PRINCETON UNIVERSITY

Galaxy modelling for all-sky multiwavelength surveys

Dan Taranu & LSST Data Management

Large Synoptic Survey Telescope

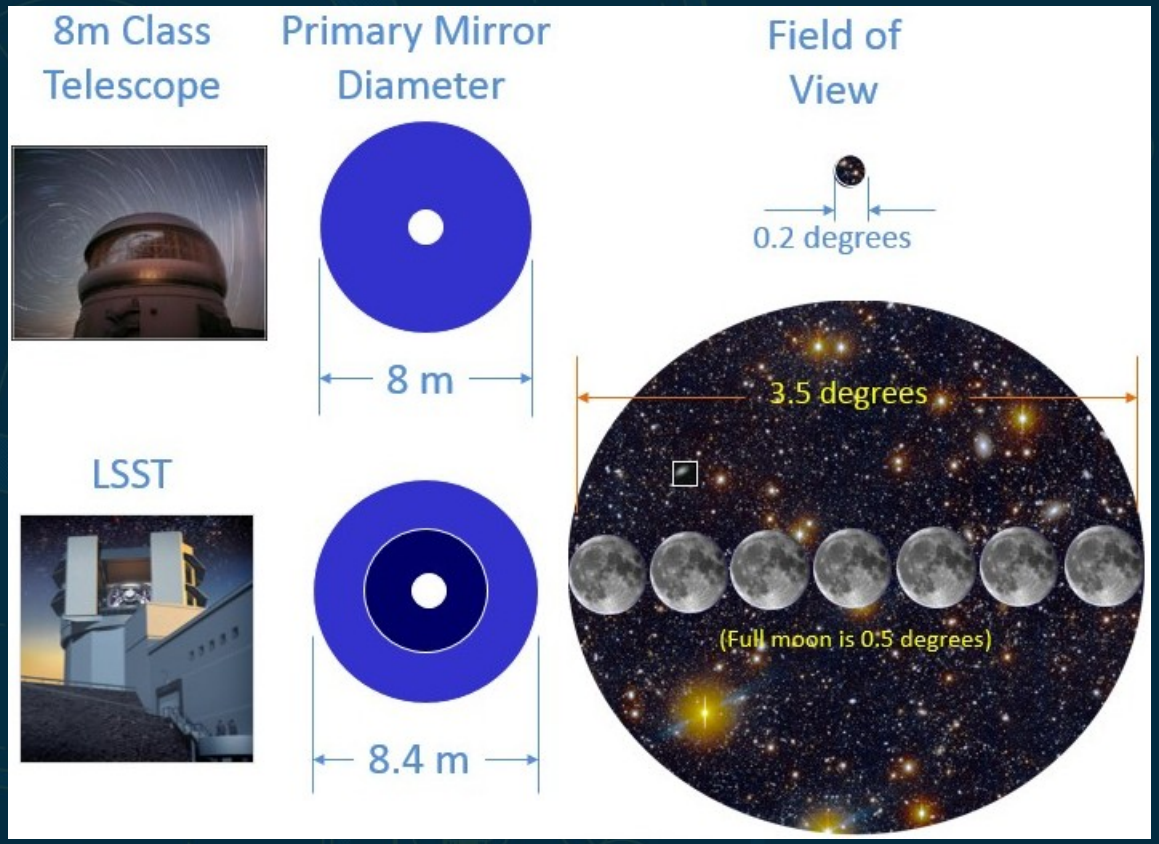
Large Synoptic Survey Telescope

Large Synoptic Survey Telescope

lsst.org/scientists/keynumbers

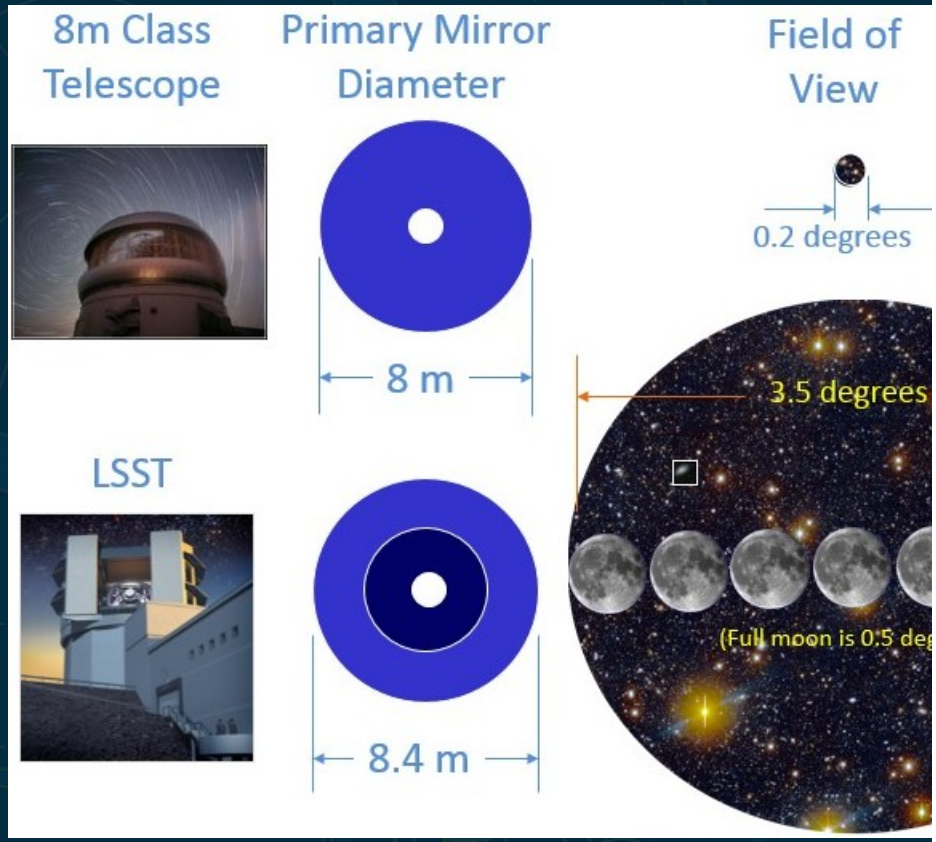
lsst.org/science

lsst.org/participate



Large Synoptic Survey Telescope

lsst.org/scientists/keynumbers
lsst.org/science
lsst.org/participate



Optical System:

Etendue (AΩ) = 319 meter^{2}degrees^2
Field of View = 3.5 degrees (9.6 square degrees)
Primary mirror diameter = 8.4 m
Effective clear aperture = 6.68 m (on-axis)
Mean effective aperture = 6.423 m (area weighted over FOV)
Final f-ratio = f/1.234
Plate scale = 50.9 microns/arcsec

Imaging System:

Pixel count = 3.2 Gpixels
Readout time = 2 sec
Dynamic range = 18 bits
Pixel pitch = 10 micron (0.2 arcsec)

Spectral Response:

Full bandpass range = 320-1050 nm
Total number of filters = 6
Number of active filters = 5
System 50% response points (nm)
u 324 395
g 405 552
r 552 691
i 691 818
z 818 921
y 922 997

Survey Area:

Fiducial main survey area = 18,000 deg2
Total covered area (est.) = 25,000 deg^2
(main survey + special programs)
Fiducial number of visits per pointing in the main survey = 825

Network bandwidths:

Summit (Pachón) to Base (La Serena) = 600 Gbps
Base (La Serena) to Archive (NCSA) = 2 x 100 Gbps

Observation Properties:

The standard visit = 2 x 15 sec. exposures
Exposure sequence = 1 s open shutter + 14 s dwell + 1 s close shutter = 16 s
Visit sequence = 16 s exp + 2 s readout + 16 sec exp = 34 s (second readout concurrent with slew)
Median slew time between visits = 4.8 s
Mean slew time between visits = 6.8 s
Visits per night = "about a 1000"
Calibration exposures = 450/day
Data collected per 24 hr period = "about 20 TB"

Data Releases:

Number of Data Releases = 11
Date of DR1 release = Date of Operations Start+ 12 months
Estimated numbers for DR-1 release
Objects = 18 billion
Sources = 350 billion (single epoch)
Forced Sources = 0.75 trillion

Estimated numbers for DR-11
Objects = 37 billion
Sources = 7 trillion (single epoch)
Forced Sources = 30 trillion
Visits observed = 2.75 million
Images collected = 5.5 million

