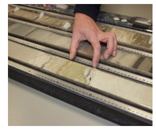
An artificial intelligence tool for complex age-depth models

Liz Bradley, Ken Anderson, Laura Rassbach de Vesine, Vivian Lai, Tom Marchitto, Tom Nelson, Izaak Weiss, and Jim White

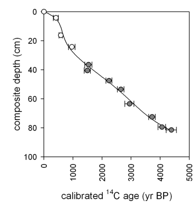
University of Colorado Boulder

INSTAAR



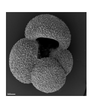
Paleoclimate proxy data

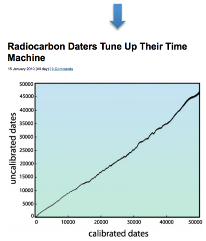
Age model

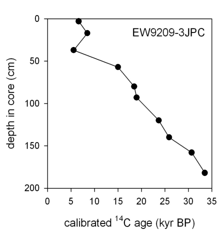


Building age models is hard

- Requires expert knowledge and forensic reasoning







Artificial Intelligence

"Artificial Intelligence is the study of ideas that enable computers to be intelligent. Intelligence includes: ability to reason, ability to acquire and apply knowledge, ability to perceive and manipulate things in the physical world, and others." (PHW 1984)

- *Symbolic AI*
  - logic systems
  - planners, theorem provers
  - rule-based systems
  - qualitative reasoning
  - …
- *Statistical AI*
  - machine learning
  - neural nets
  - support vector machines
  - Bayesian techniques
  - …

- Symbolic AI:
  - reasons generally and reports on its reasoning
  - but someone has to feed it the operative knowledge
  - and "knowledge engineering" is hard.

- Statistical AI:
  - works really well, but requires lots of information to learn from (training sets, priors, …)
  - output = statistics, not explanations

Federal 'Extreme Vetting' Plan Castigated by Tech Experts

By THE ASSOCIATED PRESS NOV. 16, 2017, 8:24 P.M. E.S.T.

Leading researchers castigated a federal plan that would use artificial intelligence methods to scrutinize immigrants and visa applicants, saying it is unworkable as written and likely to be "inaccurate and biased" if deployed.

The experts, a group of more than 50 computer and data scientists, mathematicians and other specialists in automated decision-making, urged the Department of Homeland Security to abandon the project, dubbed the "Extreme Vetting Initiative."

CCC
Computing Community Consortium
Catalyst

Big Data, Data Science, and Civil Rights

Solon Barocas, Elizabeth Bradley, Vasant Honavar, and Foster Provost

Abstract
Advances in data analytics bring with them civil rights implications. Data-driven and algorithmic decision making increasingly determine how businesses target advertisements to consumers, how police departments monitor individuals or groups, how banks decide who gets a loan and who does not, how employers hire, how colleges and universities make admissions and financial aid decisions, and much more. As data-driven decisions increasingly affect every corner of our lives, there is an urgent need to ensure they do not become instruments of discrimination, barriers to equality, threats to social justice, and sources of unfairness. In this paper, we argue for a concrete research agenda aimed at addressing these concerns, comprising five areas of emphasis: (i) Determining if models and modeling procedures exhibit objectionable bias; (ii) Building awareness of fairness into machine learning methods; (iii) Improving the transparency and control of data- and model-driven decision making; (iv) Looking beyond the algorithm(s) for sources of bias and unfairness—in the myriad human decisions made during the problem formulation and modeling process; and (v) Supporting the cross-disciplinary scholarship necessary to do all of that well.

