ASTRO 3D

AMSI
AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Geostatistics of Galaxies

Using better maths to help understand the physics of galaxy evolution

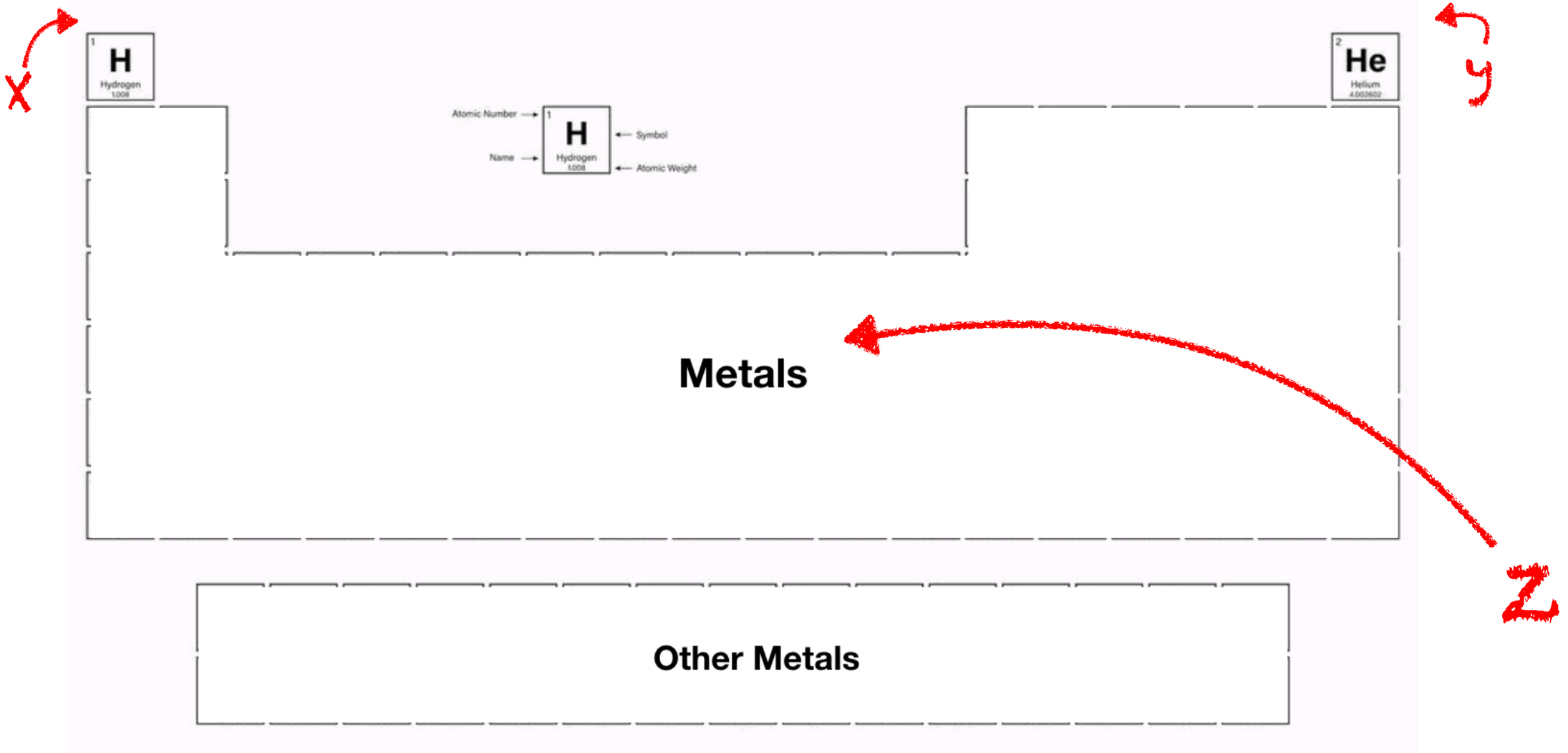
Benjamin Metha, July 19, 2021

Outline

The question: How do galaxy metallicities vary on small scales?

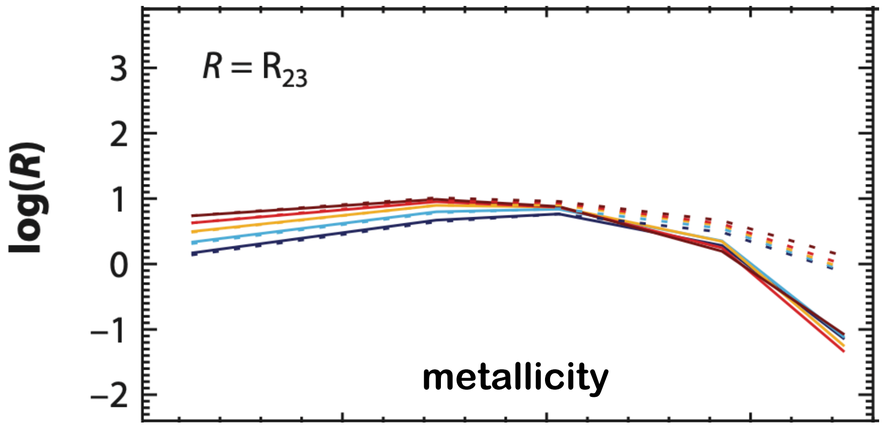
- Background
  - *The physics:* What is metallicity? How do we measure it?
  - *The maths:* Introducing Geostatistical Hierarchical Models
  - *The theory:* The KT18 model
- Early results

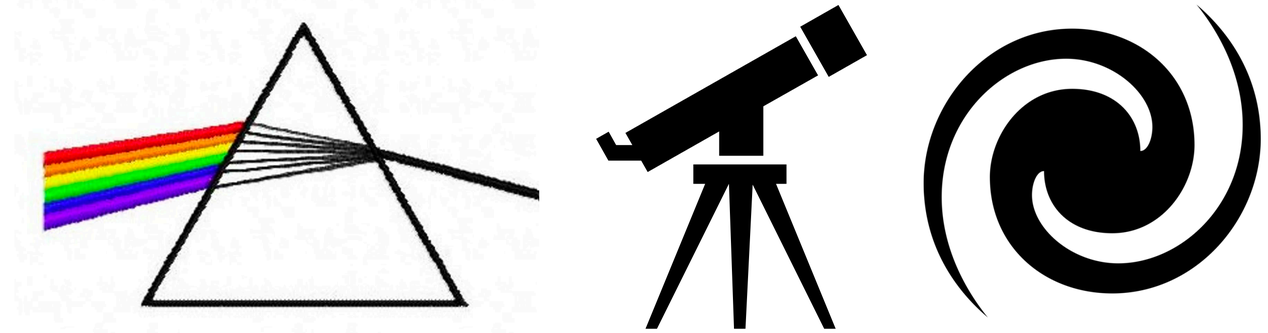
Astronomer's Periodic Table of the Elements