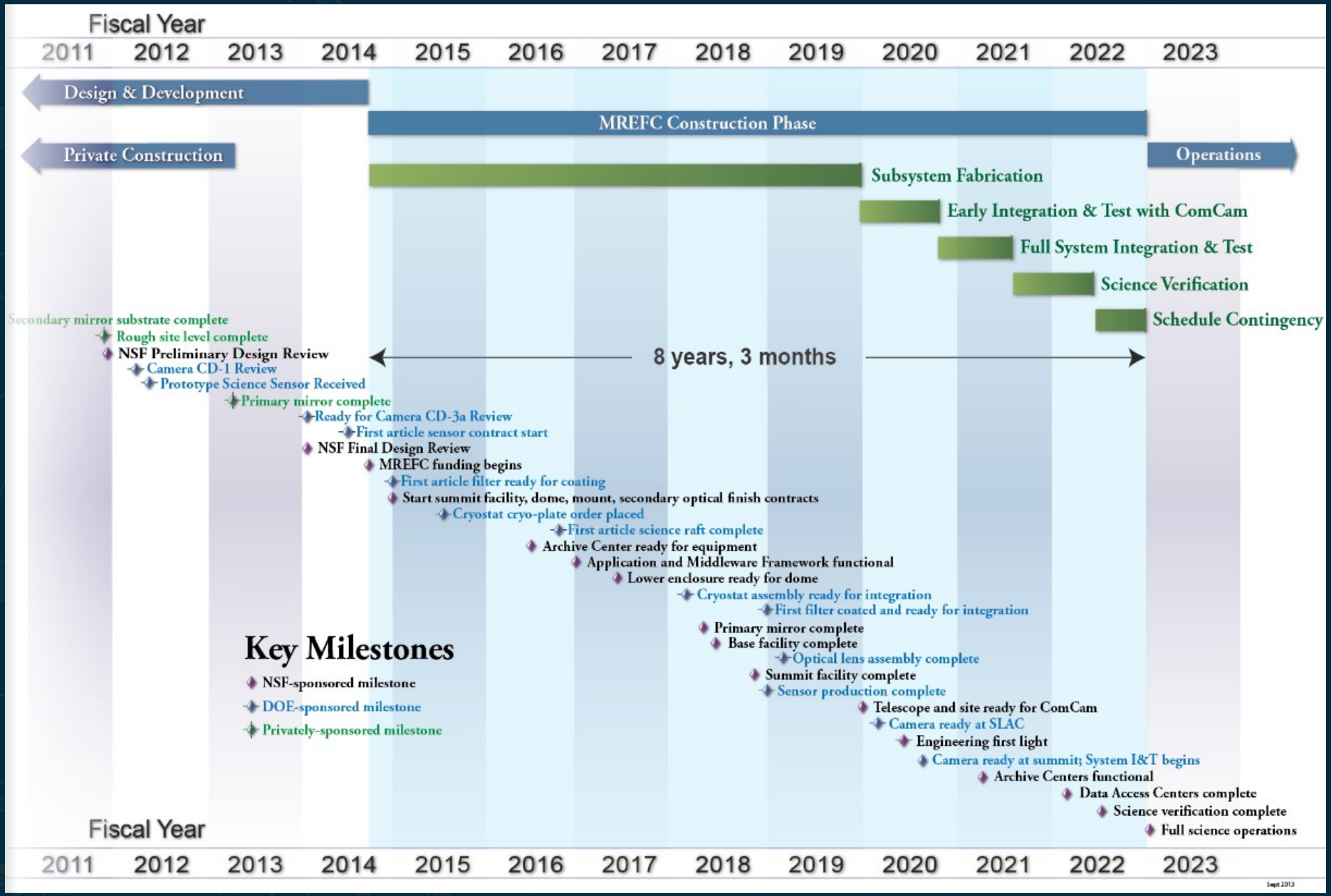
Alert Production:

Real-time alert latency = 60 seconds
Average number of alerts per night= "about 10 million"

Data and compute sizes:

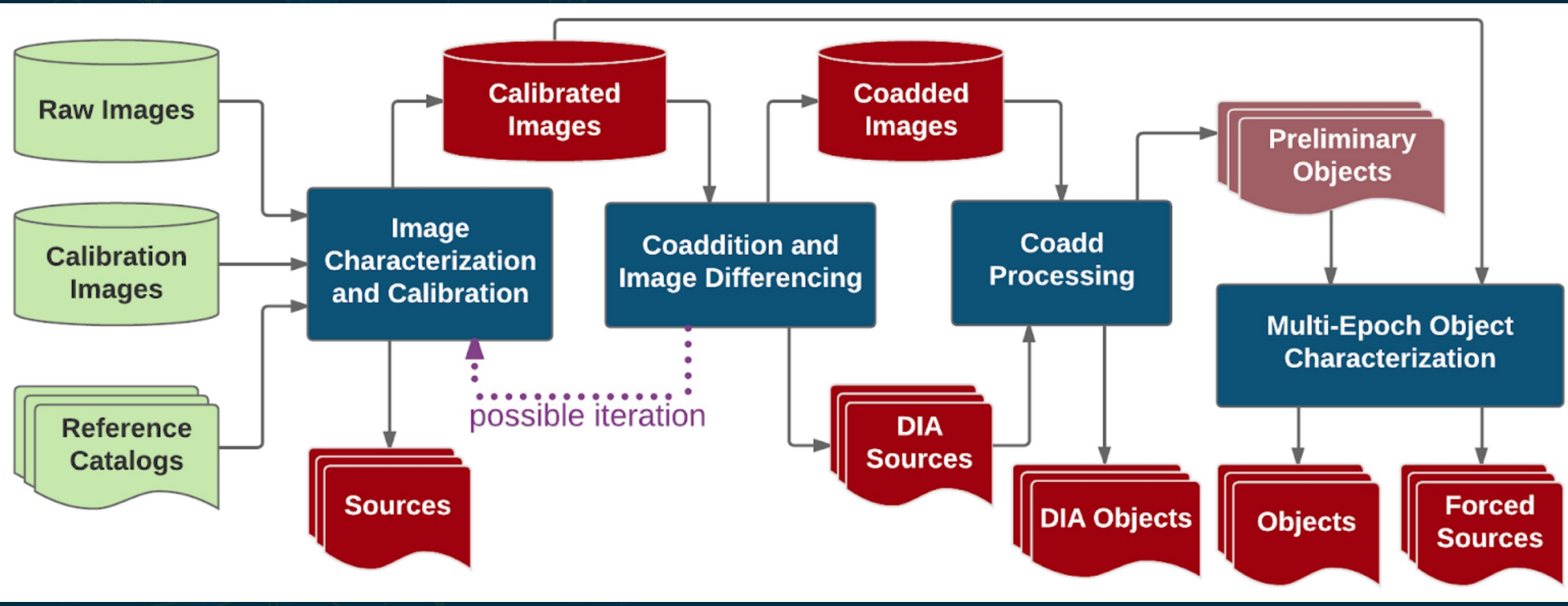
Final image collection (DR11) = 0.5 Exabytes
Final catalog size (DR11) = 15 PB
Final disk storage = 0.4 Exabytes
Peak number of nodes = 1750 nodes
Peak compute power in LSST data centers = 1.8 PFLOPS

Construction On Track



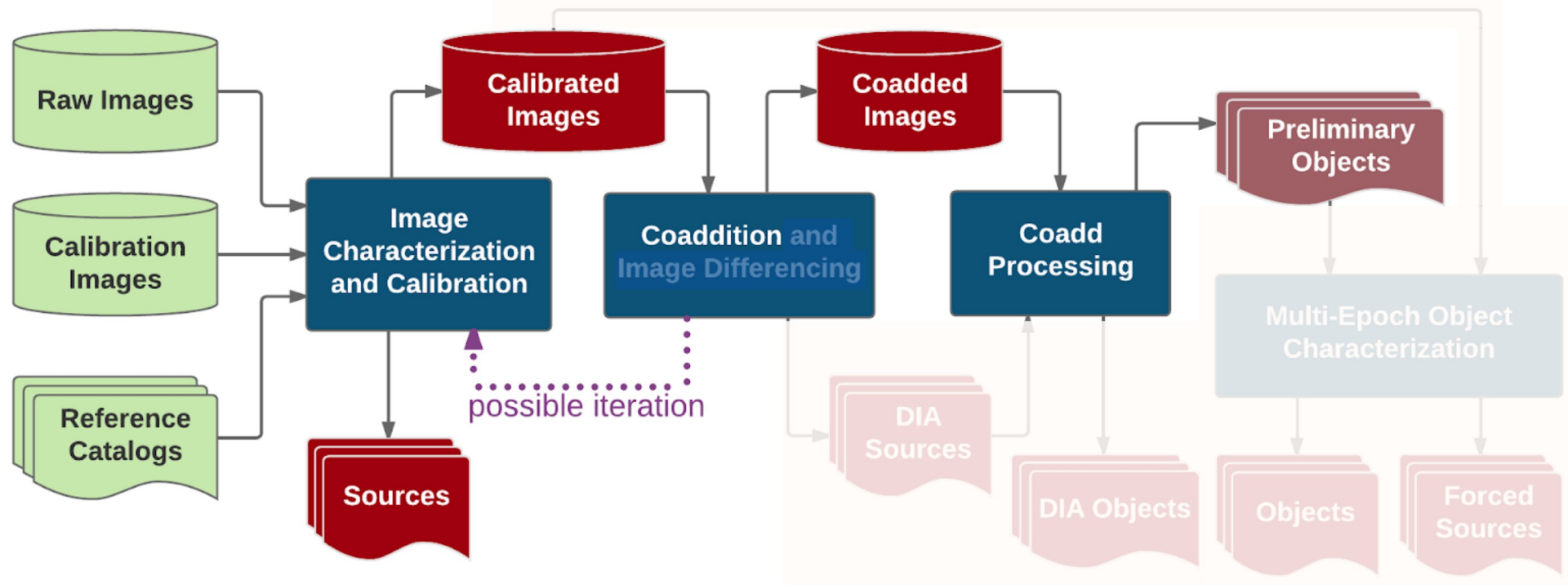
Science Pipelines Plan

- Data Management: dm.lsst.org
- Data Release Production – annual re-processing
- Current official plan – multi-epoch source modelling