- Symbolic AI:
  - reasons generally and reports on its reasoning
  - but someone has to feed it the operative knowledge
  - and "knowledge engineering" is hard.
- Statistical AI:
  - works really well, but requires lots of information to learn from (training sets, priors, …)
  - output = statistics, not explanations

Building age models is hard

- Requires expert knowledge and forensic reasoning
- Can involve subjective judgements

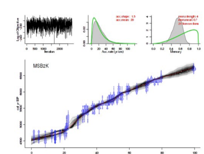
Building age models is hard

- Requires expert knowledge and forensic reasoning
- Can involve subjective judgements
- As well as some fairly complex mathematics

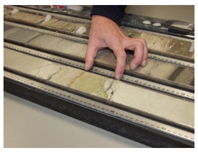
Age-modeling software is powerful, but not necessarily user-friendly

Bacon age-modelling software

If you will begin with certainties, you shall end in doubts,
but if you will content to begin with doubts, you shall end in certainties.
After Francis Bacon (AD 1561 - 1626)

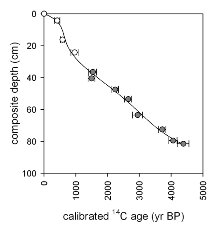



```
Bacon <- function(core="MSB2K", thick=5, prob=0.95, d.min=NA, d.max=NA,
d.by=1, unit="cm", maxcalc=500, depths.file=FALSE, acc.shape=1.5, acc.mean=20,
mem.strength=4, mem.mean=0.7, hiatus.depths=NA, hiatus.shape=1,
hiatus.mean=1000, after=.0001, cc=1, cc1="IntCal13", cc2="Marine13",
cc3="SHCal13", cc4="ConstCal", postbomb=0, d.R=0, d.STD=0, t.a=3, t.b=4,
normal=FALSE, suggest=TRUE, reswarn=c(10,200), remember=TRUE, cleanup=TRUE,
ask=TRUE, run=TRUE, defaults="default_settings.txt", sep=",", dec=".",
runname="", slump=NA, BCAD=FALSE, ssize=2000, rounded=1, th0=c(),
burnin=min(200, ssize), MinYr=c(), MaxYr=c(), find.round=4, bins=c(),
cutoff=.001, plot.pdf=TRUE, rotate.axes=FALSE, rev.yr=FALSE, rev.d=FALSE,
yr.min=c(), yr.max=c(), normalise.dists=TRUE, plot.title=core,
title.location="topleft", d.lab="Depth", yr.lab=c(), d.res=200, yr.res=200,
date.res=100, grey.res=100, width=15, dark=1, greyscale=function(x) grey(1-x),
C14.col=rgb(0,0,1,.35), C14.border=rgb(0,0,1,.5), cal.col=rgb(0,.5,.5,.35),
cal.border=rgb(0,.5,.5,.5), range.col=grey(0.5), range.lty="12",
hiatus.col=grey(0.5), hiatus.lty="12", wm.col="red", wm.lty="12", med.col=NA,
med.lty="12", mar=c(3,3,1,1), mgp=c(1.5,.7,.0), bty="l")
```



Paleoclimate proxy data

CSciBox



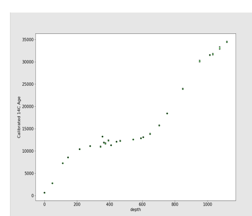
Age model

Argumentation

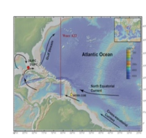
Why argumentation?

- Experts communicate in argument
  - All conclusions are defeasible
  - Multiple simultaneous hypotheses [Chamberlain]
- Shows *reasoning*, not just *answers*
  - Communicate *in the scientist's language*
- Solves the problems well
  - Partial support
  - Contradiction

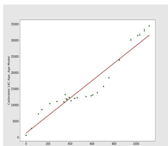
Hobbes in action



KNR166-2-26JPC
Xie et al., Paleoceanography, 27, PA3221



How about using linear regression to build the age model?



