Title: Canadian Astronomy Data Center: Tools and Analytics for Large Data Sets

Date: May 07, 2018 04:30 PM

URL: http://pirsa.org/18050043

Abstract:

Deep into Data and Distances

•••

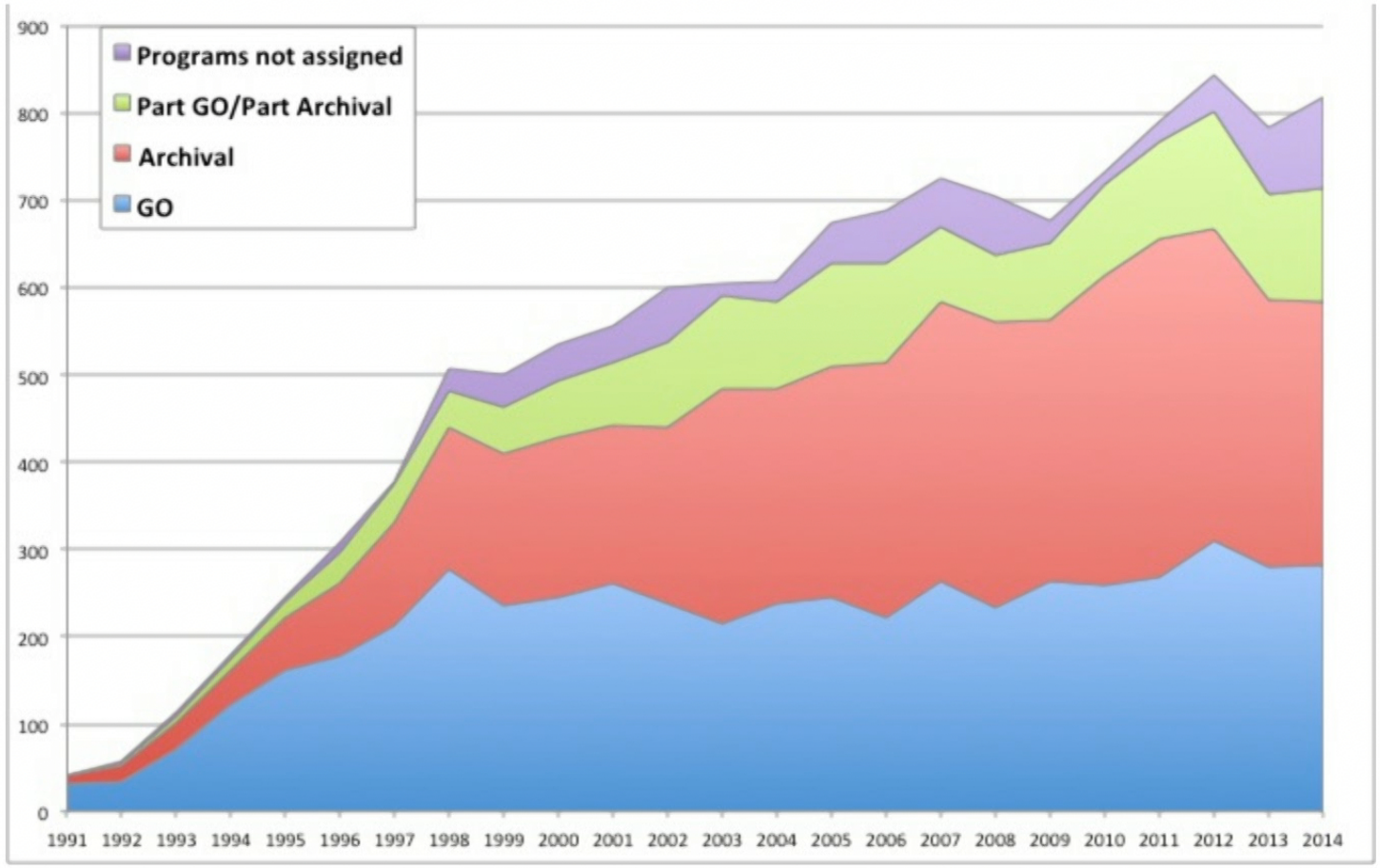
PI-NRC Meeting 7 May 2018

Sébastien Fabbro - NRC Herzberg



Data archiving and Discovery