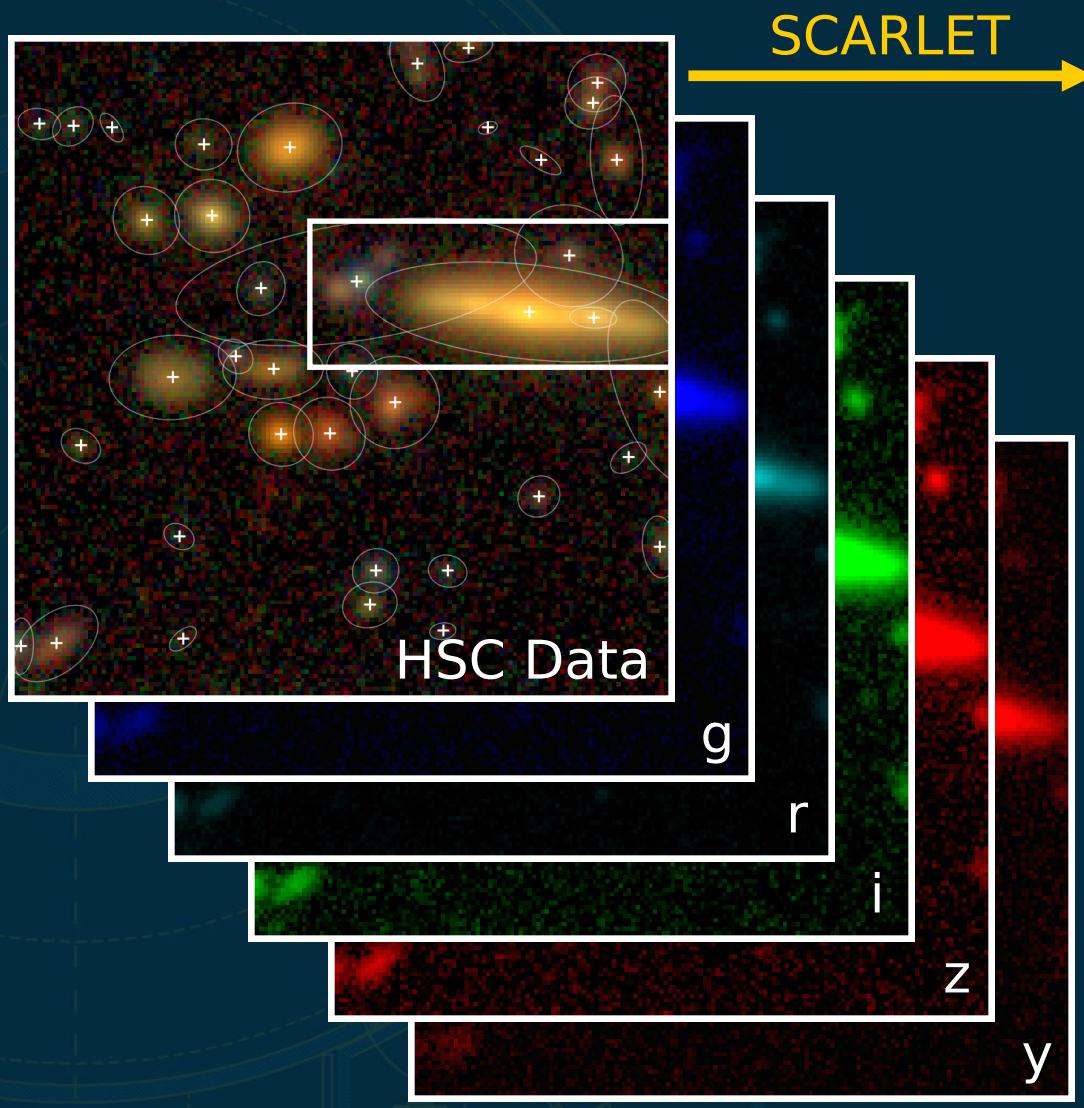


Science Pipelines Today

- Largely improvements on SDSS pipeline
- Used to process Subaru-HSC data; see Bosch+18
- Galaxy modelling on co-added single-band images



New Deblender - Scarlet



Non-parametric SED & morphology solved via matrix factorization; see Melchior+18

$$\text{scene} = \sum_k \text{SED}_k \times \text{Morphology}_k + \text{noise}$$

github.com/fred3m/scarlet

Current Galaxy Models

- Multi-Gaussian/shapelet* PSF
- Three galaxy model fits:
  - Exponential (exp., Sersic n=1): disk-like
  - de Vaucouleurs (dev., Sersic n=4) bulge-like
  - CModel – linear combination of exp+dev
    - Bulge+disk like
- Not true Sersics but multi-Gaussian approximation
  - See Hogg & Lang 13
- Why? Fast & useful approximations

*Gaussians with Hermite polynomial perturbations; see Refregier 01

Re-thinking Galaxy Modelling

- What models should we fit? My biases:
  - add free single Sersic
  - better bulge+disk
  - 3+ component models (e.g. add point source)
- How do we determine the ‘best’ models to fit?
  - Fitting idealized templates not very useful
  - Want real galaxies @ ~infinite S/N, resolution

Re-thinking Galaxy Modelling

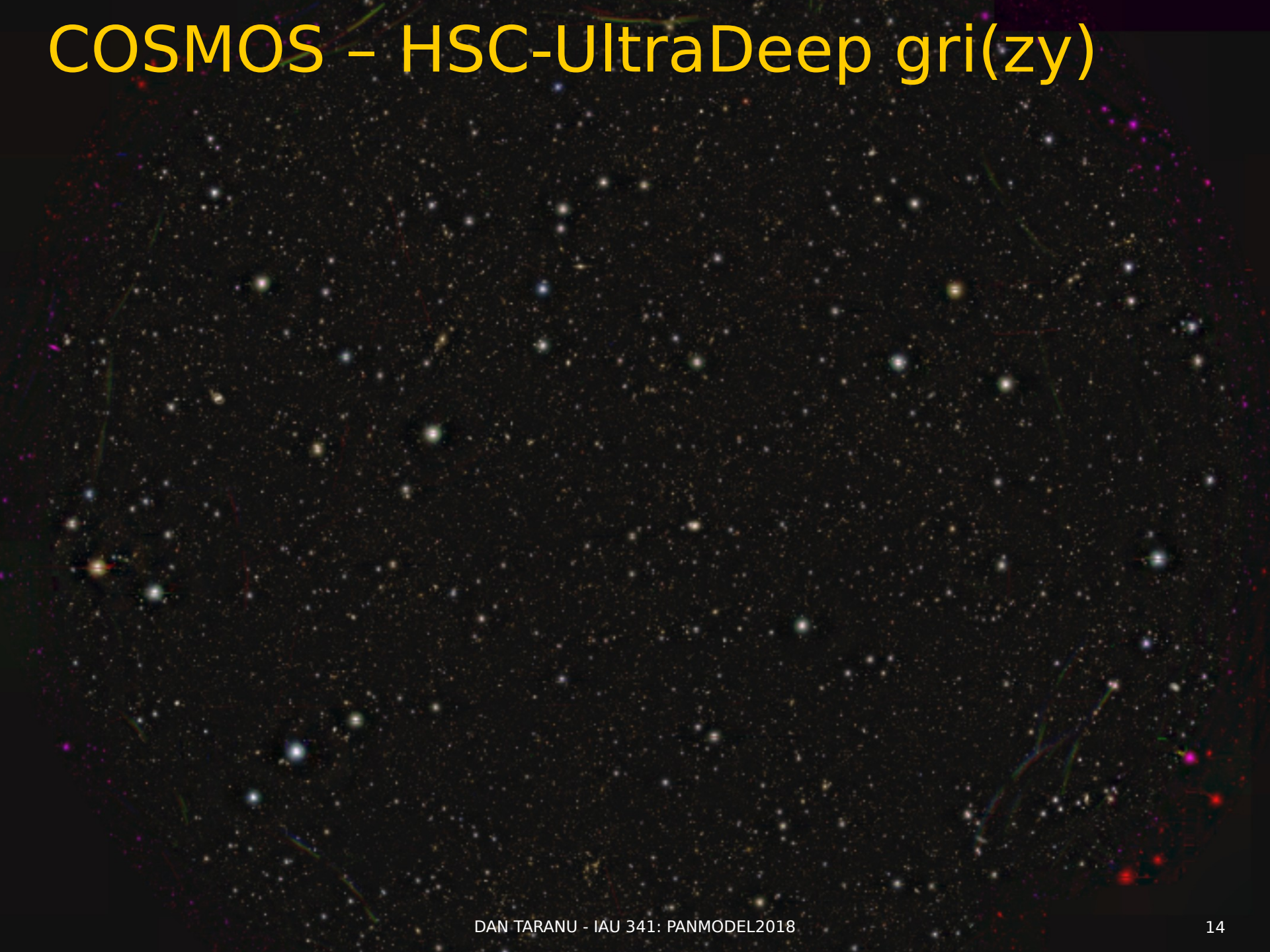
- (How) Should we do:
  - Multi-wavelength
  - Multi-object
  - Multi-epoch

