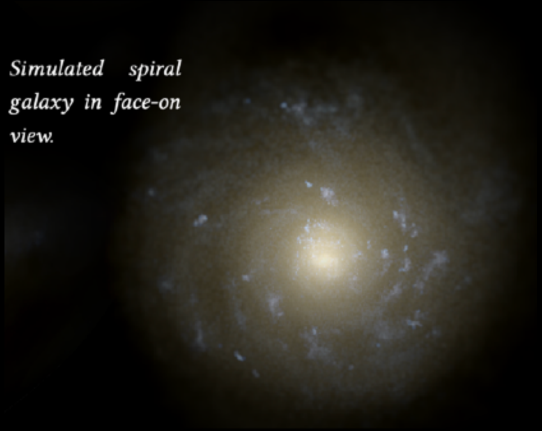
ASTR4004

COMPUTATIONAL ASTRONOMY

Week 9 https://github.com/svenbuder/astr4004_2024_week9

Spiral galaxy M74 in face-on view. ***Figure credit:*** *Gemini Observatory, GMOS Team*



Simulated spiral galaxy in face-on view.

Figure credit: *Tobias Buck*

DR SVEN BUDER | RESEARCH SCHOOL OF ASTRONOMY AND ASTROPHYSICS

Australian National University

My idea for our 3 weeks:

Assignment 3
due in Week 10a
(October 8)

Week	Summary	What I actually plan to talk about
7	Data Processing	git, csv/FITS Files, ADQL/SQL, joining & cleaning catalogues, ...
	Statistics	Fantastic Uncertainties and how to calculate/report them, ...
	Plot Clinic	How to plot well & better - with my (and hopefully your) examples, ...
8	Regression	how to fit y = f(x), if y (and even x) have uncertainties, python fitting packages and when to apply them how to which function, ...
	Dimensionality Reduction	Principal Component Analysis (PCA), tSNE, ...
9	Clustering	k-means, HDBSCAN, Gaussian Mixture Models (GMM), ...
	Model Selection	AIC & BIC, train/test sets, ...
	Interdisciplinary Thinking	How to think abstract or creative and bridge barriers/gaps: How your expertise can help other researchers/industry, ...

Useful Resources: www.astrobetter.com

AstroBetter

Tips and Tricks for Professional Astronomers

Blog About Archives Wiki

Tools for Scientist Workflows Part 5: Notion – Project Management

Author: Guest • Published: February 19, 2024 • 1 comment

Julio Morales is a second year PhD student in the Department of Astronomy at New Mexico State University, where he conducts research on the flow of plasma in the solar interior using the technique of time-distance helioseismology. This work is part of the COFFIES collaboration, whose goal is to successfully simulate the solar magnetic field [...]

Useful Resources: Bibliography management

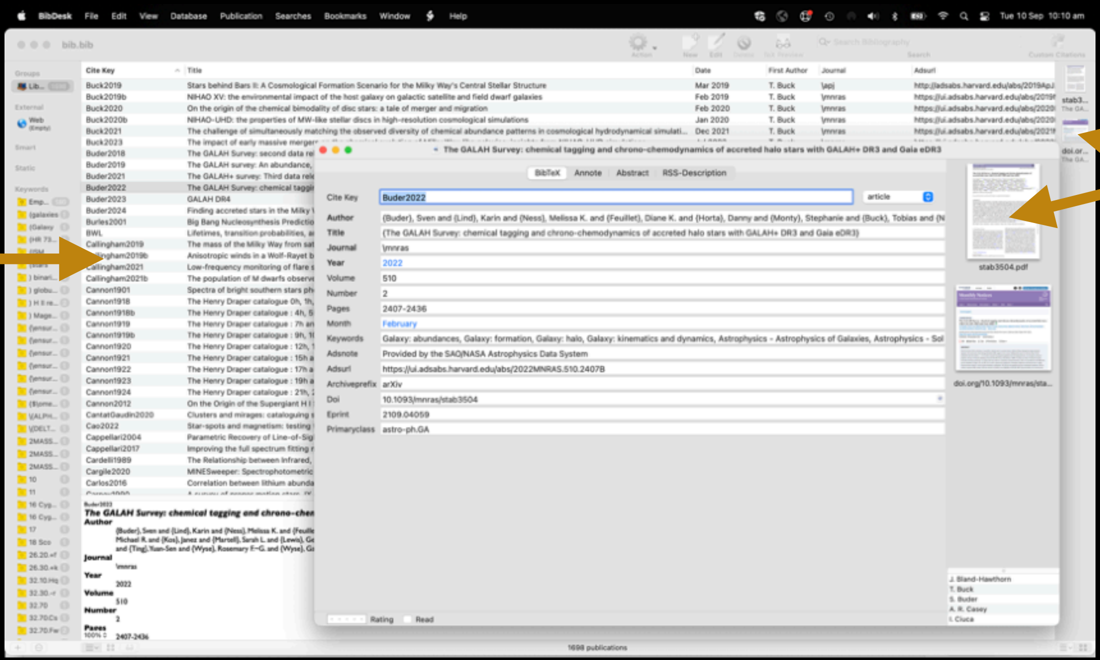
Advantage: 1 *.bib file to rule them all