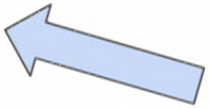
Hubble Space Telescope Publication Rate

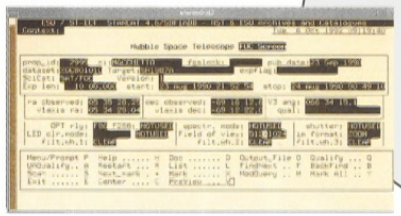


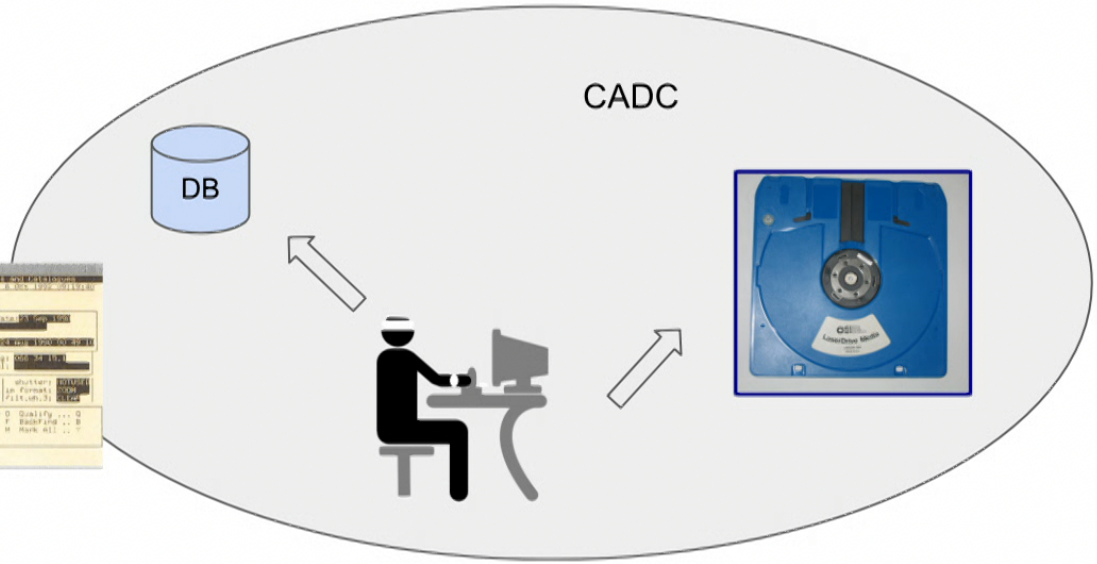
from archives

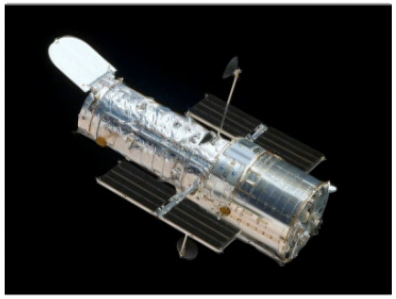


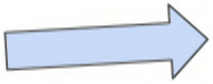
1986



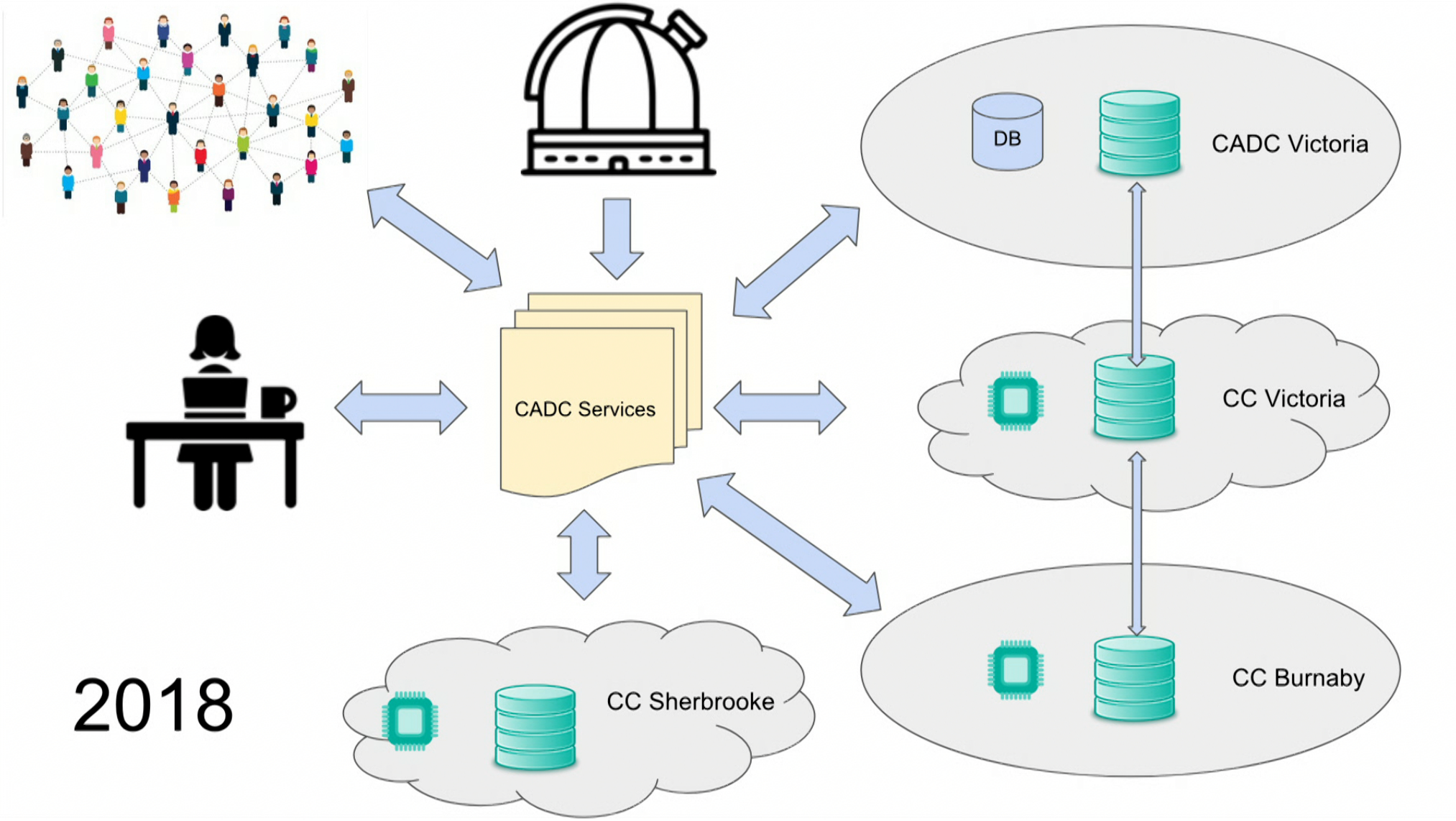