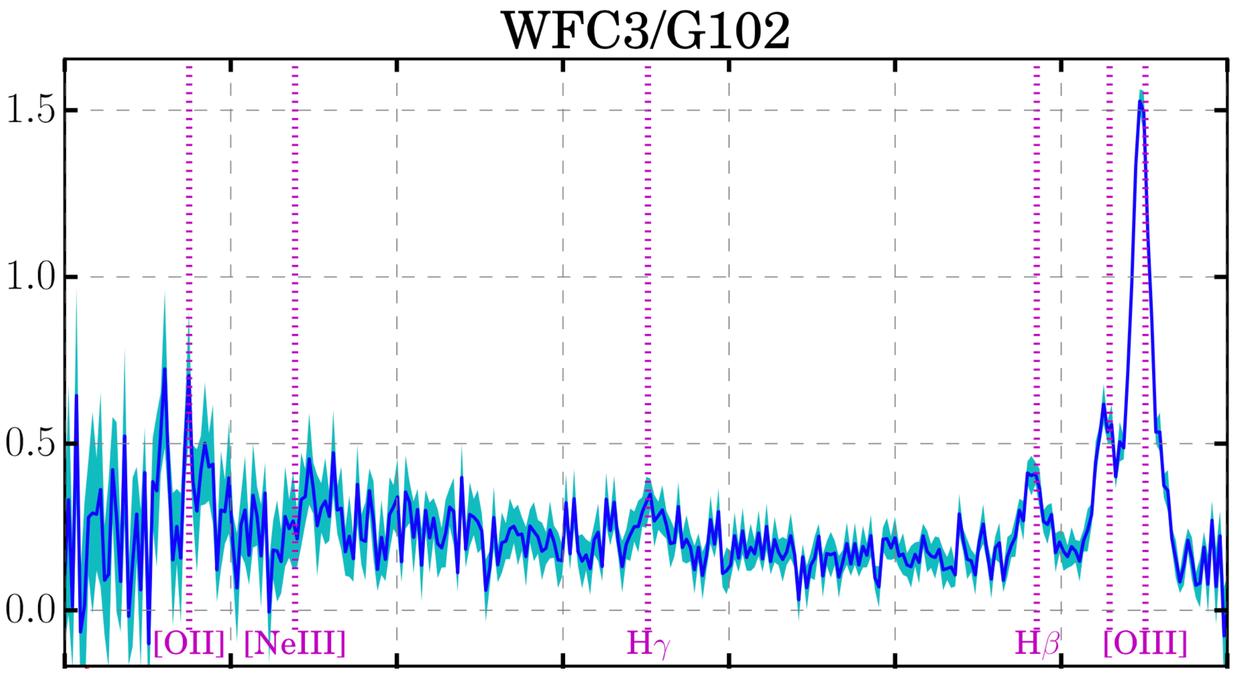


Metallicity is measured using spectroscopy

- “Spikes” in spectrum associated with emission lines
- Ratio of emission lines ~ metallicity

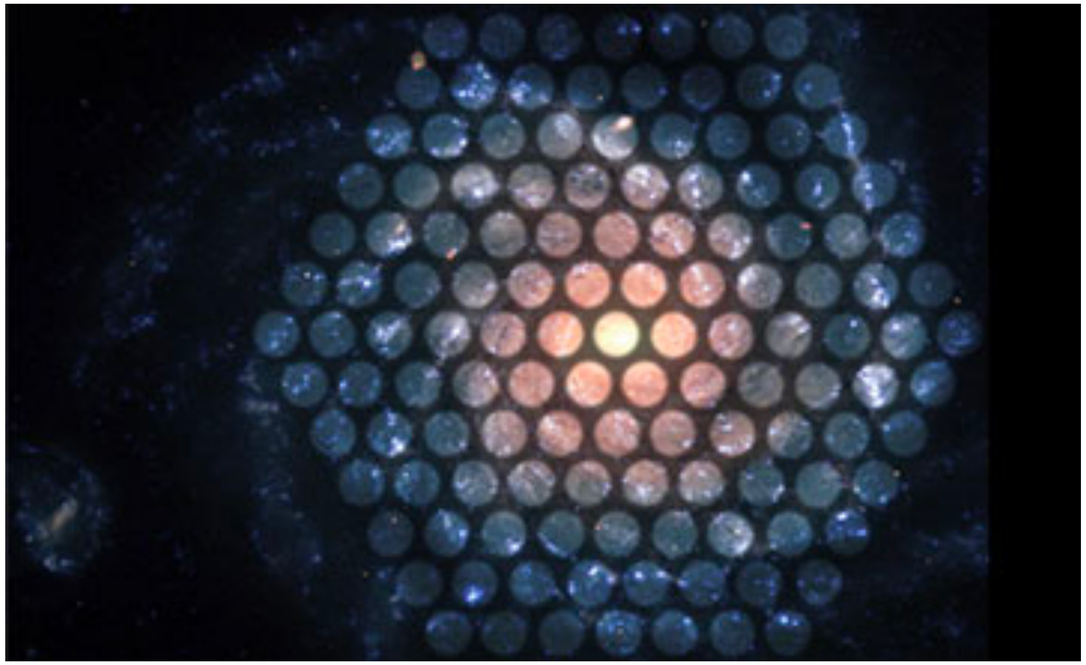


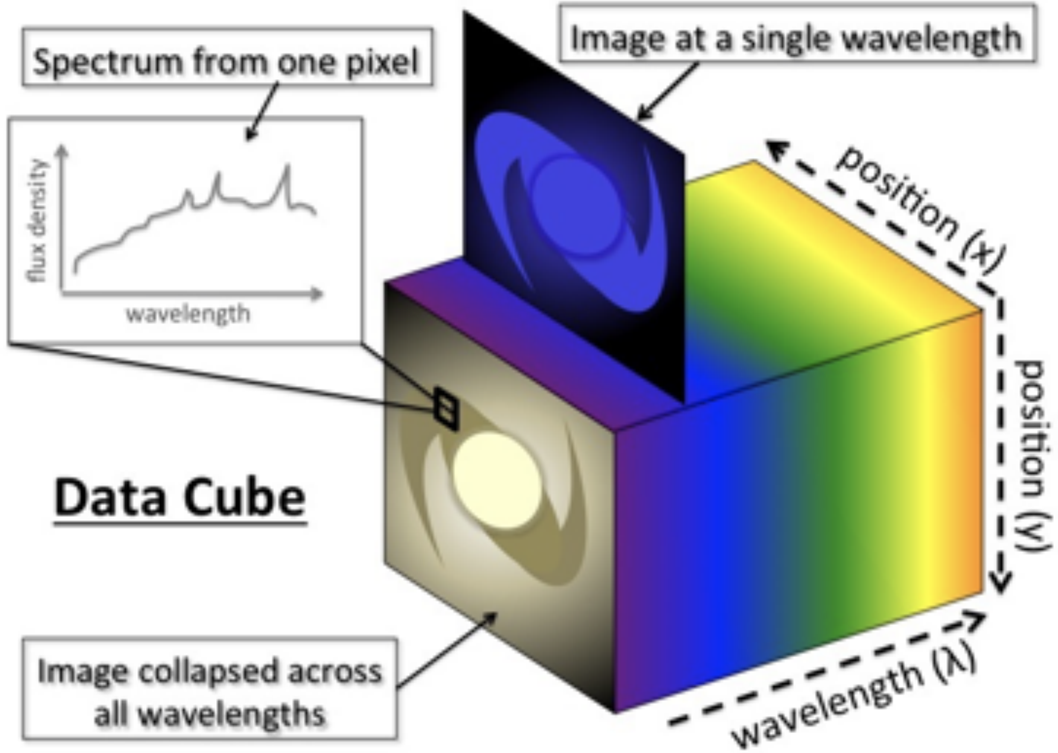




Integral Field Spectroscopy

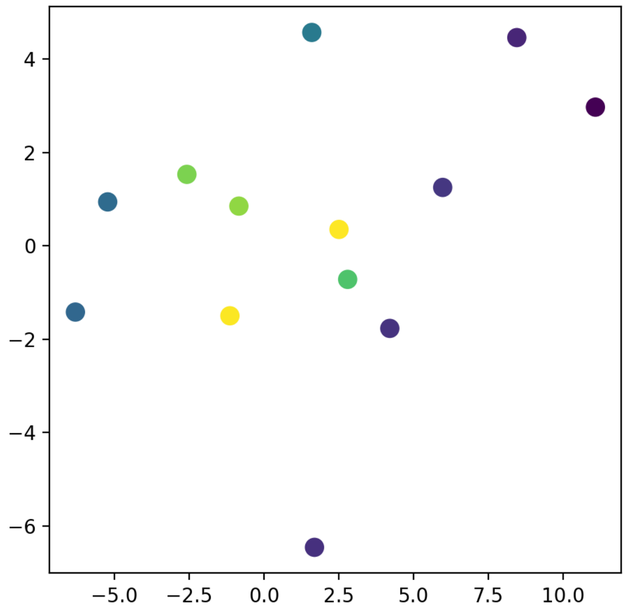
How to see internal variation in galaxy metallicities?