Easy citation:

e.g. topic

or

LastnameYear



PDF

Different options: BibDesk, Zotero, Mendeley, …

MODEL SELECTION

All models are wrong. Some are useful.

Statistics will never tell you whether the model you have is the *true* model or not. As a Bayesian all you can do is *compare* models.

Australian National University

Which model fits better?

Remember Bayes:
Posterior $\propto$ Likelihood · Prior

$$p(\mathbf{w}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{w})p(\mathbf{w})}{p(\mathcal{D})}$$

Let's assume a uniform prior:

$$\log p(\mathbf{w}|\mathcal{D}) \propto \log p(\mathcal{D}|\mathbf{w})$$

Laplace approximation:

$$\log p(\mathcal{D}|\mathbf{w}) \approx \log \hat{p}(\mathcal{D}|\mathbf{w}) - \frac{k}{2}\log N$$

Mode of the Likelihood: MLE
Maximum Likelihood Estimate

2nd derivative / curvature of
the likelihood new MLE

Bayesian Information Criterion:

$$\mathrm{BIC} = -2\log \hat{p}(\mathcal{D}|\mathbf{w}) + k\log N$$

Useful if you have models with different numbers k of parameters

WHAT MAKES A GREAT SCIENTIST?

Great scientists: Student / PostDoc / Professor

- **Curiosity and Passion:** Driven by a natural desire to discover.
- **Fundamental Knowledge:** Building a strong foundation.
- **Persistence:** Overcoming obstacles in research.
- **Collaboration:** Learning and sharing with peers.
- **Ethics:** Integrity in data collection and analysis

- **Legacy:** Shaping the future of the field and mentoring future leaders.
- **Scientific Leadership:** Influencing policy, education, and funding priorities.
- **Advocacy:** Promoting diversity, equity, and inclusion in science.
- **Lifelong Learning:** Staying curious and engaged with new developments.
- **Giving Back:** Supporting the scientific community and fostering discovery.

- **Independence:** Leading projects while collaborating.
- **Critical Thinking:** Addressing research gaps with creativity.
- **Networking:** Engaging in professional relationships.
- **Mentorship:** Starting to guide junior scientists.
- **Publishing:** Communicating results through papers and presentations

- **Leadership:** Setting a clear research vision and managing teams.
- **Mentorship:** Developing younger scientists and contributing to the community.
- **Collaboration:** Engaging in interdisciplinary research.
- **Impact:** Making significant contributions to the field and society.
- **Ethics:** Ensuring the societal benefit of research.

- **Research Identity:** Establishing a niche in the field.
- **Grant Writing:** Securing funding and managing projects.
- **Balancing Teaching and Research:** Inspiring students while advancing science.
- **Mentoring:** Supporting the development of junior researchers.
- **Ethical Responsibility:** Ensuring research integrity and reproducibility.

AND OUTSIDE OF ACADEMIA?

