks uncertainty. It **explores the space of possible models** thoroughly and quickly, and always **explains its reasoning** to the scientist, including steps that are conflicting, uncertain, or otherwise problematic.

Argumentative reasoning

Hobbes **considers arguments for and against** each theory, and weighs their strength.

- This **mimics the way scientists think**, and the way they communicate about it with others—and thus streamlines communication between Hobbes and its users
- It also allows the scientist to **consider multiple age models** at once, and enables Hobbes to handle conflicting arguments.
- At this point in the process, for instance, the scientist can choose one or the other of the suggested firn models—or may choose todirect Hobbes to explore both.
- This is the place to put that material about scientific disagreement and conflicting rules

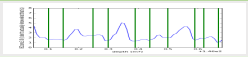
Two components of confidence in the reasoning

- Confidence in the scientific knowledge (validity): Is this rule of inference valid? Validity ranges from a universally accepted rule (accepted), to a gut-level feeling (plausible).
  Accepted > Sound > Probable > Plausible
  A validity level is attached to each rule, and is **determined by the expert geoscientists** on the team. The **discrete levels** of validity correspond to the weights scientists give to their reasoning, and are easier to obtain from experts than continuous values would be.
- Confidence in the evidence (applicability): How well does our situation match the rule of inference, as shown by our dataset?
  Highly applicable > mostly applicable > partly applicable
  Hobbes determines the applicability of a piece of evidence, by analyzing the data.
- When weighing the evidence for or against a theory, Hobbes is designed to intelligently consider the validity and applicability of each argument..

Example: Dating NEEM-S1 ice core

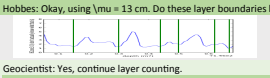
Step 1: Use StratiCounter [Winstrup 15] for automated layer counting

Hobbes: I need a parameter \mu, for the average layer thickness. I guessed \mu= 7 cm, based on a similar core. Do these layer boundaries look correct?



Geoscientist: No, there are too many layer boundaries. Use \mu = 13 cm.

Hobbes: Okay, using \mu = 13 cm. Do these layer boundaries look correct now?



Geoscientist: Yes, continue layer counting.

Hobbes: Okay *[saves \mu=13 cm in the metadata and proceeds]*

Step 2: Align the gas record and the ice record

Hobbes: We need a firn densification model to estimate age difference between gas bubbles and surrounding ice. The standard model is by Herron&Langway, but it assumes there was no snow melt. I will evaluate whether this model is appropriate.

Evidence AGAINST snow melt
Average temperature below freezing
(weaker evidence)

Evidence FOR snow melt
Temperature frequently above freezing
(stronger evidence)

Argument FOR snow melt
(only moderately certain)

Hobbes: Judging by the current temperature values at the location of the NEEM-S1 core (average -3.34 C and standard deviation 22.97), the Herron&Langway model may not be appropriate for this core. Should I use the <name other model here> instead?

Step 3: Stop layer counting after annual signals fade

Hobbes: We are at depth 3525 m. I will check whether the annual signals have faded to see if I should keep layer counting.

Argument: average layer thickness is normal
mostly applicable plausible

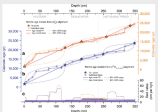
Argument: number of peaks per series is normal
partly applicable sound

Conclusion: continue layer counting
mostly applicable sound

Hobbes: We can continue layer counting after depth 3525 m.

Final output:

A plot produced by CSciBox should go here. (not this exact plot, because I just stole it off the internet) But it will show the finished age models and their uncertainties



Reasoning behind age models:
- Performed automated layer counting between depth 50 m and 4020 m <these are made up!>
- Used Dansgaard-Johnsen flow model for dating depths below 4020 m I don't see that above. I think you're missing a 4th step
- Used two different firn densification models: Herron&Langway (orange curve) and ? (blue curve)
- What about the gas record alignment?
- Etc (more details would be included in the

Hobbes's Knowledge Base

Hobbes is aiming to encode the vast knowledge geoscientists have built up through years of experience dating cores. Hobbes's collection of rules springs out of long, detailed discussions between the AI scientists and the geoscientists on the team. These conversations focus on

1. The methods of reasoning geoscientists use to create & evaluate age models
2. Specific rules of inference geoscientists use in particular situations while creating an age model

Examples of rules obtained from these conversations include:

- The more measurements we have of a core, the more certain we'll be about the resulting age model.
- StratiCounter works well for layer counting if neighboring layers have roughly the same thickness.
- If a sediment layer contains fossils of terrestrial life, it must have been deposited on land.
- If the sedimentation rate changes abruptly, the age model is probably wrong.

The AI scientists then encode each rule of inference into an algorithm Hobbes can understand.

In the works:

- Say what the future work landscape looks like
- Here are some suggestions:
  - Making StratiCounter & Bacon both work
  - Helping Hobbes make smart decisions about when to run expensive computations
  - Adding extensively to Hobbes' rule base

CSciBox is open source

- Source code (python) available on github: `github.com/ldevesine/cscience`
- But you don't have to know python to run it; we also have one-click installers there too
- GNU public license; free to modify/extend/use as you see fit

We are trying to build— and support — a user community. Please join us!
www.cs.colorado.edu/~lizb/cscience.html

This material is based upon work sponsored by the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.