Re-thinking Galaxy Modelling

- (How) Should we do:
  - Multi-wavelength – Yes; free mag. per band
  - Multi-object – Yes; crowded fields?
  - Multi-epoch – No; unless necessary

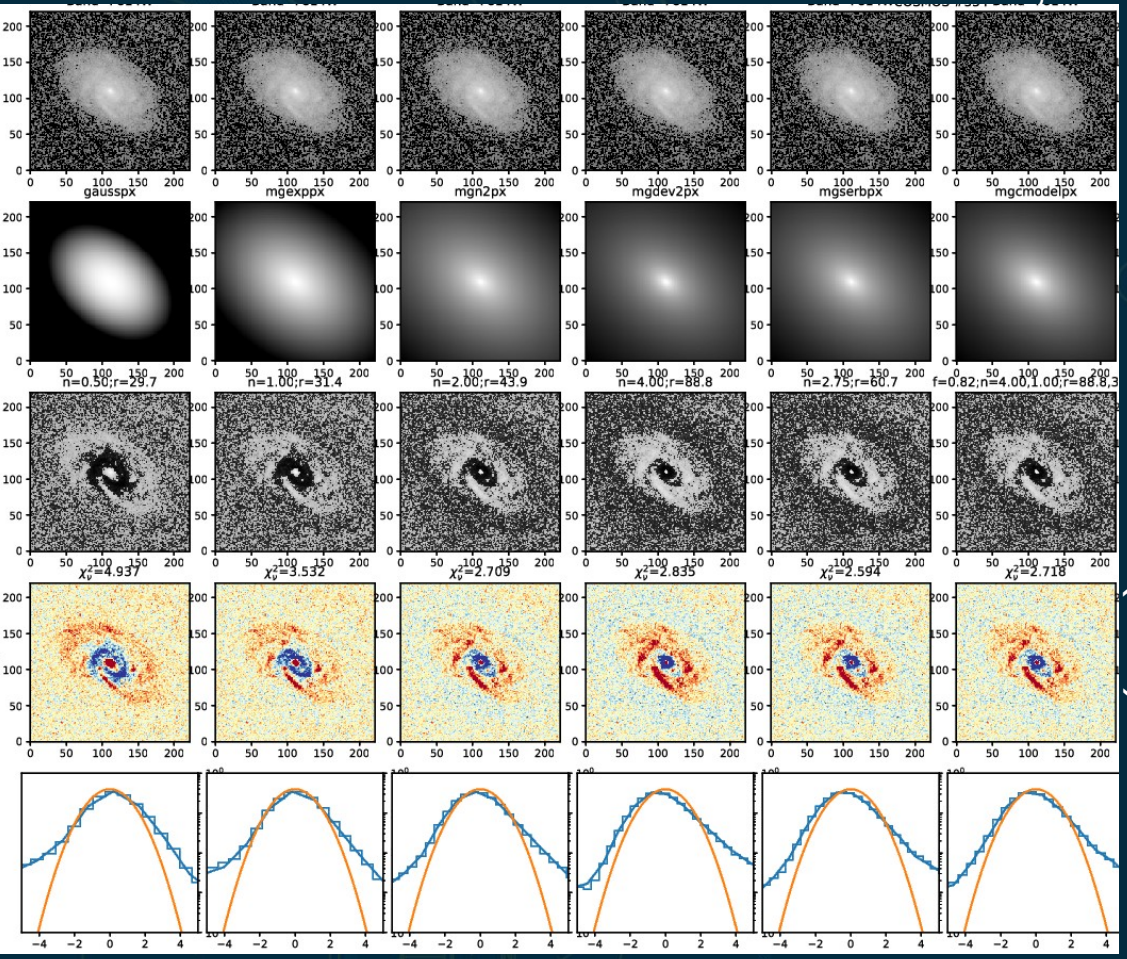
COSMOS – HST F814W (i+z)

COSMOS – HSC-UltraDeep gri(zy)



Galaxy Models

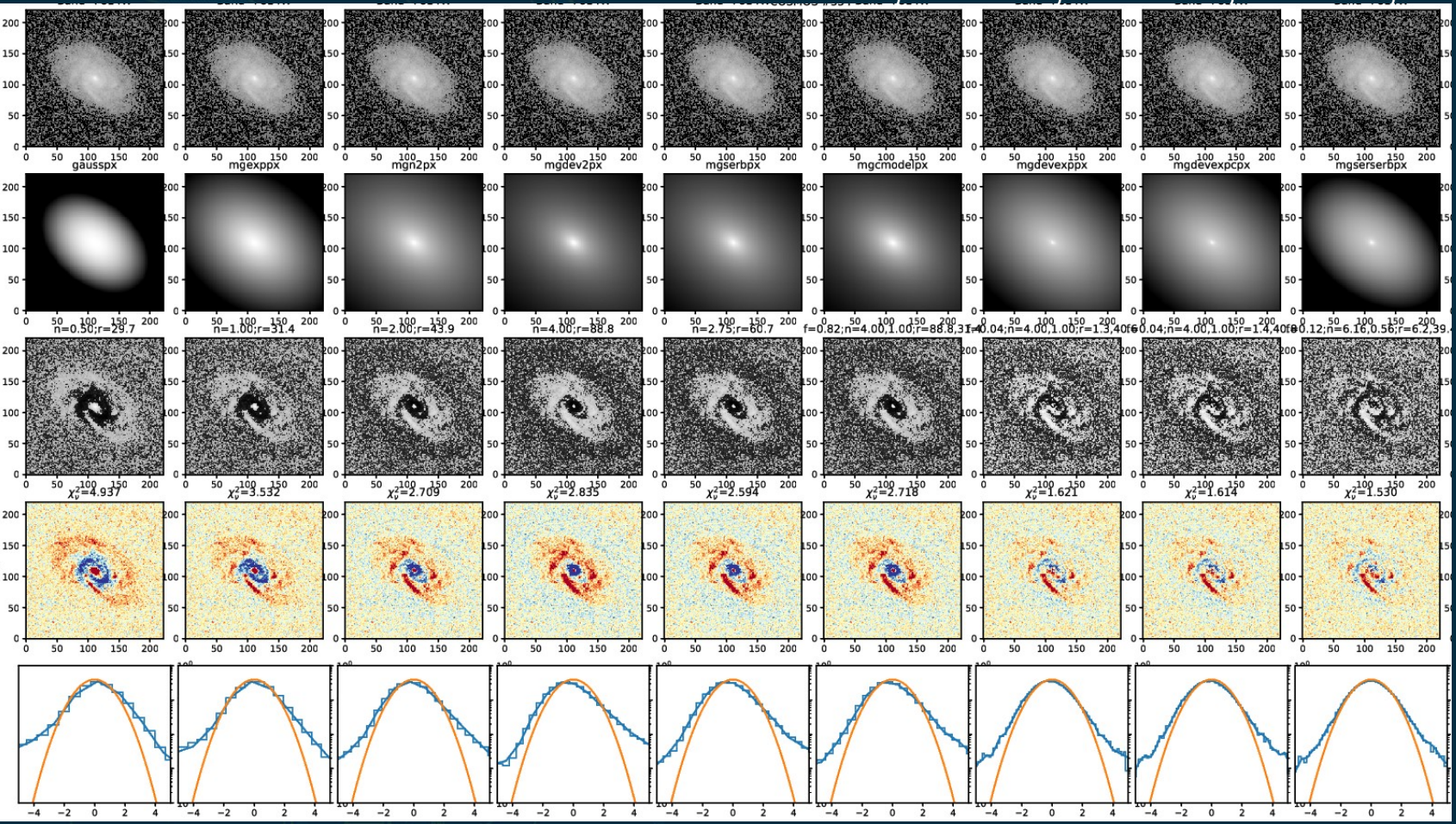
Gauss.	Exp.	n=2	deV.	Sersic	cModel f=0.82
n=0.5	1	2	4	2.75	4, 1
Re=29.7	31.4	43.9	88.8	60.7	89, 31



χ^2_ν=4.937	3.532	2.709	2.835	2.594	2.718

Galaxy Models

Gauss.	Exp.	n=2	deV.	Sersic	cModel f=0.82	devExp f=0.04	devExpc f=0.04	2-Sersic f=0.12
n=0.5	1	2	4	2.75	4, 1	4, 1	4, 1	6.2, 0.56
Re=29.7	31.4	43.9	88.8	60.7	89, 31	1.3, 41	1.4, 41	6.2, 40