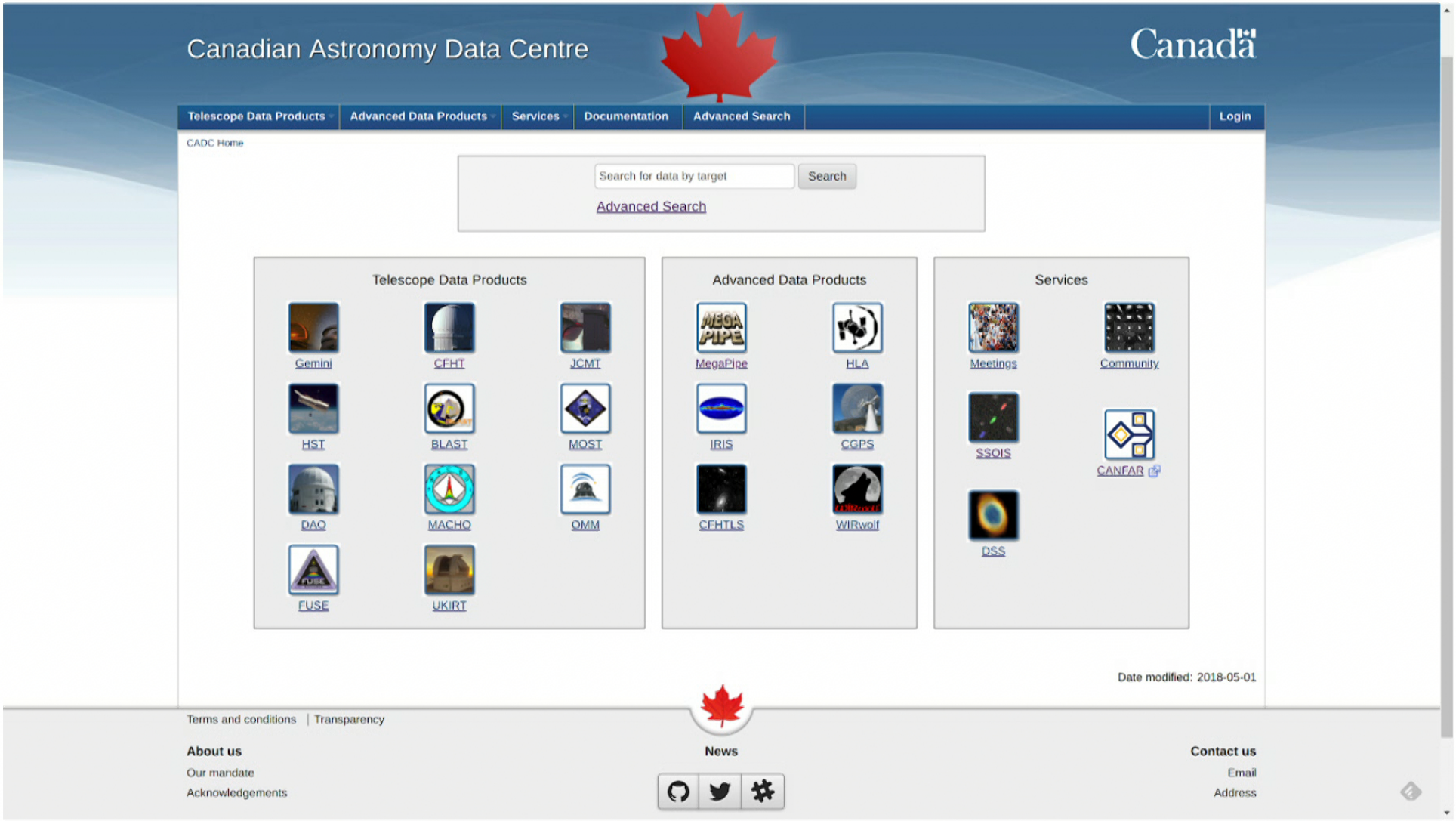




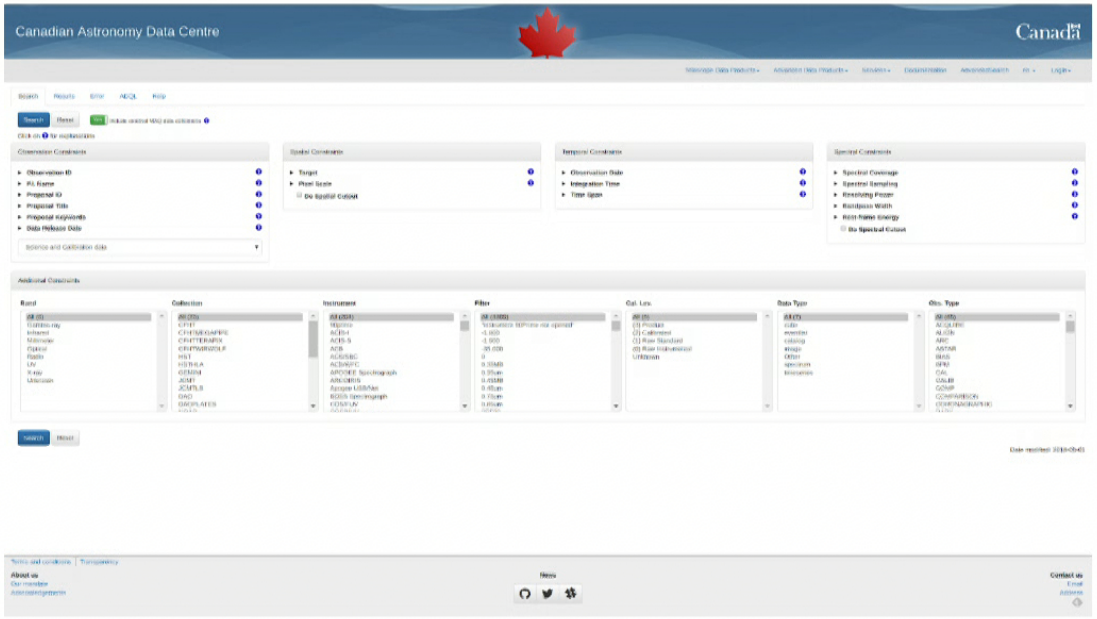





DB
CADC Victoria
CADC Services
CC Victoria
CC Sherbrooke
CC Burnaby
2018


Canadian Astronomy Data Centre
Canada
Telescope Data Products
Advanced Data Products
Services
Documentation
Advanced Search
Login
CADC Home
Search for data by target
Search