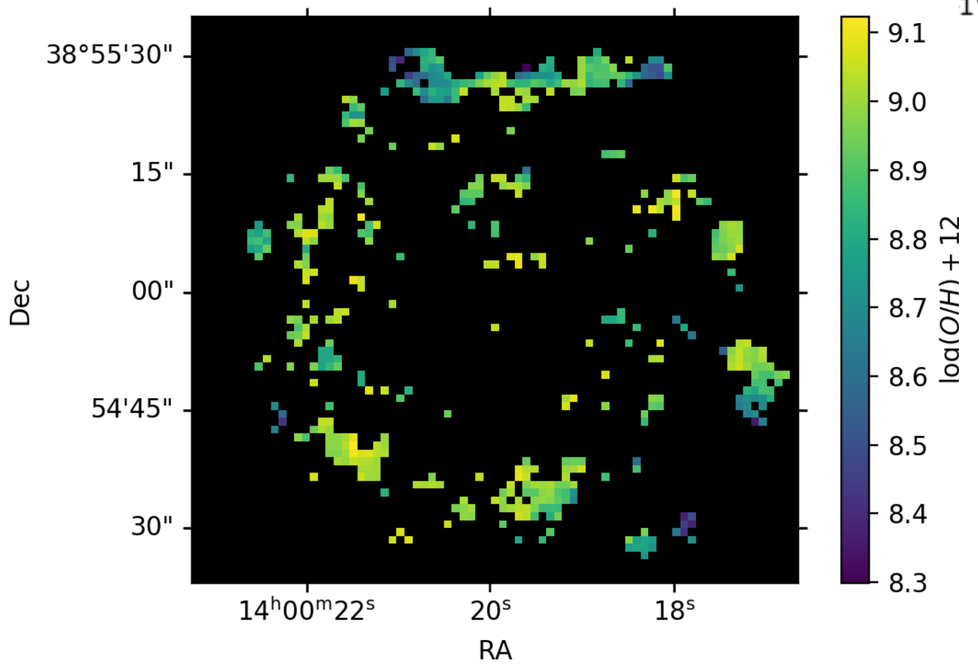
Over time, our pictures of galaxies have gotten better...

M101 - Searle, 1971



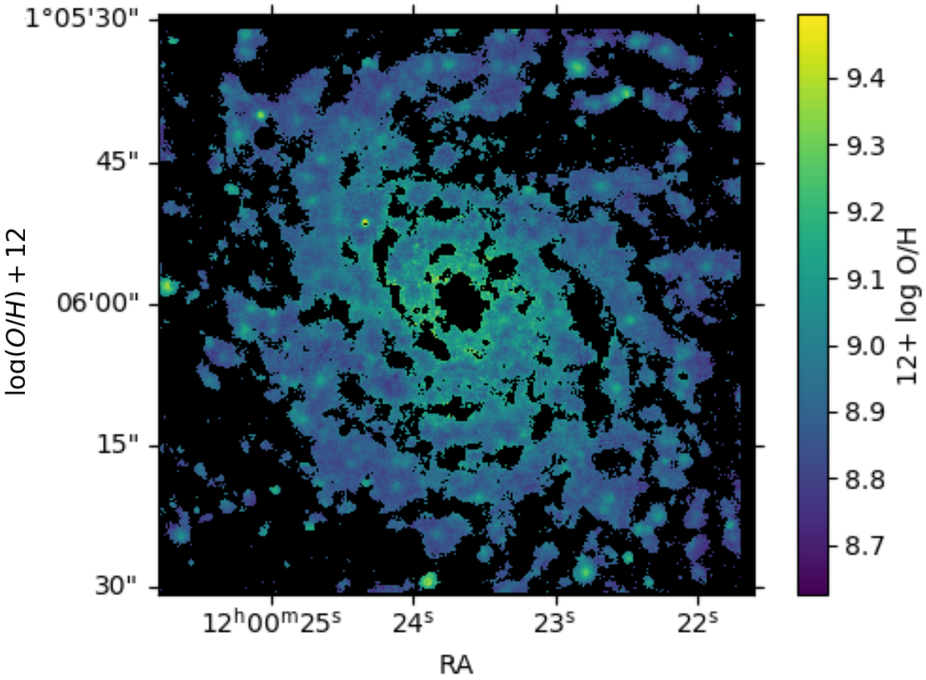
~10 data points per galaxy

NGC5406 - CALIFA, 2014



Resolve down to ~1 kpc scales

NGC4030 - MUSE, 2019

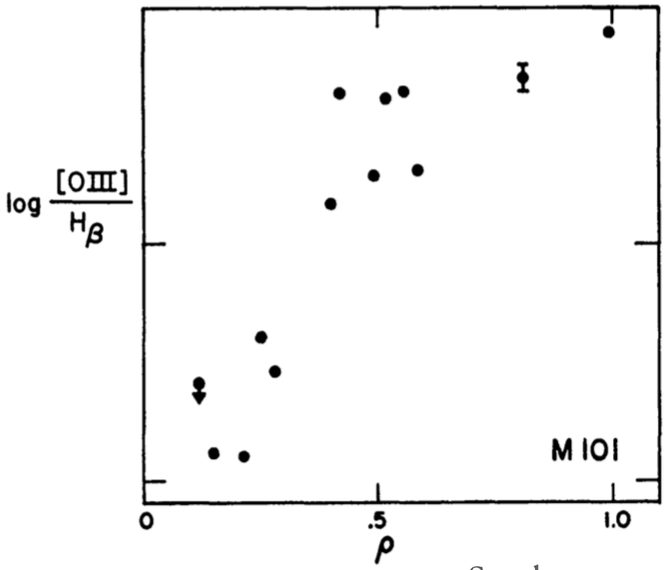


Resolve down to ~50 pc scales

But our data analysis has not improved

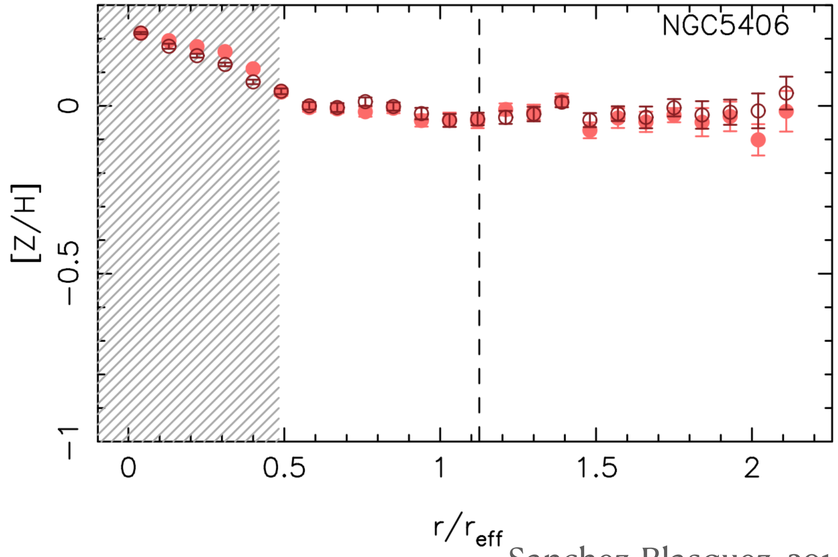
(linear regression to find metallicity vs radius)