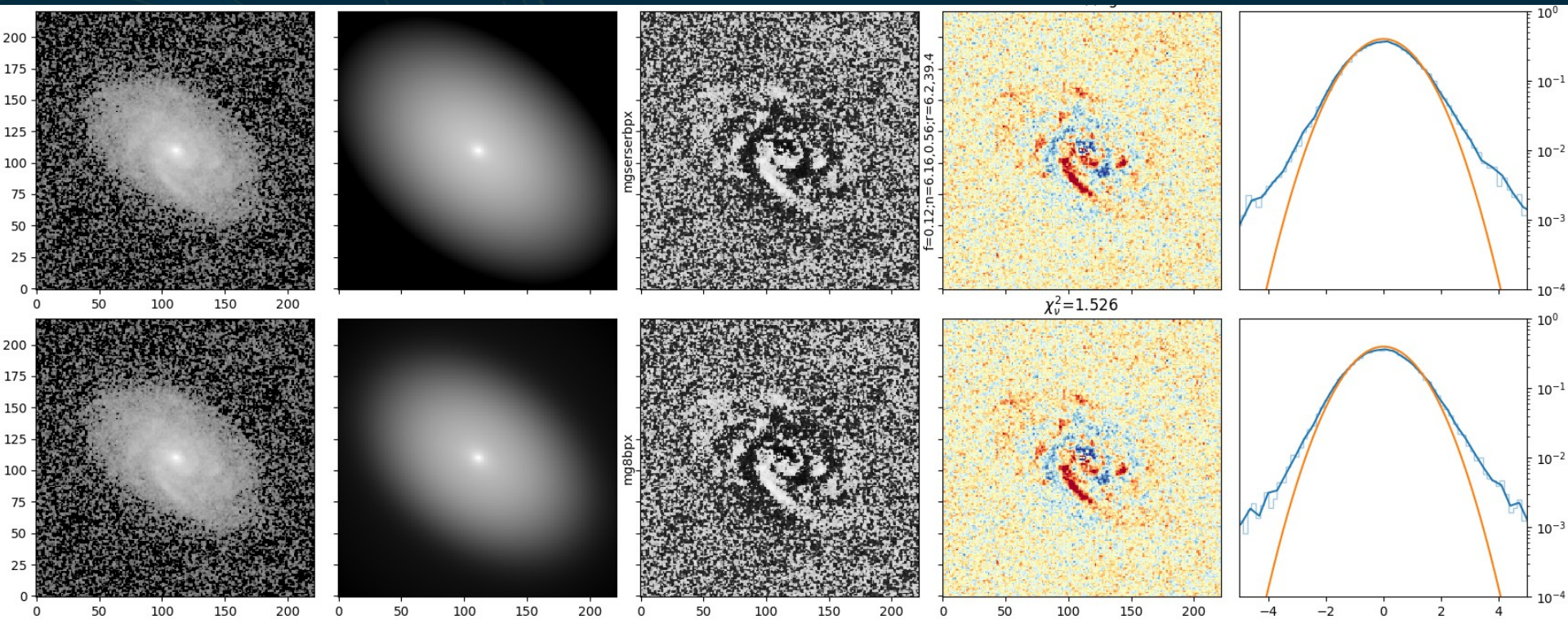


χ^2_ν=4.937 3.532 2.709 2.835 2.594 2.718 1.621 1.614 1.530

*NOT 8-Gaussian approx.

8-Gaussian Mixture Model

2-Sersic model χ^2_ν=1.530 (free ellipses)

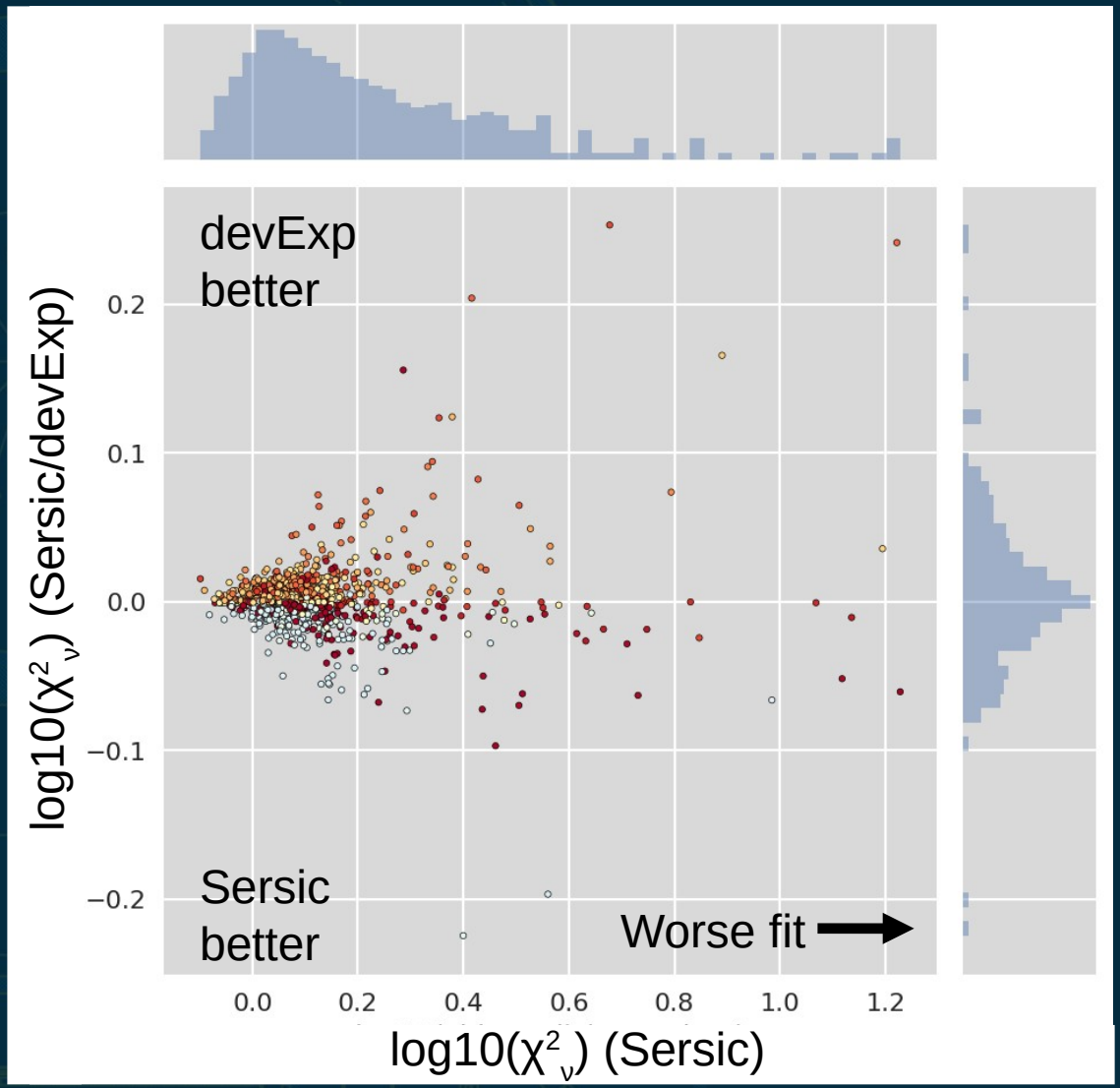


8-Gaussian model χ^2_ν=1.526 (shared ellipse)