Advanced Search
Telescope Data Products
Gemini
CFHT
JCMT
HST
BLAST
MOST
DAO
MACHO
OMM
FUSE
UKIRT
Advanced Data Products
MegaPipe
HLA
IRIS
CGPS
CFHTLS
WIRwolf
Services
Meetings
Community
SSOIS
CANFAR
DSS
Date modified: 2018-05-01
Terms and conditions | Transparency
About us
Our mandate
Acknowledgements
News
Contact us
Email
Address

Astronomy Data Discovery at CADC



Search observatory-based data by:

- time
- space
- wavelength / energy

Latest news: April 2018

- search external archives!
- increase 115 to 168 collections
- 36M data sets now discoverable

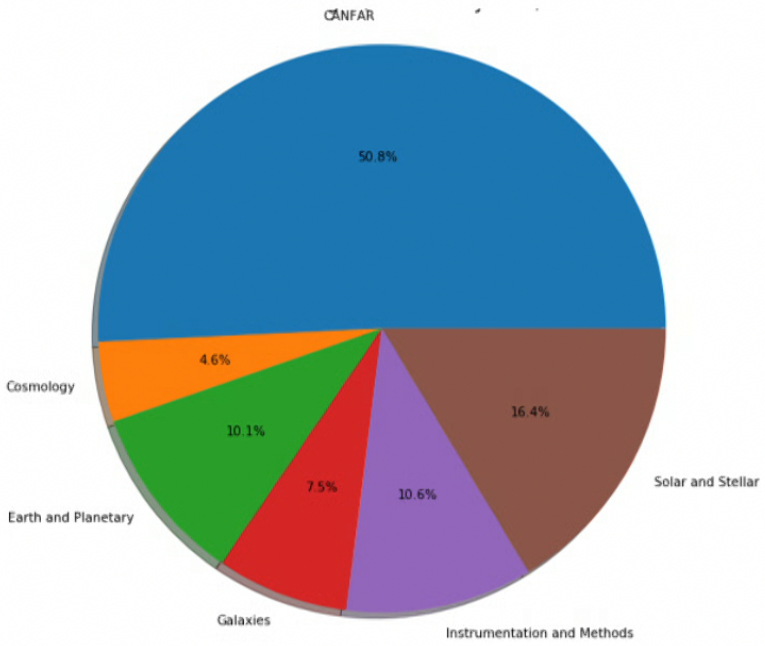
Canadian Advanced Network for Astronomical Research