M101 - Searle, 1971



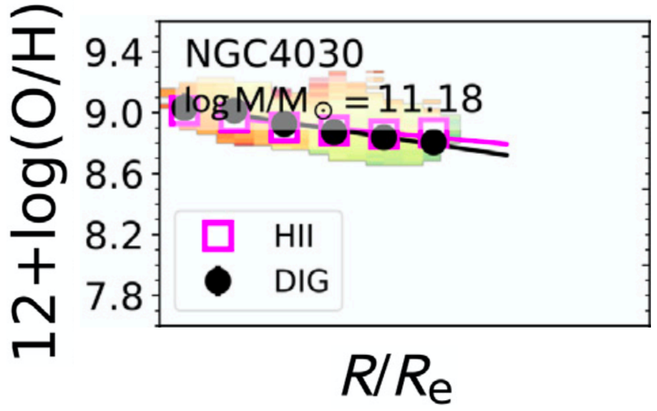
Searle, 1971

NGC5406 - CALIFA, 2014



Sanchez-Blasquez, 2014

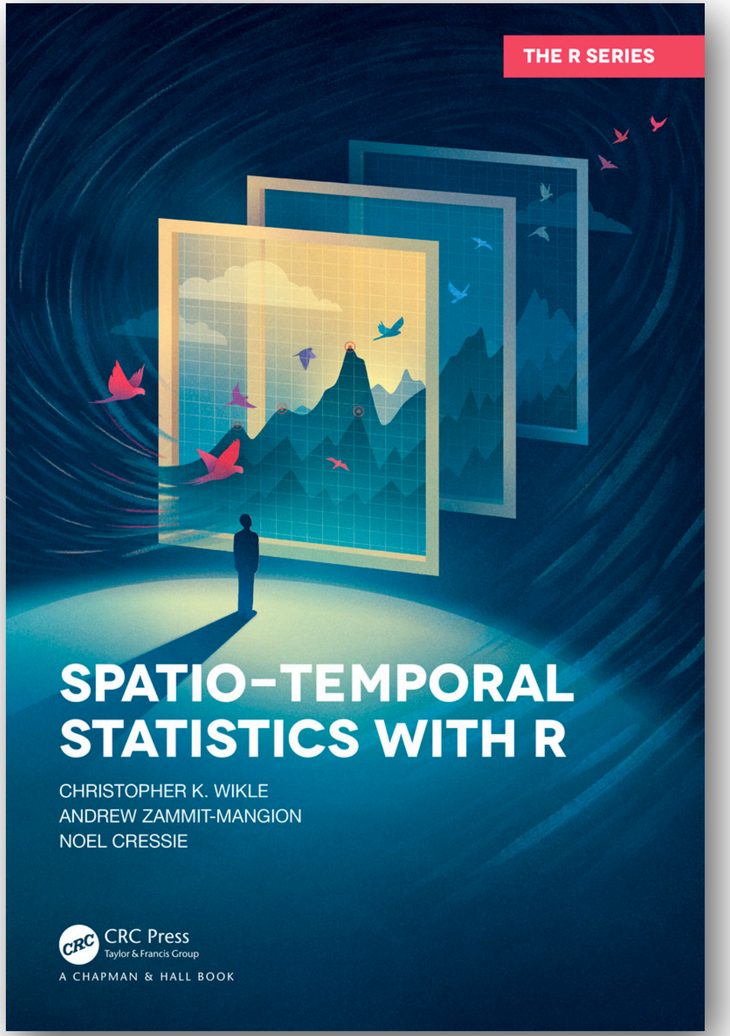
NGC4030 - MUSE, 2019



Erroz-Ferrer, 2019

Spatio-Temporal Statistics

- Studies random fields that vary over space/time in a predictable way.
- Three goals:
  - **Prediction** at unmeasured locations
  - **Forecasting** future values
  - **Scientific inference**: What makes the process happen?
- Modern mathematics; invented in the 1960s.
  Still being improved today



Hierarchical Modelling

What do you need to model a spatially-varying, random process?

$$Z_{obs}(x) = Z_{true}(x) + \epsilon(x)$$

Observation error:
uncorrelated white noise

$$Z_{true}(x) = \mu(x) + \eta(x)$$

Process mean:
accounts for metallicity gradient

Random fluctuations:
nearby points are correlated.

If we can model $\mu(x)$, $\eta(x)$, and $\epsilon(x)$, then we can model the galaxy!

Measurement error, $\epsilon(x)$

- **Large errors are associated with metallicity measurements**
- Amount of error known from S/N of data
- Safe to assume that observation errors are not spatially correlated.

Process mean, $\mu(x)$

- Model using additive linear functions:

$$\mu(x) = \beta_0 + \beta_1\varphi_1(x) + \cdots + \beta_p\varphi_p(x)$$

- $\varphi(x)$: something that Z depends on that varies over space
- Possible $\varphi(x)$ include: r, SFR, density, whether or not we're in a spiral arm, etc.
- For now, just use $Z = \beta_0 + \beta_1 r$ (metallicity gradient model)

Small-scale random fluctuations, $\eta(x)$

- This is hard to model!
- Luckily, someone has already tried:

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **475,** 2236–2252 (2018)
Advance Access publication 2017 December 22

doi:10.1093/mnras/stx3286

Metallicity fluctuation statistics in the interstellar medium and young stars – I. Variance and correlation

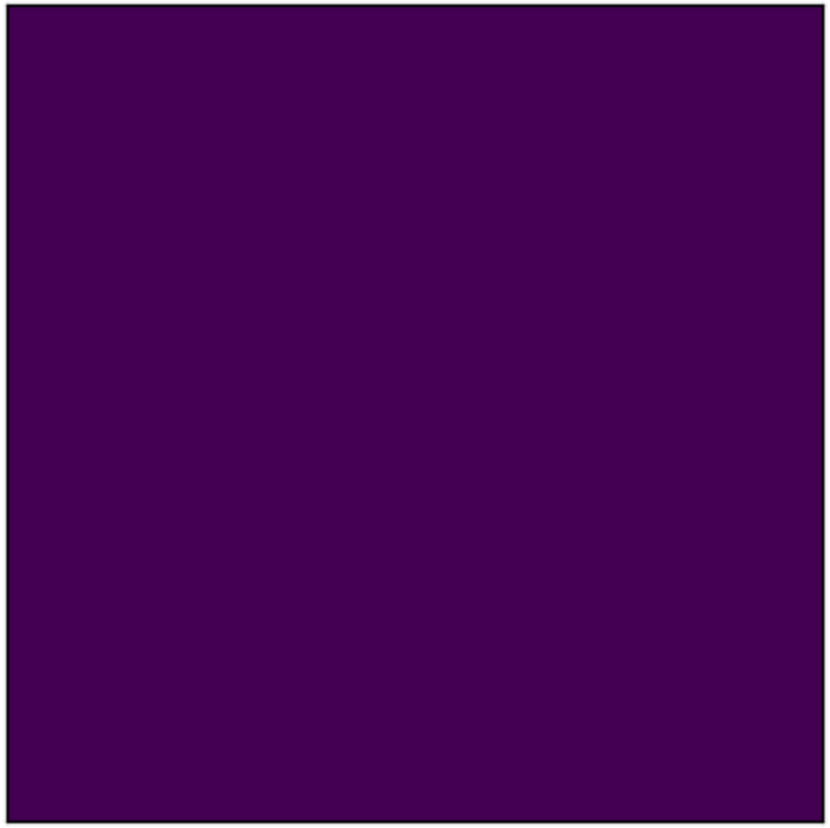
Mark R. Krumholz[1★] and Yuan-Sen Ting[1,2,3,4]

[1]*Research School of Astronomy & Astrophysics, Australian National University, Canberra, 2612 ACT, Australia*
[2]*Institute for Advanced Study, Princeton, NJ 08540, USA*
[3]*Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*
[4]*Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA*

Accepted 2017 December 18. Received 2017 December 17; in original form 2017 August 14

Overview: the KT model

- Stochastic partial differential equation
- Assumptions:
  - Constant, uniform star formation history
  - **Linear diffusion**
  - No stellar winds
- Outputs:
  - Correlation structure for $\eta(x)$



Overview: the KT model

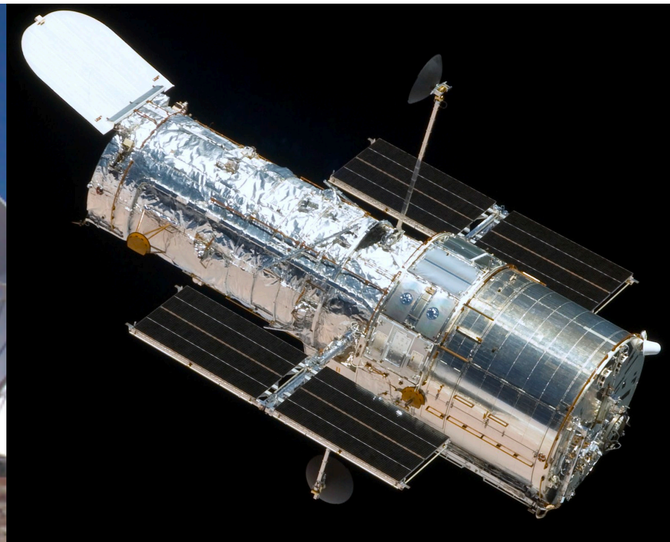
5 Parameters, but none are free

Model parameter	Explanation	Depends on
Supernova Rate	Rate of supernovas	SFR
Star Formation Timescale	How long has the galaxy been producing stars/metals for?	SFR, M_*
Diffusion Coefficient	How fast do metals spread within a galaxy?	Gas density, gas velocity, gas fraction
Injection Width	How far do supernovae expand before mixing into the ISM?	Gas velocity, H density
Injection Mass Variance	How much does the amount of metals added by a supernova vary?	—

Introducing: PHANGS!