one component, non-parametric radial profile

Metrics for Model Fits

- COSMOS-HST: Sersic better than devExp for ~50%



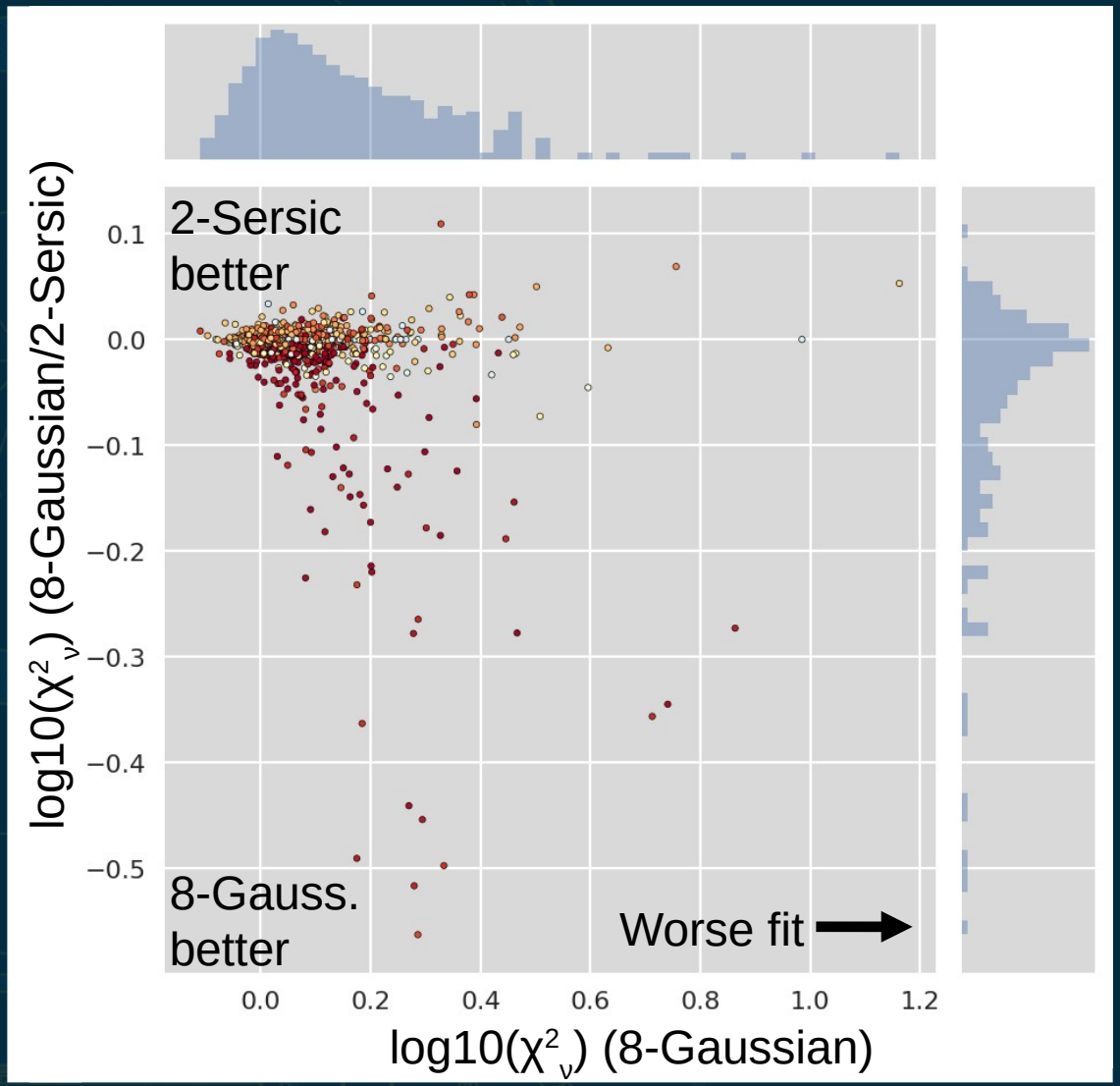
Sersic n=0.5

Sersic n=2

Sersic n=6

Metrics for Model Fits

- COSMOS-HST: 8-Gauss (same ellipses) vs 2-Sersic

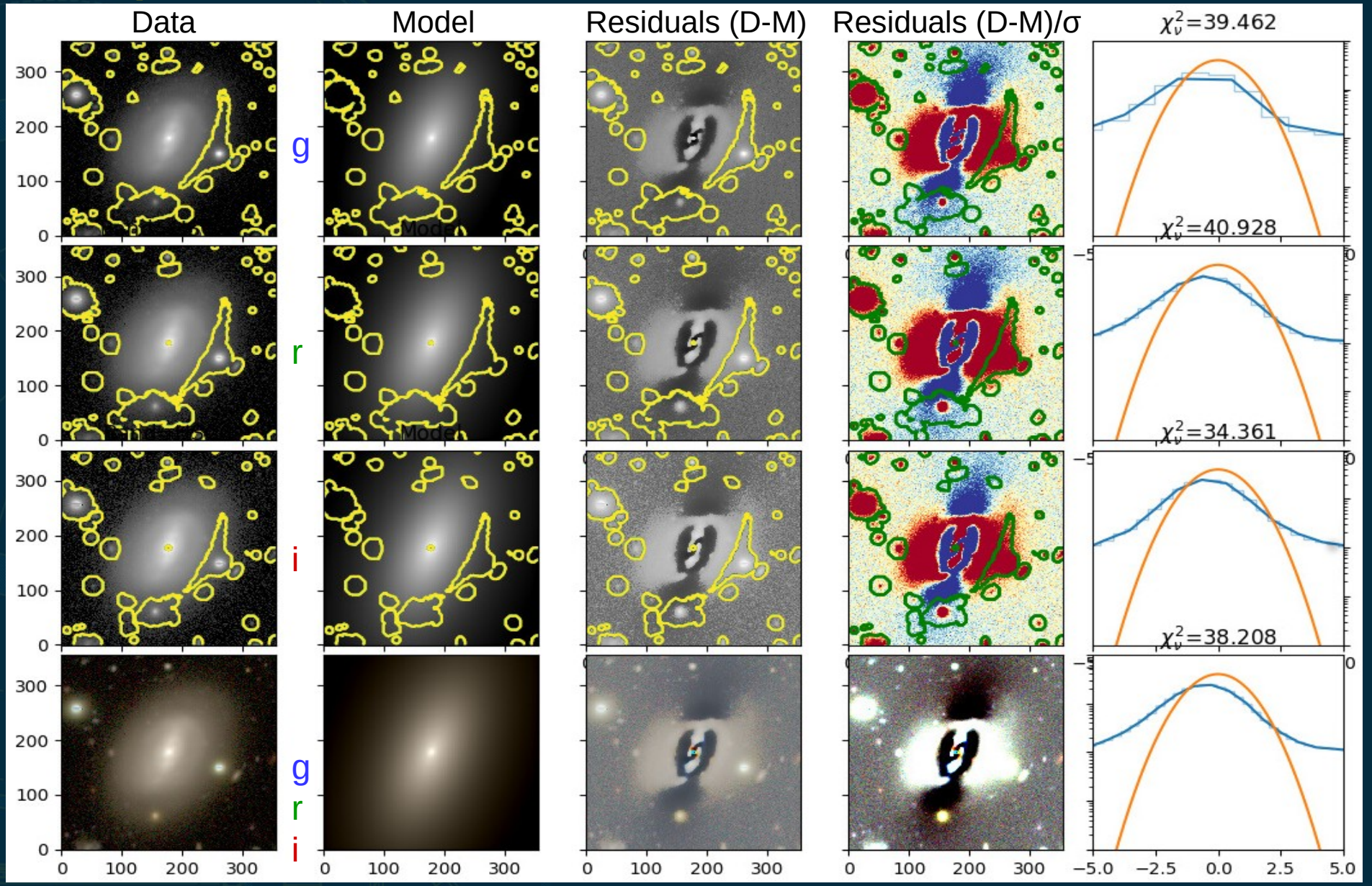


Sersic n=0.5

Sersic n=2

Sersic n=6

Experimental Multi-band Fits

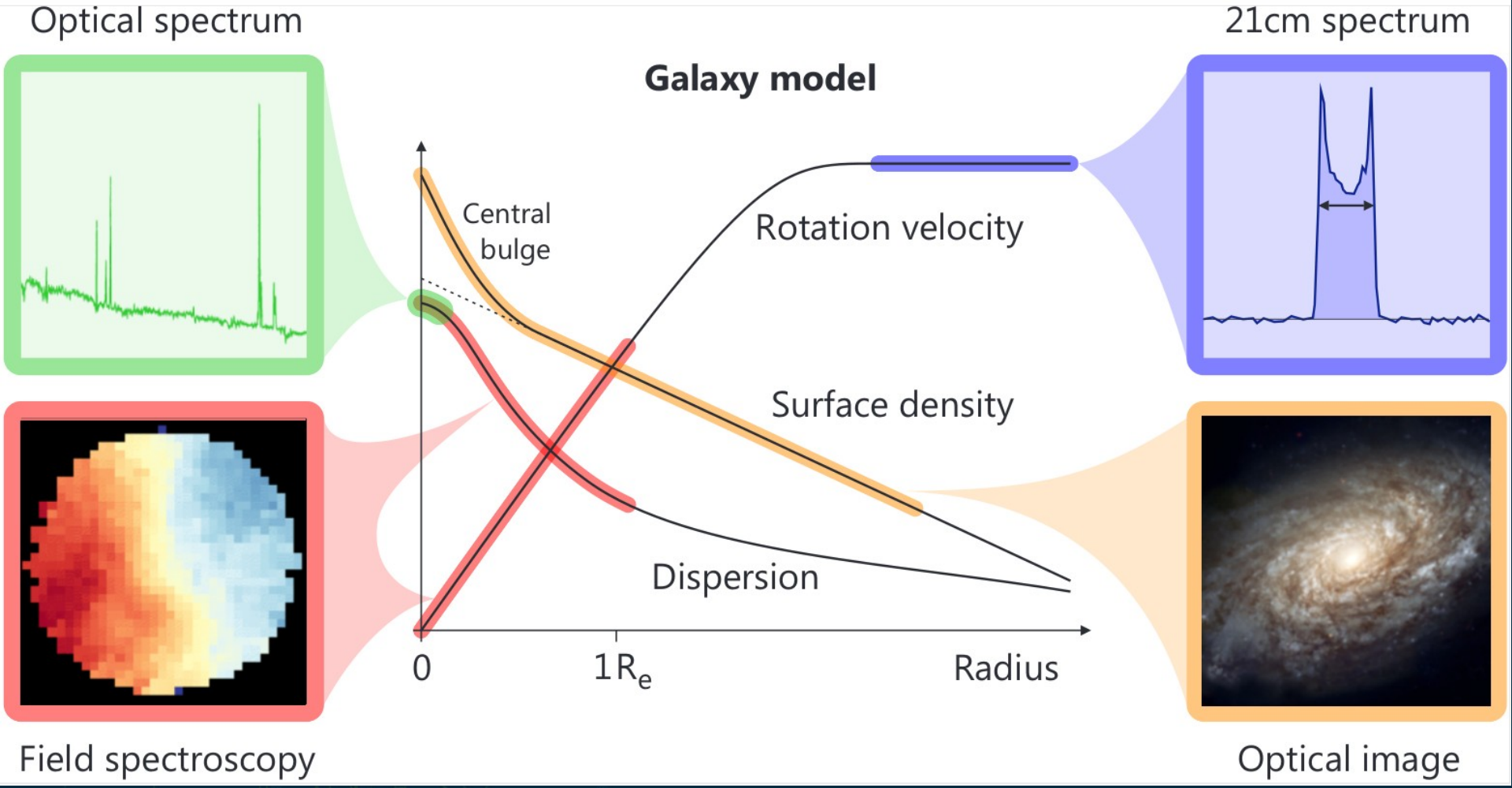


Conclusions

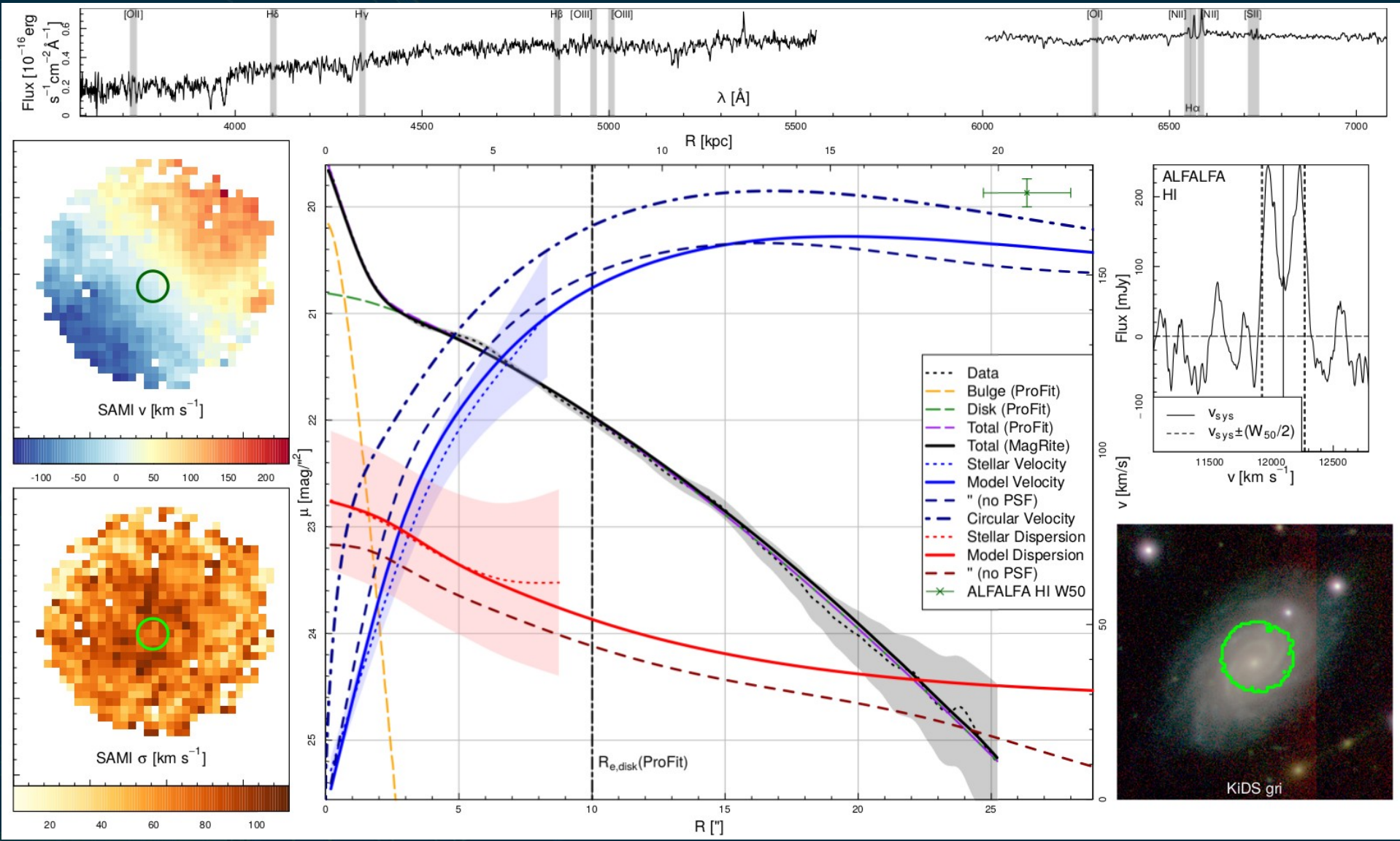