Talking with a former colleague who is a data scientist now

"To be honest, I think that this is one thing which sets apart people who have done actual research from those who haven't: knowing that the tools you use (and might learn in a computational astrophysics seminar!) are just tools. The really hard thing is translating the big picture questions — 'why do galaxies look the way they do', or 'why are the chemical abundances of stars in the MW the way they are' — into something concrete you can use those tools on."

"Obviously I understand that this is very hard to teach, and probably only comes with experience. It's definitely much easier to teach people how to run a convolutional neural network on the MNIST data set, or how to use k-means clustering, or whatever. But this means that people who leave uni with a masters in data science (say) and who've never done real research end up thinking that the algorithms they've learned are the core thing about doing science. When in fact it's all the thinking that comes before that!"

A few other things astro research has taught me/helped me with:

- Coding is a big thing, especially getting comfortable translating thoughts into code. Also even the basics of things like using unix, the terminal, etc.
- Coming up with the right plot to prove (or even better- disprove) your point, as well as being able to convince others!
- The experience writing academic papers has come in very handy. I now have to write short reports, which are a bit like papers but get finished in ~weeks rather than months/years.
- The ability to pick up the things I need from a paper/source without getting really bogged down in all the details. Skim-reading, I guess.
- Working with imperfect data.
- Doing a "good-enough" job at something, and understanding the ways that your model/results/etc will be wrong. Again, I don't think many undergrad courses really go into this, as you end up using perfectly clean data and get a 99% accurate result.

What my colleague has learn when transitioning out of research:

"As part of the interview process, the current place I'm working at gave me a take-home task to complete. I was given two time series of data and asked to investigate whether it was possible to predict future data."

"I actually thought about the problem a bit more and ended up using basic linear regression and a simple rolling correlation measurement. This went down really well and they offered me the job."

"In terms of things I wish I was taught and wasn't, the main area would be time series analysis. I found this book helpful at the start, even if it is geared at people coding in R: https://bookdown.org/mpfoley1973/time-series/. The smoothing chapter in particular was very helpful- we use exponential moving averages for so many things at work."

Also check out ASTRO 3D's beyond-astro video series: https://astro3d.org.au/beyond-astro/

TIME-SERIES ANALYSIS

Australian National University

INTERDISCIPLINARY THINKING

Introduction to Interdisciplinary Research in Astronomy

- What is interdisciplinary research?
  - Collaboration across subfields within astronomy (e.g., spectroscopy, simulations).
  - Integration of astronomical research with other scientific disciplines (e.g., climate science, biology).

- Why is it important?
  - Expands knowledge, enhances problem-solving, and drives innovation.
  - Tackles complex questions that require cross-disciplinary approaches.

Hurdle 1: Divergent Data Types and Techniques

- Challenges:
  - Different subfields use different data types (e.g., high-resolution spectra vs. large-scale simulations).
  - Varied techniques across stellar, galactic, and cosmological research.
- **Example**: Stellar spectroscopy vs. cosmological simulations.
- Solution:
  - Establish unified data platforms for sharing across subfields (e.g., joint data repositories).
  - Encourage collaborative workshops and research projects.
- **Example**: Unified stellar abundance measurements (e.g., GALAH data) used in both simulations and observational astronomy.

Hurdle 2: Language and Terminology Differences

- Challenges:
  - Different fields/subfields use specialized terminology.
  - Researchers may not immediately understand each other.

- **Example**: Jargon in cosmology vs. spectroscopy.

- Solution:
  - Use glossaries in collaborative papers.
  - Promote cross-field education through interdisciplinary workshops or presentations.

- **Example**: A shared glossary in collaborative papers or training students in interdisciplinary terminology.

Hurdle 3: Data Scalability Issues

- Challenges:
  - Astronomers studying small-scale phenomena (e.g., stars) work differently from those studying large-scale structures.
  - Different needs in data handling and tools for different scales.

- **Example**: stellar spectroscopy vs. photometric quasars

- Solution:
  - Develop analysis tools that work across scales (small to large datasets).
  - Open-access archives for easy retrieval and comparison of data.

- **Example**: Gaia data applied in multiple research contexts, from individual stars to quasars and galaxy evolution.