The data

- **P**hysics at **H**igh **A**ngular resolution in **N**earby **G**alaxies **S**urvey
- Multi-wavelength campaign using VLT/MUSE, ALMA, and the HST
- Aim: to understand the interplay of the small-scale physics of gas and star formation with galactic structure and galaxy evolution



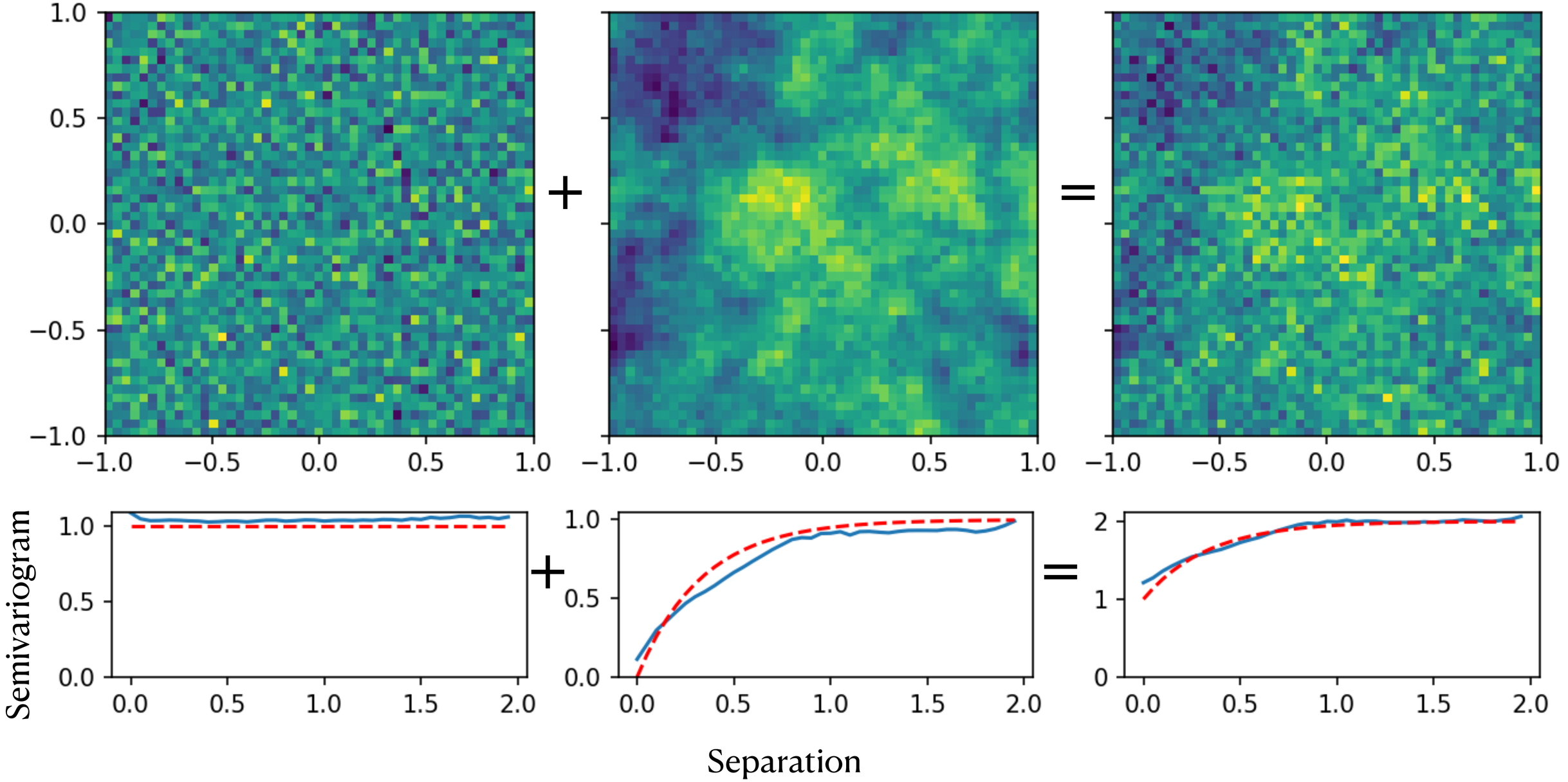
Observationally, how can you separate $\epsilon(x)$ from $\eta(x)$?

The Semivariogram

- How does **variance in data points** depend on **separation between data points?**

$$\gamma(h) = \frac{1}{2}\text{Var}\left(Z(\vec{x}) - Z(\vec{y})\right)$$

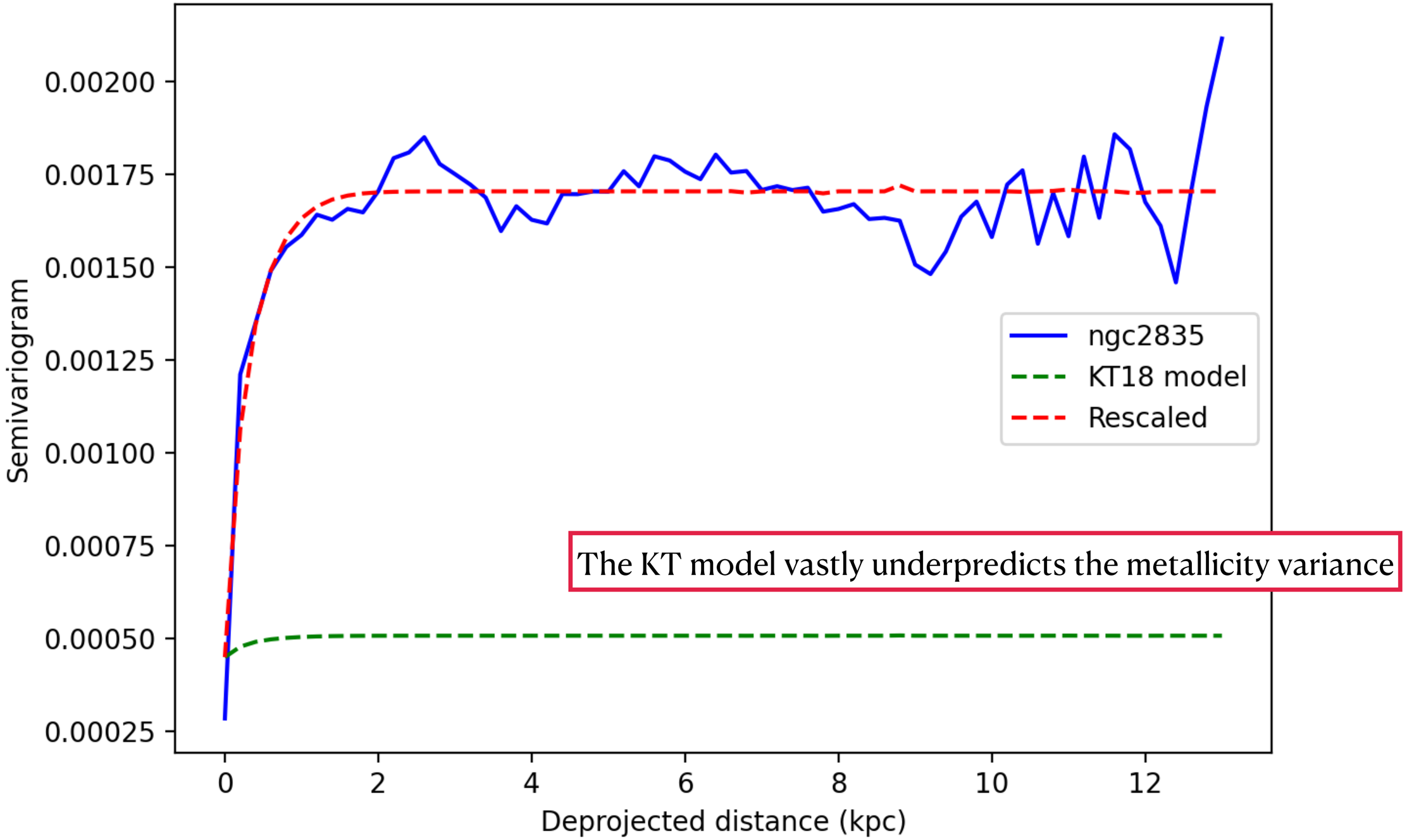
$$h - \delta h \leq |\vec{x} - \vec{y}| \leq h + \delta h$$

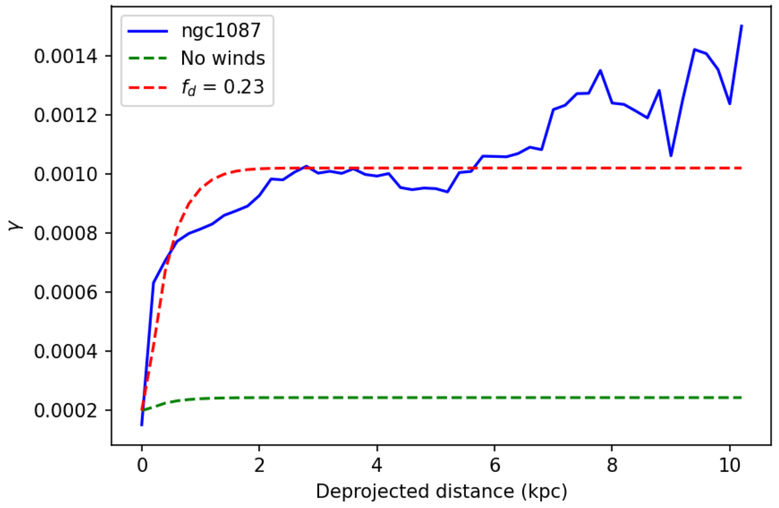

1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
+
=
Semivariogram
2
1
0
Separation