- LSST will detect billions of galaxies
  - The data will be big but manageable
  - The data is complex
- Data Management developing tools for turning pixels → (physical) parameters
  - New deblender Scarlet and MultiProFit (not TM)
  - Everything is open source!
    - github.com/lsst – Science Pipelines
- See lsst.org/participate and lsst.org/science
  - Community-driven requirements guide design

But what about...

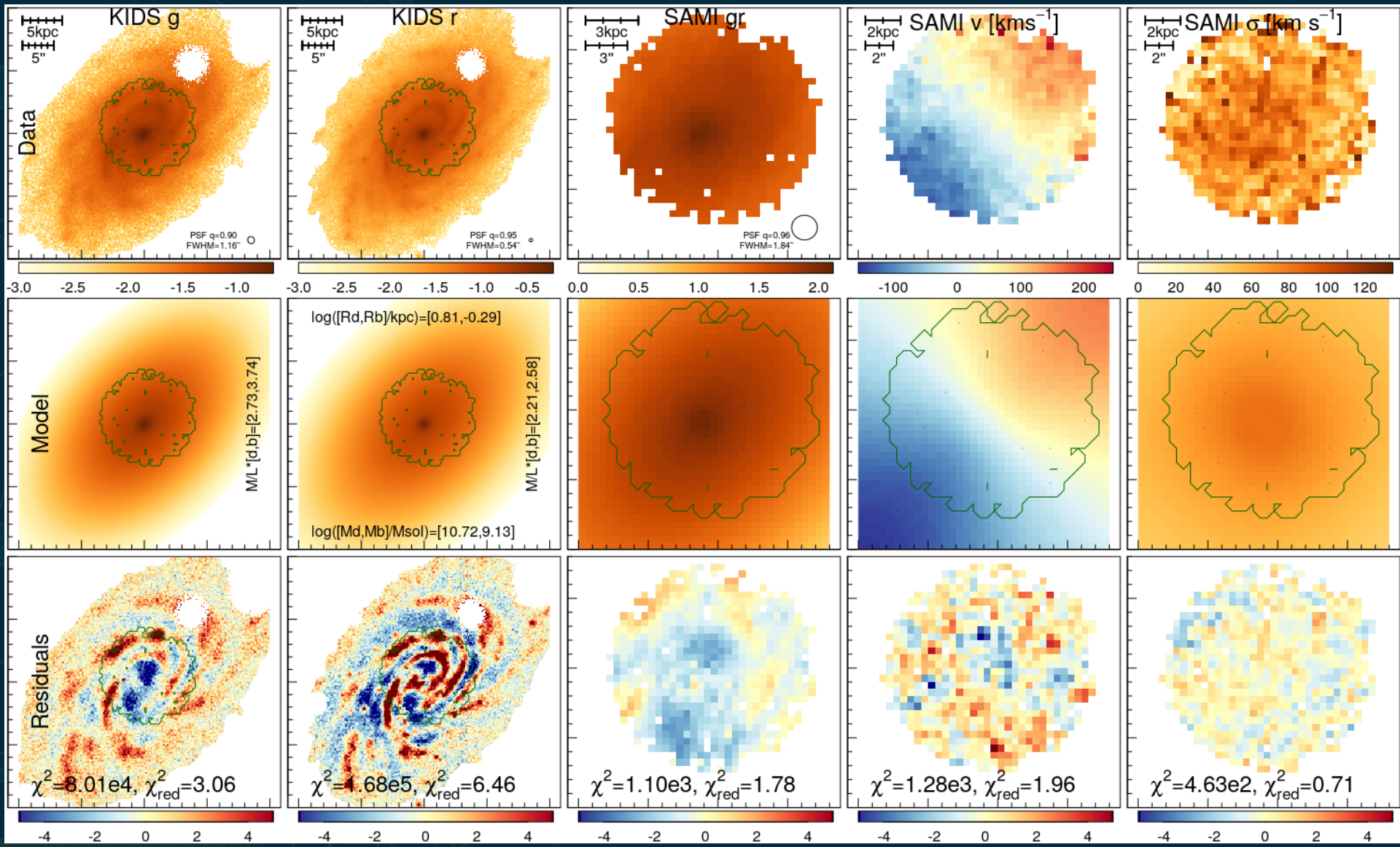
- I promised something about dynamical modelling:



MagRite (Taranu+17)

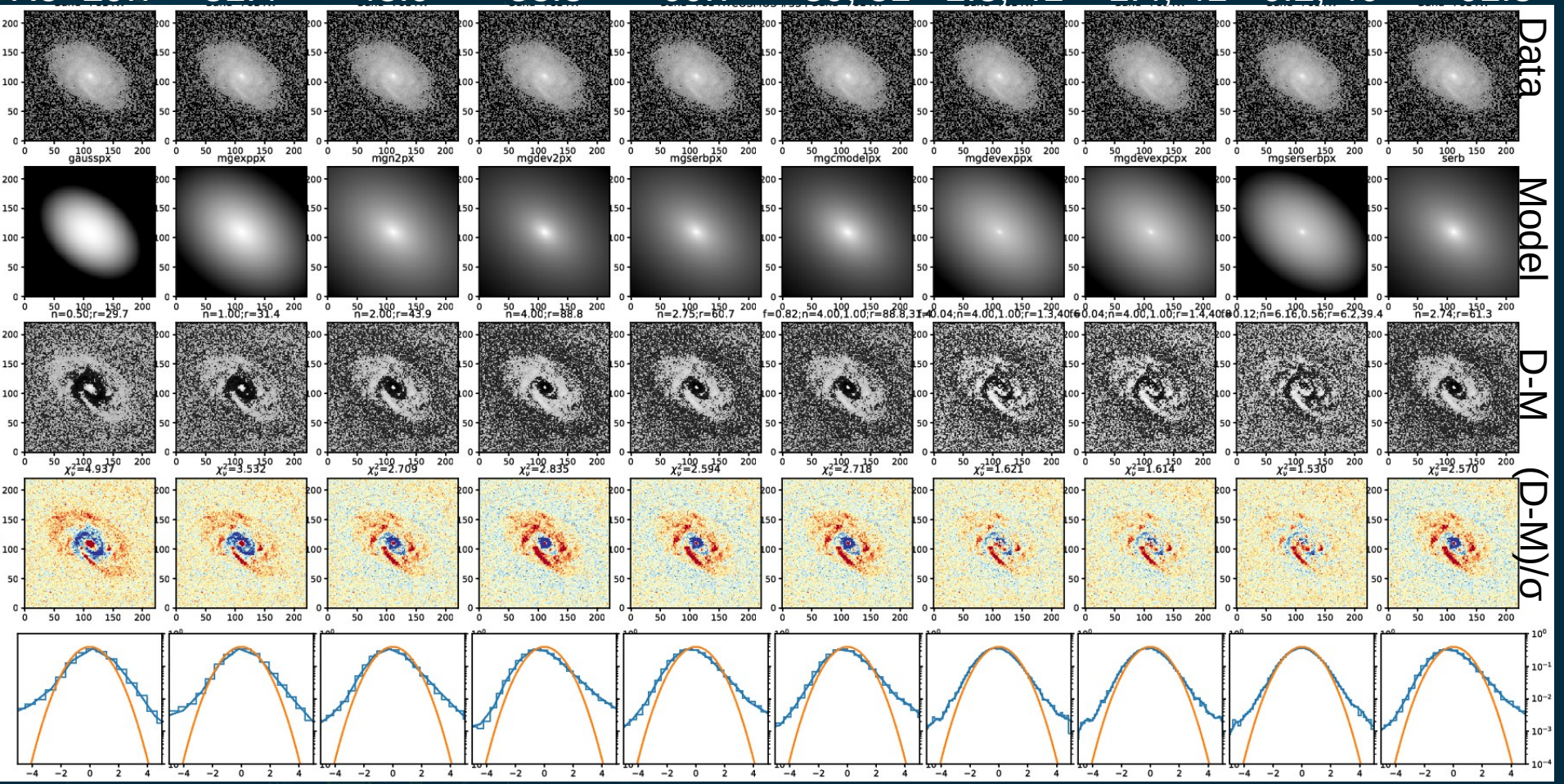


MagRite (Taranu+17)



Galaxy Models

Gauss.	Exp.	n=2	deV.	Sersic	cModel f=0.82	devExp f=0.04	devExpc f=0.04	2-Sersic f=0.12	Sersic*
n=0.5	1	2	4	2.75	4, 1	4, 1	4, 1	6.2, 0.56	2.74
Re=29.7	31.4	43.9	88.8	60.7	89, 31	1.3, 41	1.4, 41	6.2, 40	61.3

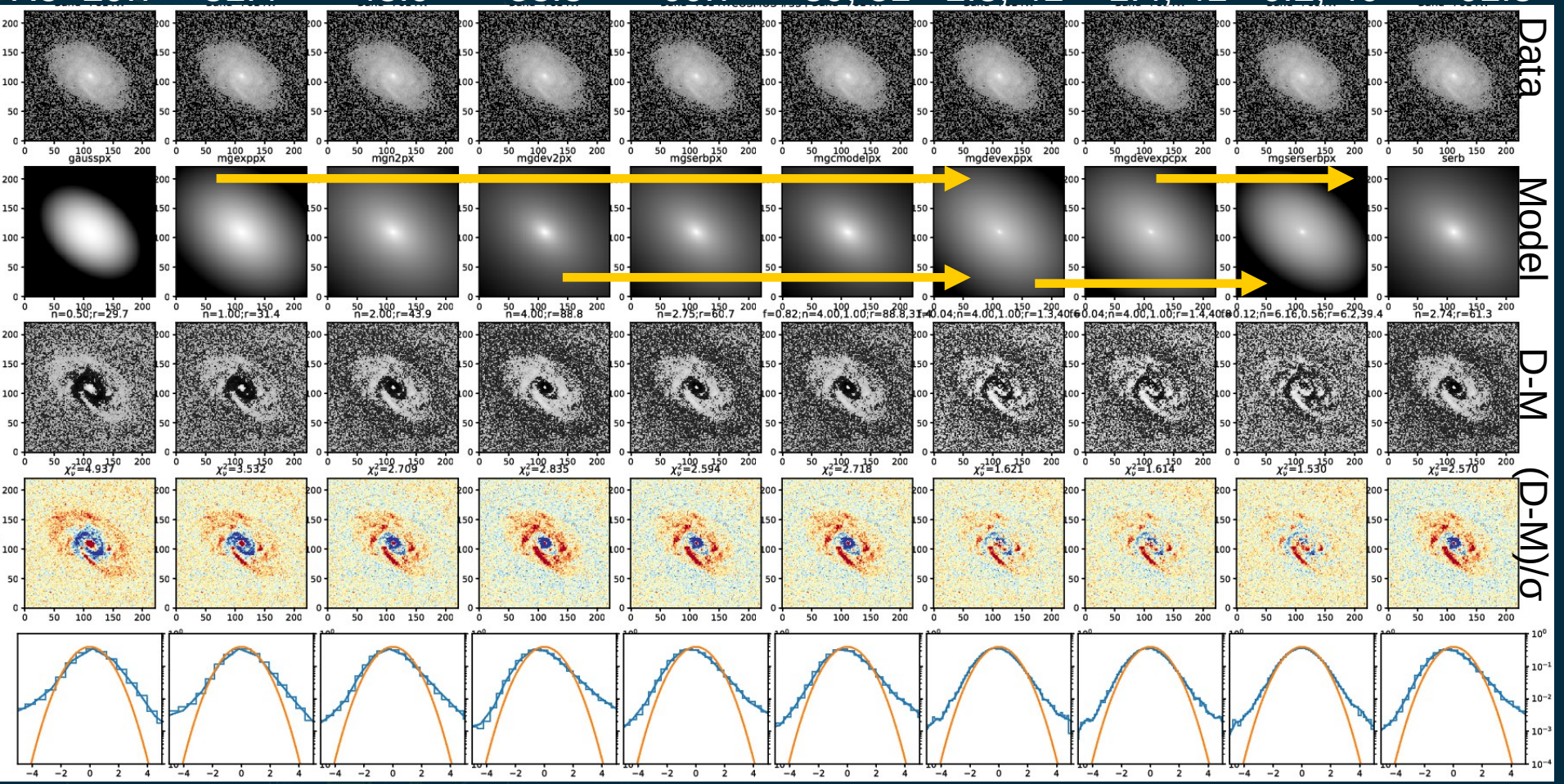


χ^2_ν=4.937 3.532 2.709 2.835 2.594 2.718 1.621 1.614 1.530 2.570

*NOT 8-Gaussian approx.

Galaxy Models

Gauss.	Exp.	n=2	deV.	Sersic	cModel f=0.82	devExp f=0.04	devExpc f=0.04	2-Sersic f=0.12	Sersic*
n=0.5	1	2	4	2.75	4, 1	4, 1	4, 1	6.2, 0.56	2.74
Re=29.7	31.4	43.9	88.8	60.7	89, 31	1.3, 41	1.4, 41	6.2, 40	61.3



χ^2_ν=4.937 3.532 2.709 2.835 2.594 2.718 1.621 1.614 1.530 2.570

*NOT 8-Gaussian approx.