Hurdle 4: Different Research Goals Across Disciplines

- Challenges:
  - Astronomy’s long-term goals vs. immediate societal challenges in fields like environmental science.
  - Differences in priorities, funding cycles, and timescales.

- **Example**: Astronomy’s focus on cosmic questions vs. environmental science addressing urgent issues like climate change.

- Solution:
  - Align short-term practical outcomes with long-term scientific goals.
  - Use astronomy’s methods for immediate applications (e.g., bushfire detection through leaf spectroscopy).

- **Example**: Collaborating on leaf reflectance spectra to detect bushfire risk by using astronomical spectral analysis techniques.

Hurdle 5: Methodological and Cultural Gaps

- Challenges:
  - Astronomy’s hypothesis-driven approach vs. experimental methods in other fields.
  - Different expectations for data interpretation and outcomes.

- **Example**: Computational modelling in astronomy vs. field experiments in biology.

- Solution:
  - Form interdisciplinary teams with experts who can "translate" between fields.
  - Foster long-term collaboration to build shared methodologies.

- **Example**: Astronomers working together with evolutionary biologists on phylogeny (evolutionary trees) of plants or animals, and stars
- **Example:** Computer scientists working alongside astronomers or environmental scientists

Hurdle 6: Technical and Infrastructure Barriers

- Challenges:
  - Specialized tools and software in astronomy may be inaccessible to other disciplines.
  - Computational requirements for processing astronomical data differ from those in other fields.

- **Example**: Spectroscopic analysis tools not readily available to environmental scientists.

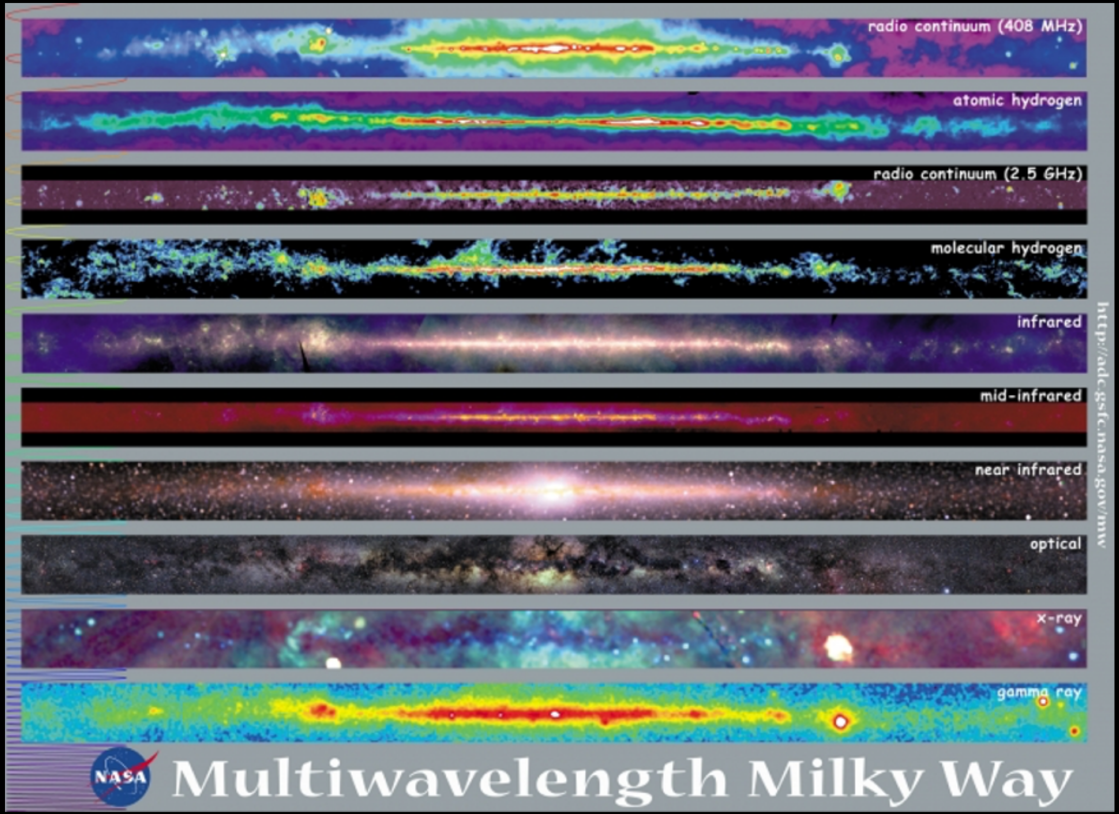
- Solution:
  - Develop open-source tools that are adaptable to multiple fields.
  - Create simplified interfaces for non-astronomers to use astronomical tools.