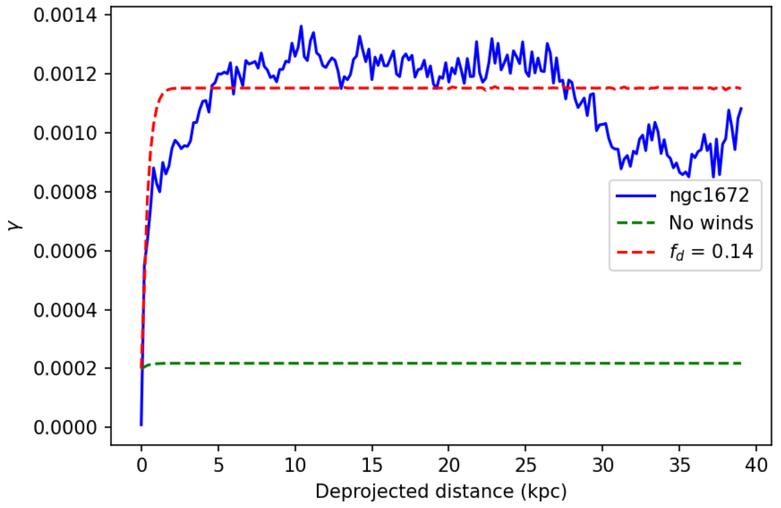
Do the semivariograms predicted by KT18 match those observed in real galaxies?

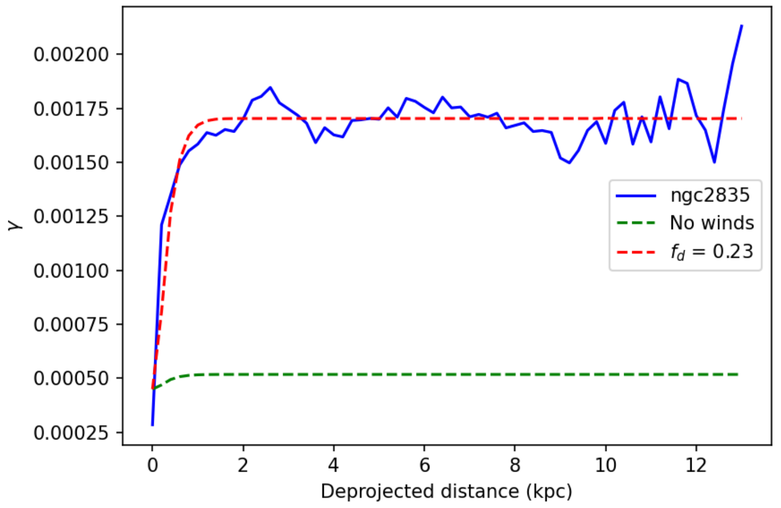
Methods

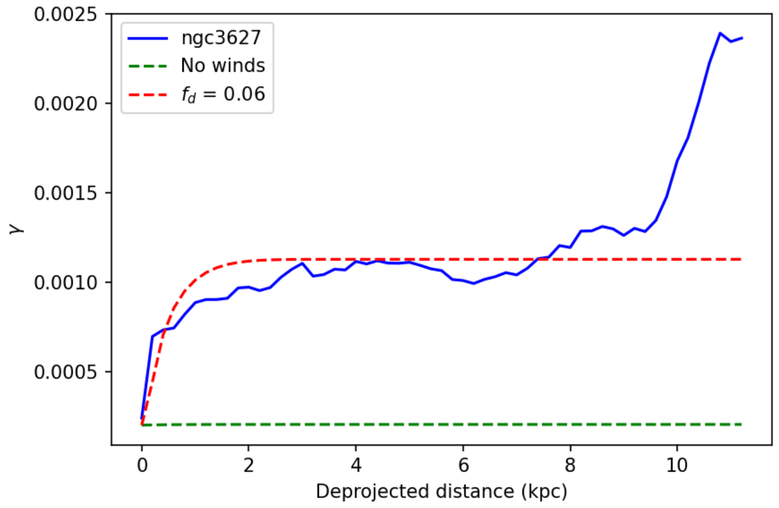
- Using global galaxy properties, set KT model parameters.
- Compute theoretical semivariogram from KT model
- Add variance due to observation error
- Accounting for this error, fit metallicity gradient using GLS
- Compare the semivariogram of the residual small-scale variation to the theory...

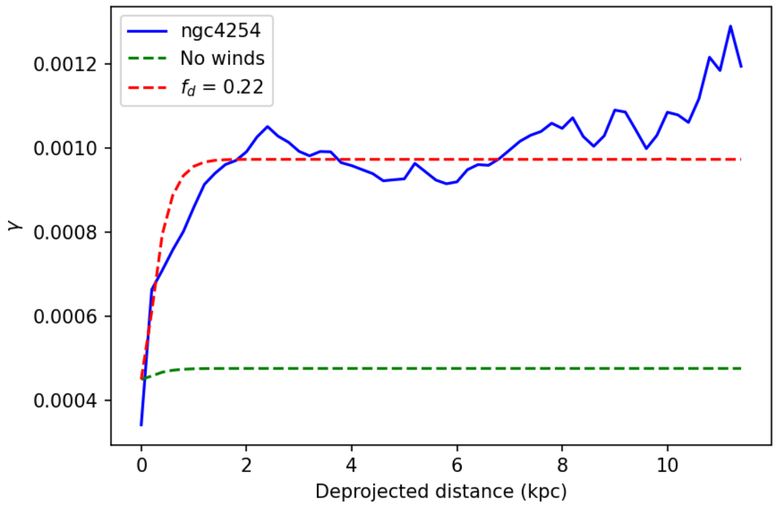

0.00200
0.00175
0.00150
0.00125
0.00100
0.00075
0.00050
0.00025
Semivariogram
0
2
4
6
8
10
12
Deprojected distance (kpc)
ngc2835
KT18 model
Rescaled
The KT model vastly underpredicts the metallicity variance

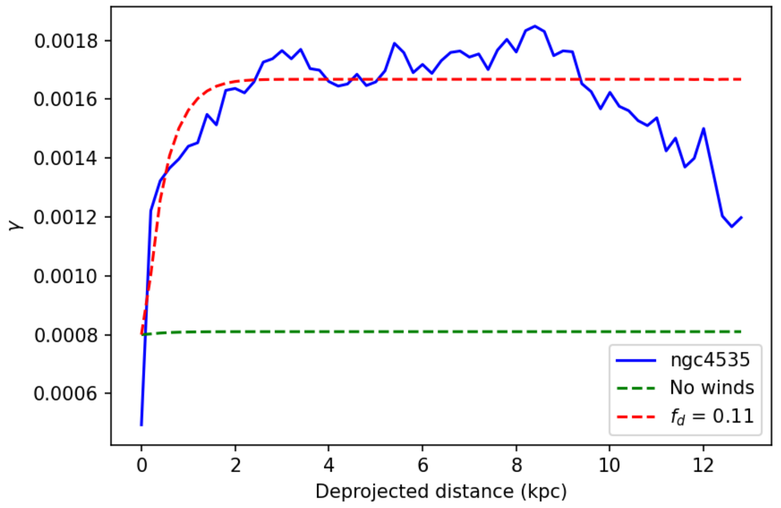

ngc1087
No winds
f_d = 0.23
γ
Deprojected distance (kpc)


ngc1672
No winds
f_d = 0.14
γ
Deprojected distance (kpc)


ngc2835
No winds
f_d = 0.23
γ
Deprojected distance (kpc)


ngc3627
No winds
f_d = 0.06
γ
Deprojected distance (kpc)


ngc4254
No winds
f_d = 0.22
γ
Deprojected distance (kpc)


ngc4535
No winds
f_d = 0.11
γ
Deprojected distance (kpc)

Why won't it fit?