- Observed 2nd derivative of the model is small everywhere ➔ slope is consistent ➔ weak argument **in favor of** this model
- No observed reversals in model ➔ very weak argument **in favor of** this model
- Observed residuals are large ➔ very strong argument **against** this model

The strength of the argument **against** this model is stronger than the combined strength of the arguments **for** it, so this is judged to be a bad model

Under the hood…

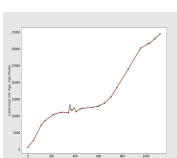
```
Evaluating Linear Regression Model:

Argument FOR Linear Regression Model (weak):
 Evidence FOR Linear Regression Model (weak) <=
        Argument FOR consistent slope (very_strong)
               (Evidence FOR consistent slope (very_strong) <=
                observed 2nd derivative < 0.2)

Argument FOR Linear Regression Model (very_weak):
 Evidence FOR Linear Regression Model (very_weak) <=
        Argument AGAINST reversals (very_strong)
               (Evidence AGAINST reversals (very_strong) <=
                observed reversals found < 1)

Argument AGAINST Linear Regression Model (very_strong):
 Evidence AGAINST Linear Regression Model (very_strong) <=
        Argument AGAINST good data fit (very_strong)
               (Evidence AGAINST good data fit (very_strong) <=
                NOT observed residuals < 0.2)
```

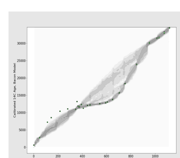
Or maybe piecewise-linear interpolation?



- Observed 2nd derivative of the model is not small everywhere ➔ slope is not consistent ➔ weak argument **against** this model
- Several observed reversals in model ➔ very strong argument **against** this model
- Observed residuals are small ➔ weak argument **for** this model

The combined strength of the arguments **against** this model is (far) stronger than the strength of the argument **for** it, so it too is judged to be an even worse model

What about a BACON model?

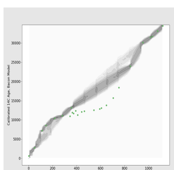


- Observed 2nd derivative of the model is small everywhere ➔ slope is consistent ➔ very weak argument **for** this model
- No observed reversals in model ➔ very weak argument **for** this model
- Model age not within error bounds ➔ weak argument **against** this model
- Model not converging to a single distribution ➔ weak argument **against** this model

Still not a good model…

What if we increased the BACON number-of-iterations parameter?

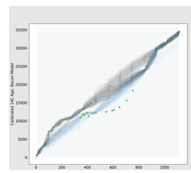
```
Argument FOR Increase Bacon Iterations (strong)
 Evidence FOR Increase Bacon Iterations (weak) <=
       model age not within error bounds
 Evidence FOR Increase Bacon Iterations (weak) <=
       model not converging to a single distribution
```



Reversal-free, has consistent slope, and now converges to a single distribution, but the age points are further outside the error bounds, so it's not a better model.

What if we then increased the BACON section-width parameter?

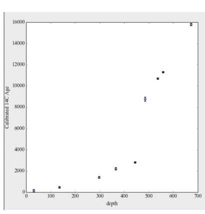
```
Argument FOR Decrease Section Width (weak)
 Evidence FOR Decrease Section Width (weak) <=
       model age not within error bounds
```



10 cm section width in black
5 cm section width in blue

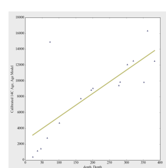
The age points are closer to the error bounds and all of the other properties (reversals, slope, single distribution) are still good, so this one is better...

Reasoning about hiatuses



```
r('Hiatus', arg("hiatus at", "d_i"), very_strong)
r(('hiatus at', 'd_i'), arg('vertical jump', 'd_i'), strong)
r(('vertical jump', 'd_i'),
       calc('percent_change',
             calc('local_slope','d_i'),
             calc('avg_slope')),
       very_strong)
```

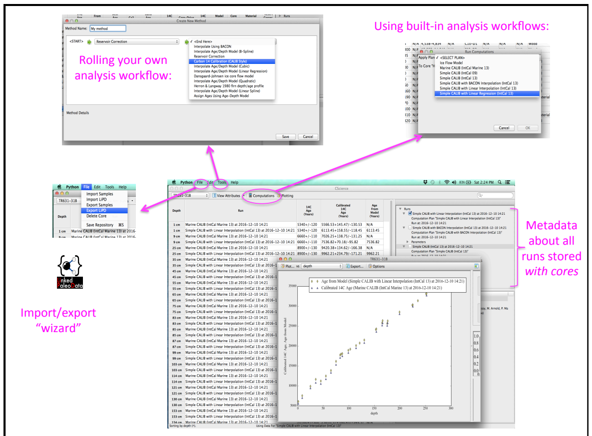
Reasoning about outliers



```
r((’outlier', 'd_i'),
       arg(’err_anomaly', 'd_i'),
       weak)

r((’outlier', 'd_i'),
       arg(’different_material', 'd_i'),
       strong)
```