- **Example**: Open-source software like Python packages for spectroscopy that can be applied to environmental studies.

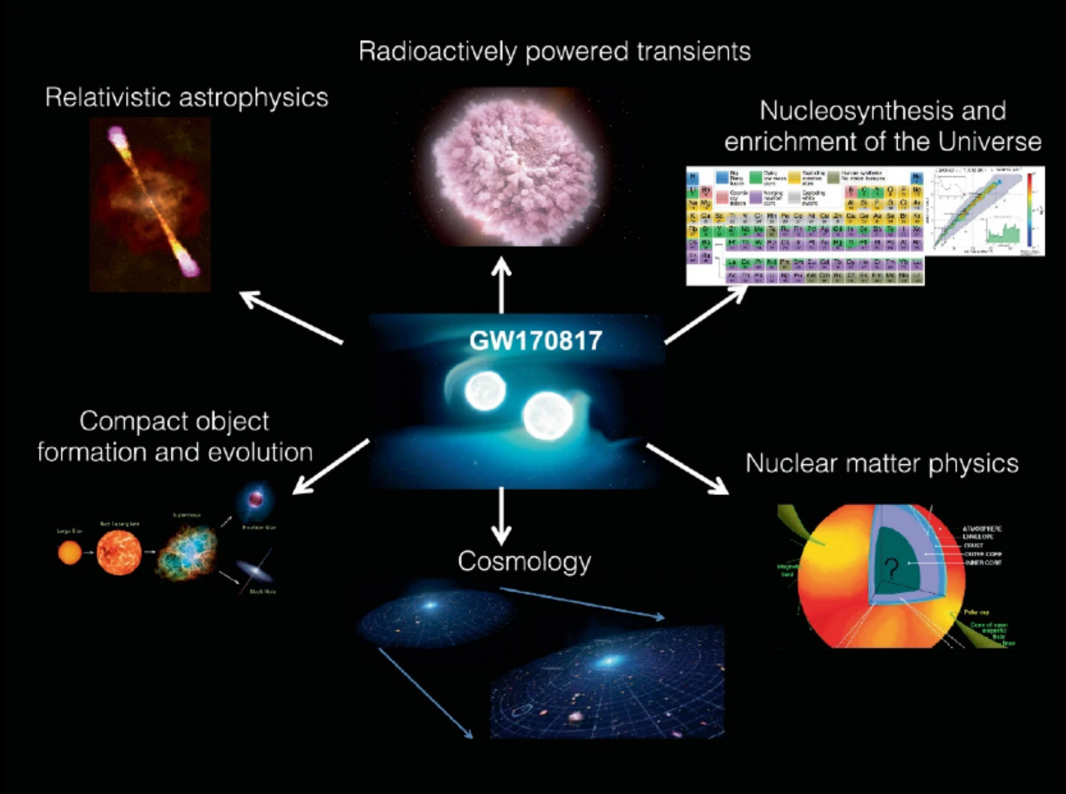
INTERDISCIPLINARY THINKING:

OVERCOMING LANGUAGE AND TERMINOLOGY DIFFERENCES

Understanding our Universe

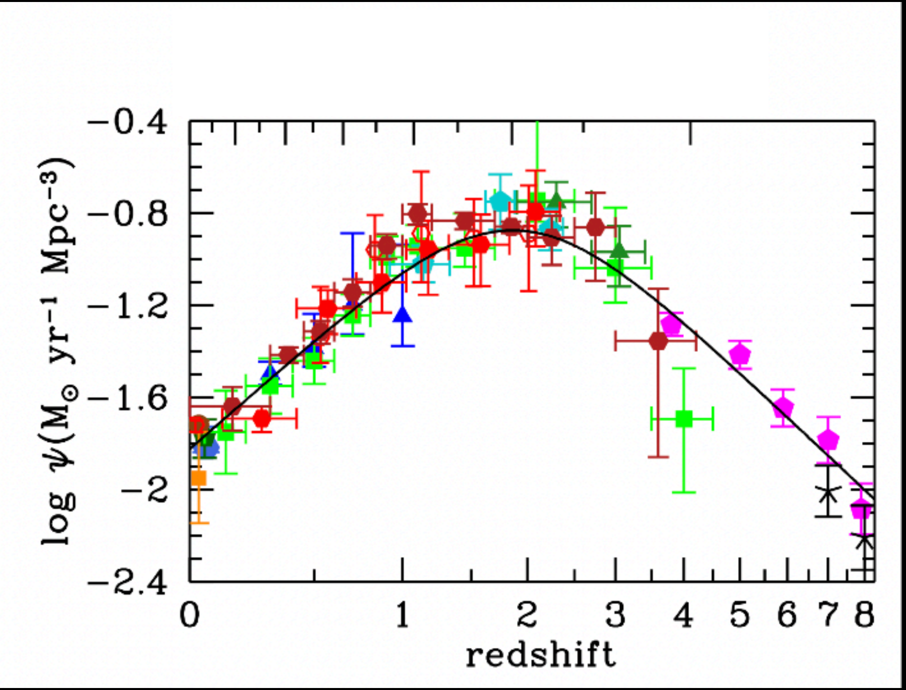


Understanding our Universe



"Time" in context: redshift *z* vs. *stellar age* / Gyr

Logarithmic Star Formation Rate Density

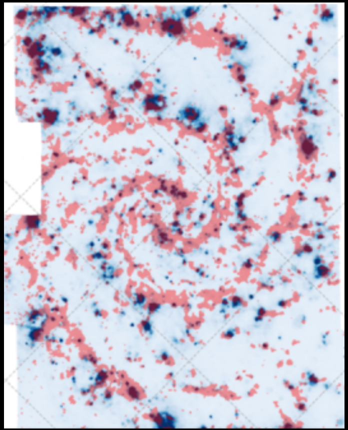


Madau & Dickinson (2014): 4577 citations

"Size" in context: pixels vs. deg vs. kpc vs. R_e

MUSE

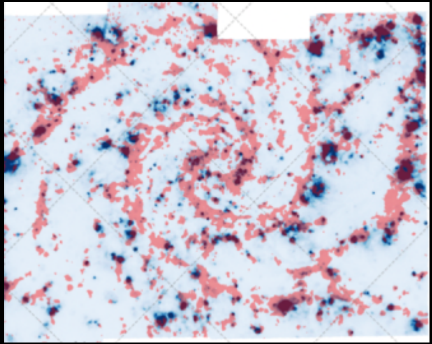
Y Pixels



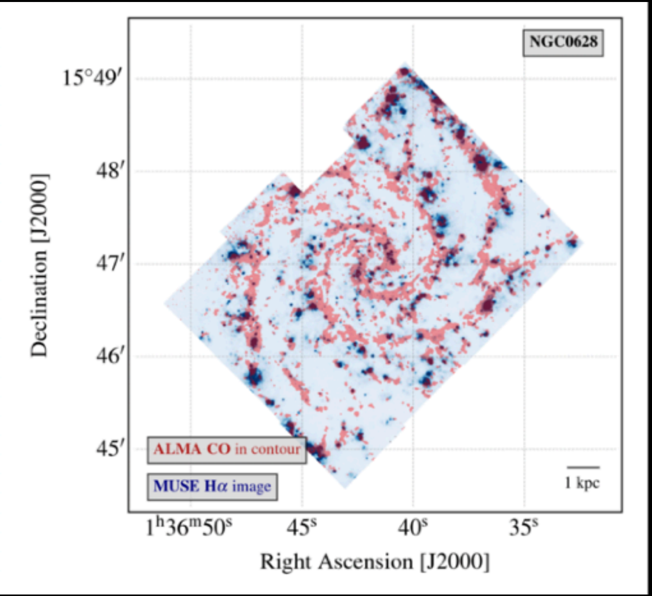
X Pixels

ALMA

Y Pixels



X Pixels

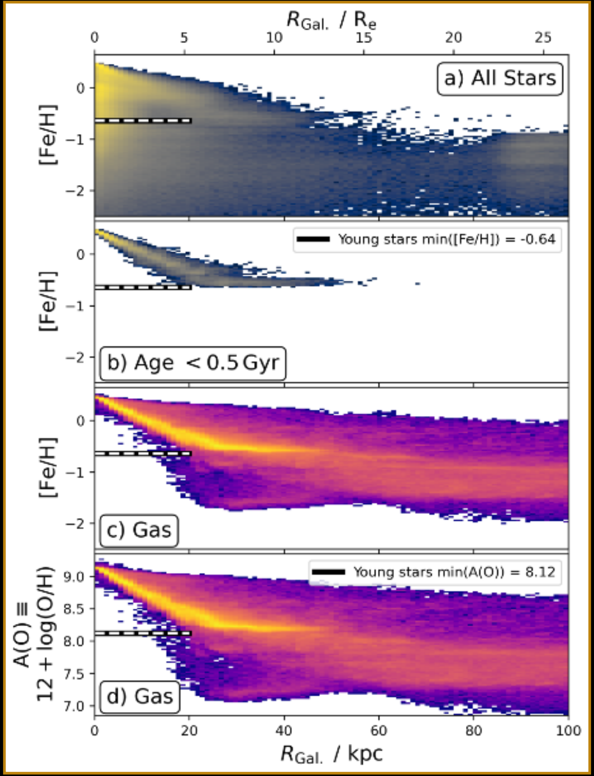


Remember to rotate everything to the same reference frame (e.g. ra/dec)