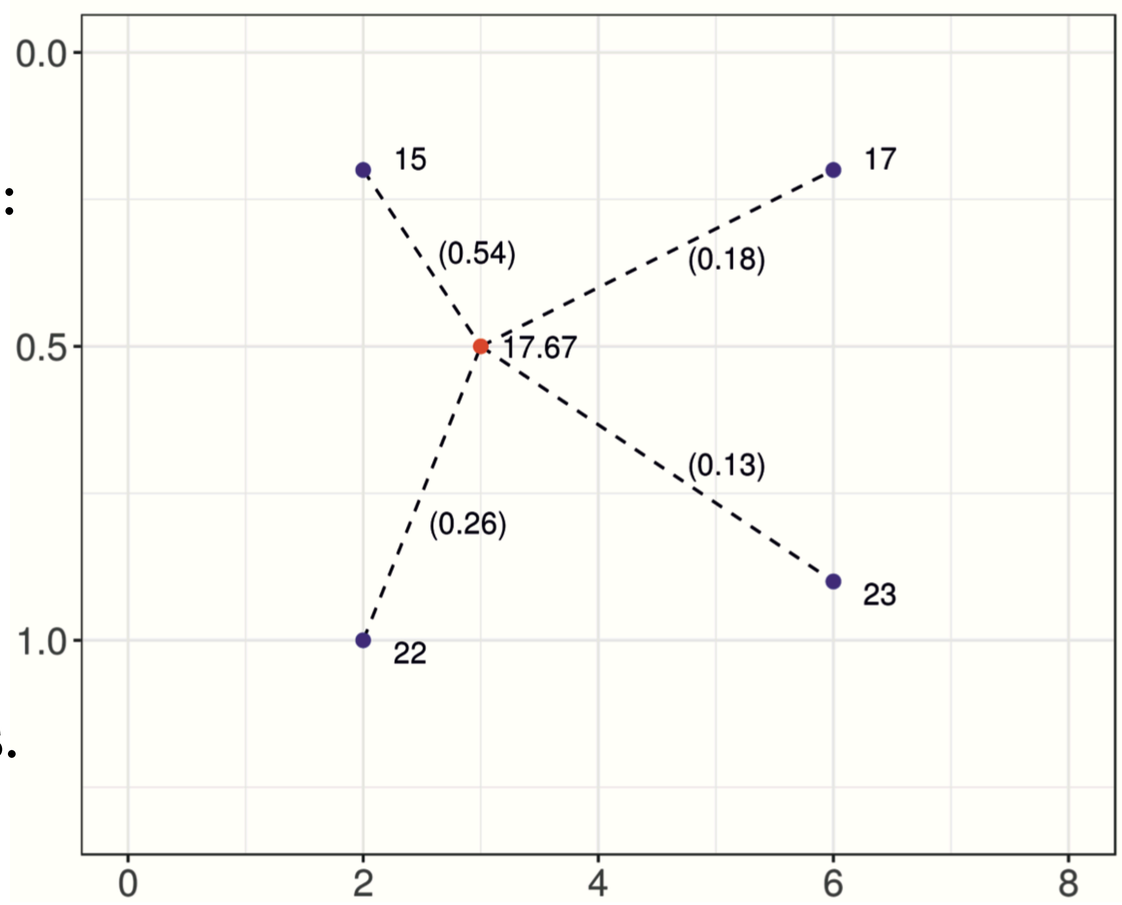
Some ideas; but here is my favourite

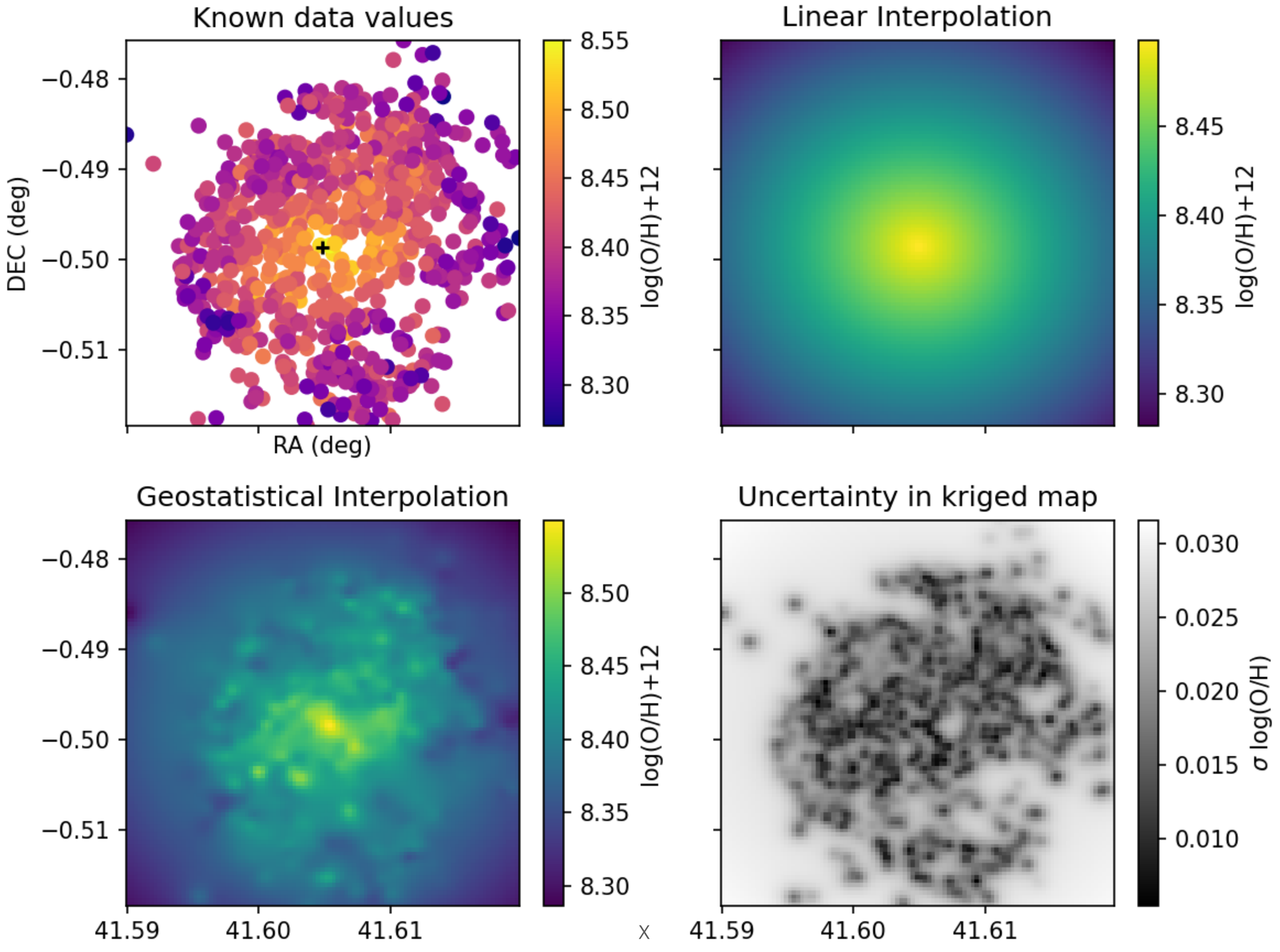
- In the KT model, turbulence is treated as linear diffusion
- This may lead to overestimating amount of smoothing (Pan + Scannapieno, 2010)
- Turbulence matters! Can allow small-scale inhomogeneities to persist for a long time before getting washed out (de Avillez + Mac Low, 2002)
- Turbulence is complicated! May lead to model for which semivariogram cannot be calculated analytically :(

Kriging

This model is not useless!