Depth	Run	Sample Id	Material Dated	14C Age (Years)	Calibrated 14C Age (Years)
22.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89203	Charcoal	380+/-30	425.63+73.37/-91.63
36.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	109356	Charcoal	1245+/-25	1195.26+65.74/-20.26
46.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89204	Charcoal	1570+/-40	1463.63+57.37/-49.63
65 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89205	Charcoal	2700+/-15	2797.39+42.61/-32.39
73.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	131490	Wood	12670+/-380	14949.47+601.53/-674.47
101 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89206	Charcoal	4160+/-45	4697.9+123.1/-70.9
165 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89207	Charcoal	6960+/-70	7796.05+114.95/-95.05
195.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89208	Charcoal	8000+/-100	8855.72+152.28/-138.72
200.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89209	Charcoal	8130+/-60	9088.02+100.98/-90.02
277.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89210	Charcoal	8450+/-80	9439.2+93.8/-29.2
281 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89211	Charcoal	8820+/-160	9890.81+262.19/-200.81
302.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	109357	Charcoal	10310+/-160	12070.34+340.66/-253.34
320.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	109358	Charcoal	10690+/-190	12538.32+259.68/-291.32
351.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	109359	Charcoal	8770+/-370	9863+506/-528
362.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	89212	Charcoal	13560+/-100	16349.34+147.66/-168.34
383.5 cm	Simple CALIB (IntCal 13) at 2017-11-24 10:47	109360	Charcoal	10740+/-310	12519.29+483.71/-345.29



- Graphical User Interface, powerful plotter, lots of built-in tools, can compose your own analysis workflows, …
- Documentable, reproducible, interoperable
- Speak to me after the session for a demo (and/or help getting it installed on your machine)
- The CSciBox code* is open source and freely available on `github`

* We're still busy breaking the AI version every other day, so I wouldn't advise grabbing it unless you have a lot of CS experience

Who & how

Geoscience: Jim White, Tom Marchitto *Software Engineering:* Viv Lai, Izaak Weiss, Suyog Soti, Ken Anderson
AI: Tom Nelson, Laura Rassbach de Vesine
Funding: US National Science Foundation CREATIV/INSPIRE #ATM-0325929

undergrads

Knowledge engineering:

Dave Anderson, Maarten Blaauw, Sze Ling Ho, Colin Lindsay, Amy Myrbo, Tyler Jones, Kira Rehfeld

Thanks!

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.

Forensic paleo reasoning

- The data that you have:
  - Physical & chemical properties of some stuff
- What you want to figure out:
  - The past history of that stuff:
    - When & how it got there
    - What happened since then
- What you know:
  - A set of processes that may have acted upon that stuff
- What you don't know:
  - Which of those processes really *were* involved, and what the parameter values were
- How you proceed:
  - Multiple simultaneous hypotheses

Can we automate that reasoning?

What's hard about automating forensic paleo reasoning

- Combinatorial explosion of scenarios
- Which may involve processes with continuous-valued parameters
- So can't just do brute-force abduction
- Knowledge engineering is a challenge…

What's hard about automating forensic paleo reasoning, cont.

- Representation & reasoning issues
  - Expert reasoning involves lots of hypotheses & heuristics
  - It's often contradictory
  - It's not absolute; several weaker conclusions can defeat a stronger one
  - So most of the standard AI solutions won't work
  - And scientists are often suspicious of automated reasoning results
  - One nice solution to all of that: argumentation