Sky coordinates: ra/dec or l,b in °/'/''/mas
Physical scales: AU/pc/kpc/Mpc
Relative size: R_{25}/R_e (effective/half light/mass R)

PHANGS-ALMA (Leroy et al., 2021)

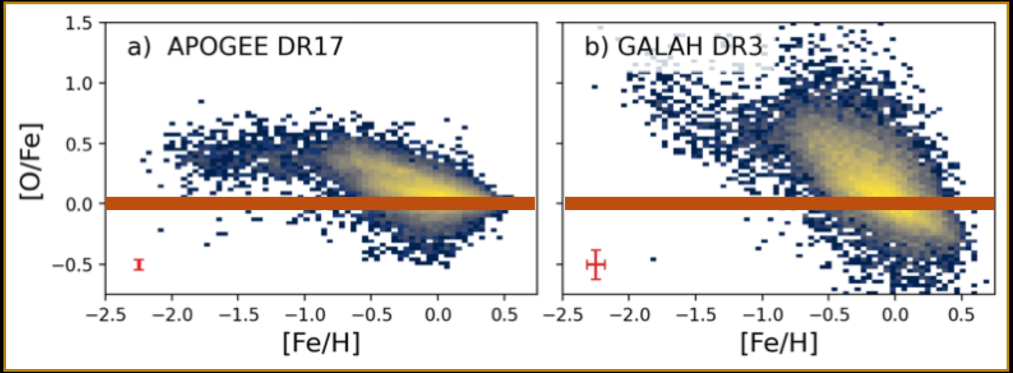
"Metallicity" in context: [M/H] vs. [Fe/H] vs. log10(O/H)+12



"Chemistry" is difficult!

We often use different tracers of "chemistry"

Historically, we have not been able to measure precisely enough for difference to matter



But clearly: O is not Fe! And this does matter!

INTERDISCIPLINARY THINKING:

NOW IT’S YOUR TURN TO PITCH YOU SCIENCE TO OTHERS:

WHY IS YOUR SCIENCE INTERESTING? WHAT PROBLEM ARE YOU TRYING TO SOLVE HOW?