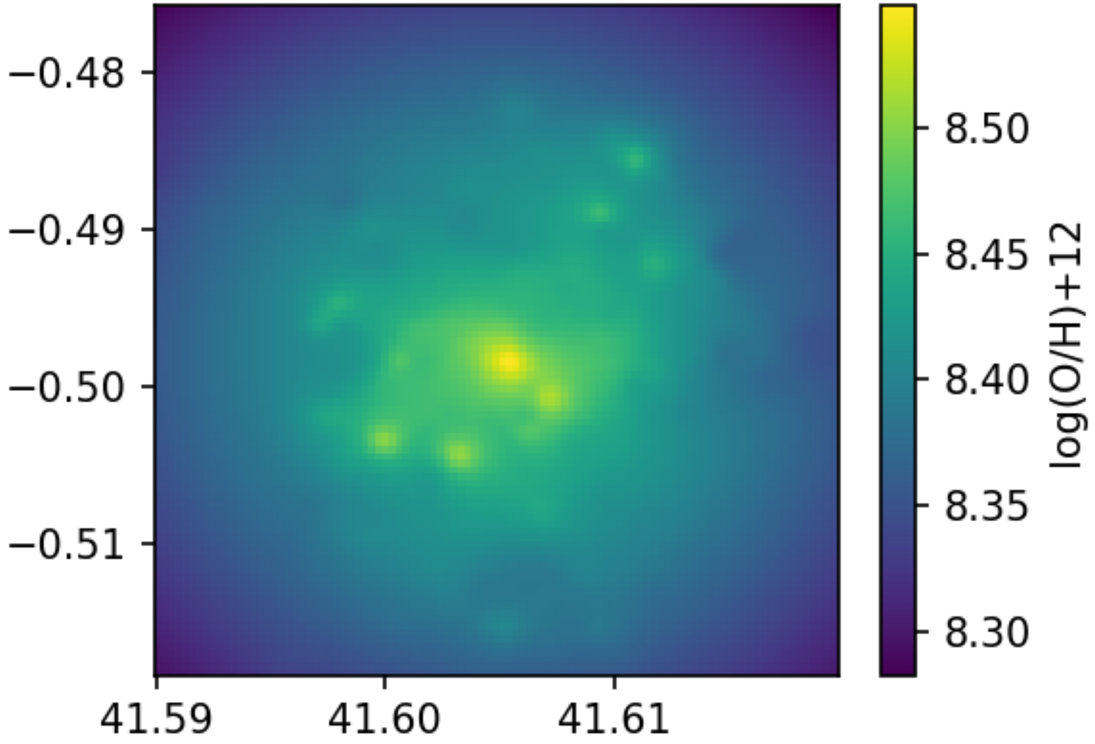
- In galaxies, metallicity can't be measured:
  - around AGN
  - in diffuse regions
- The KT model cannot explain the variance
- But it can be used, after rescaling, to **predict metallicities** at unknown points.



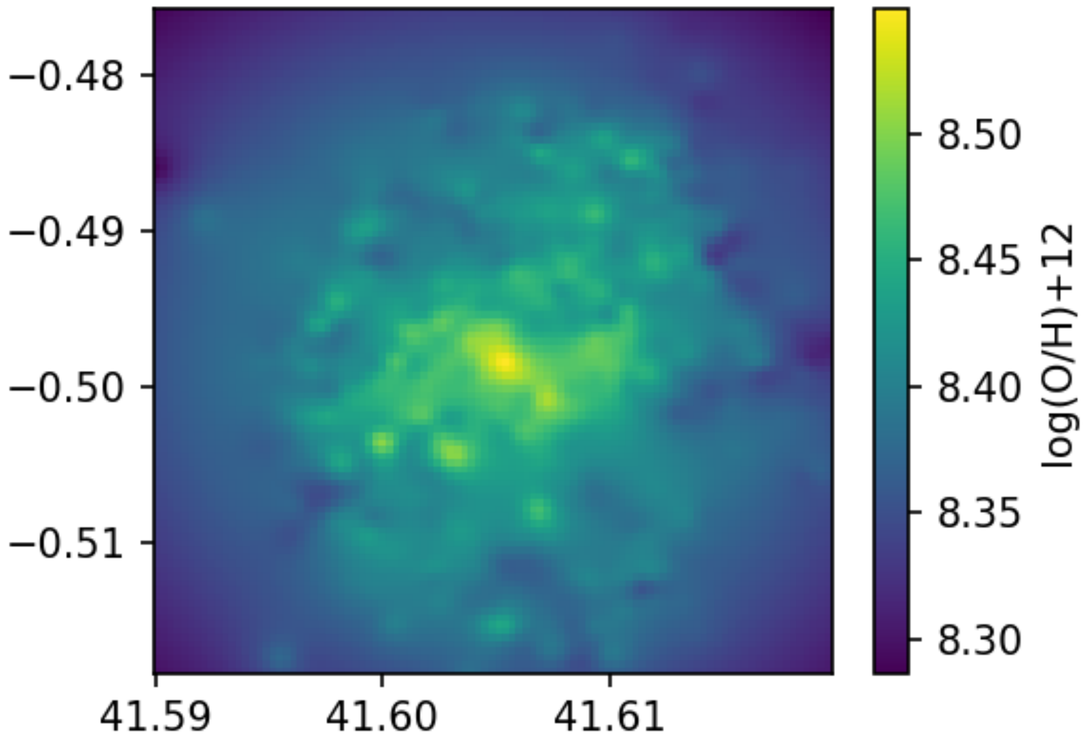

Known data values
Linear Interpolation
Geostatistical Interpolation
Uncertainty in kriged map
DEC (deg)
RA (deg)
−0.48
−0.49
−0.50
−0.51
41.59
41.60
41.61
log(O/H)+12
8.55
8.50
8.45
8.40
8.35
8.30
σ log(O/H)
0.030
0.025
0.020
0.015
0.010

Rescaling matters!

Before rescaling



Rescaled to Best Fit

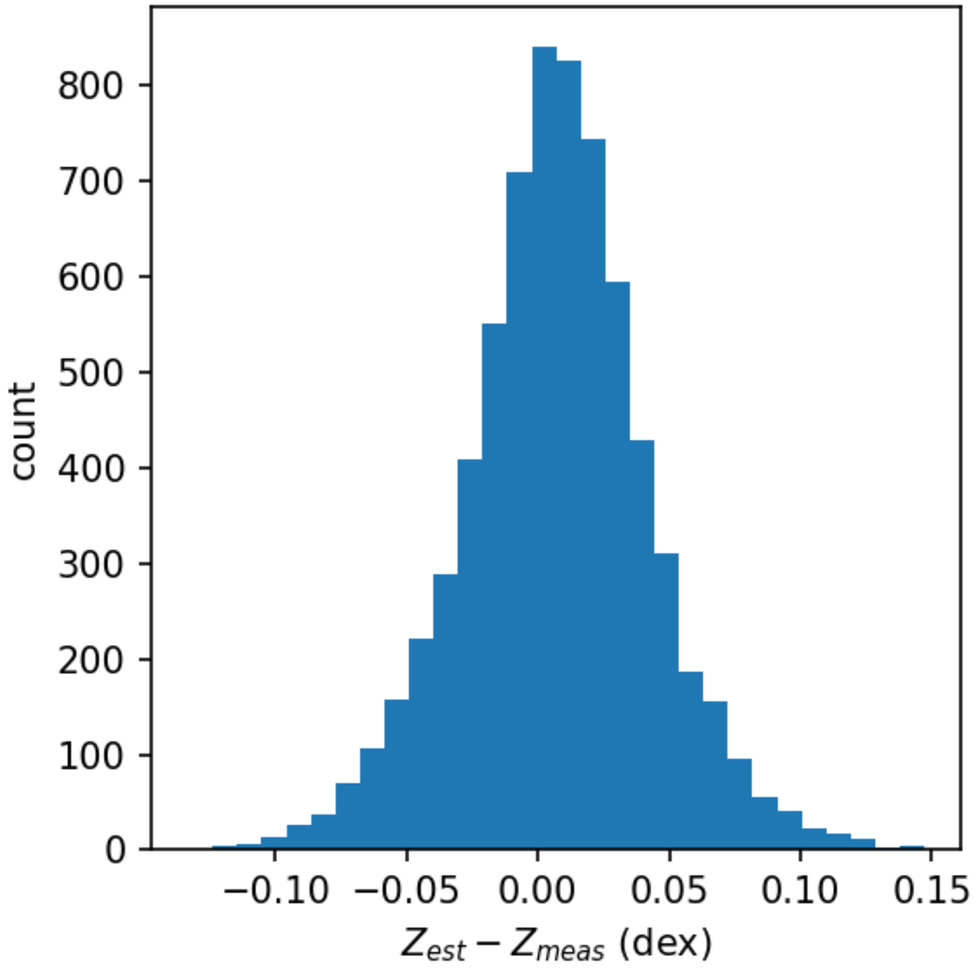


X

But are these predictions valid?

Introducing K-fold cross-validation

- Split data into 10 bins
- For each bin:
  - Fit model using other 9 bins (training set)
  - Compare predictions to excluded bin (testing set)



Conclusions

Do you like stats? Come talk to me!

- High-resolution datacubes contain lots of information
- Geostatistical methods can separate uncorrelated measurement error from true correlated fluctuations
- The KT model for metal transport underpredicts small-scale metallicity variance
- **Good stats matters!**