Ecosystem for data intensive astronomy

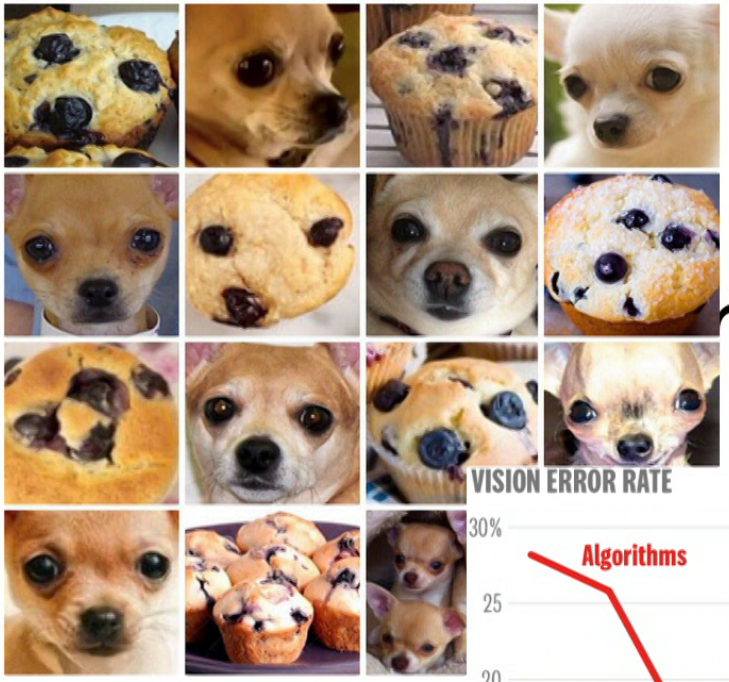
NRC + Universities + Compute Canada (+ CANARIE)

In 2017:

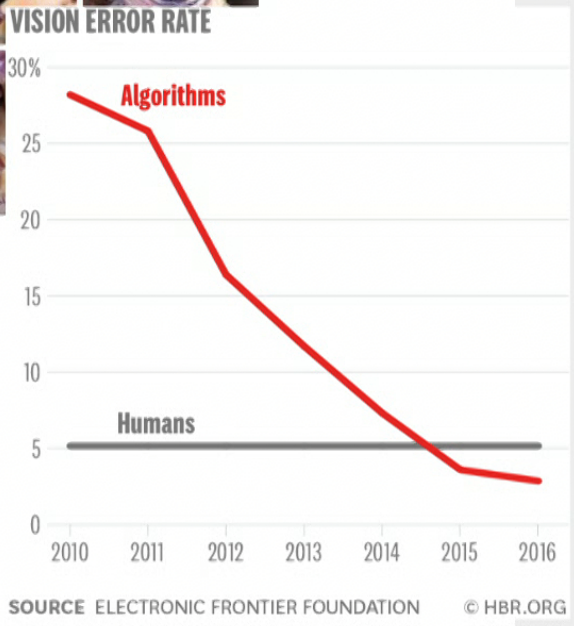
- 212 collaborations on 200TB data
- 43 interactive analysis projects
- 13M batch processing jobs





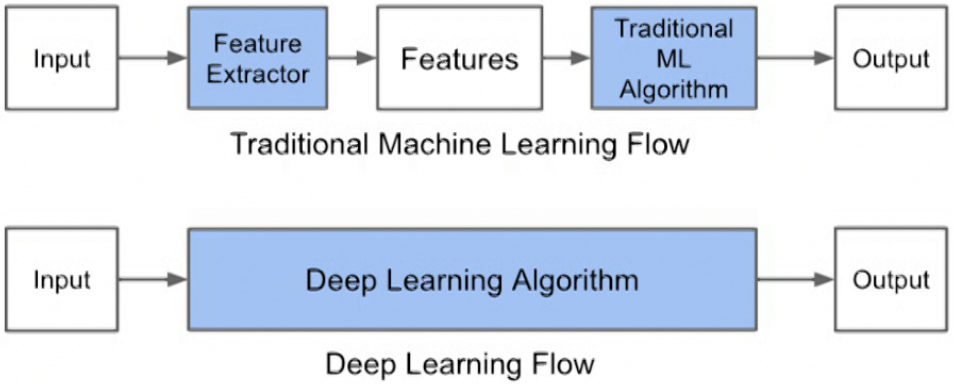


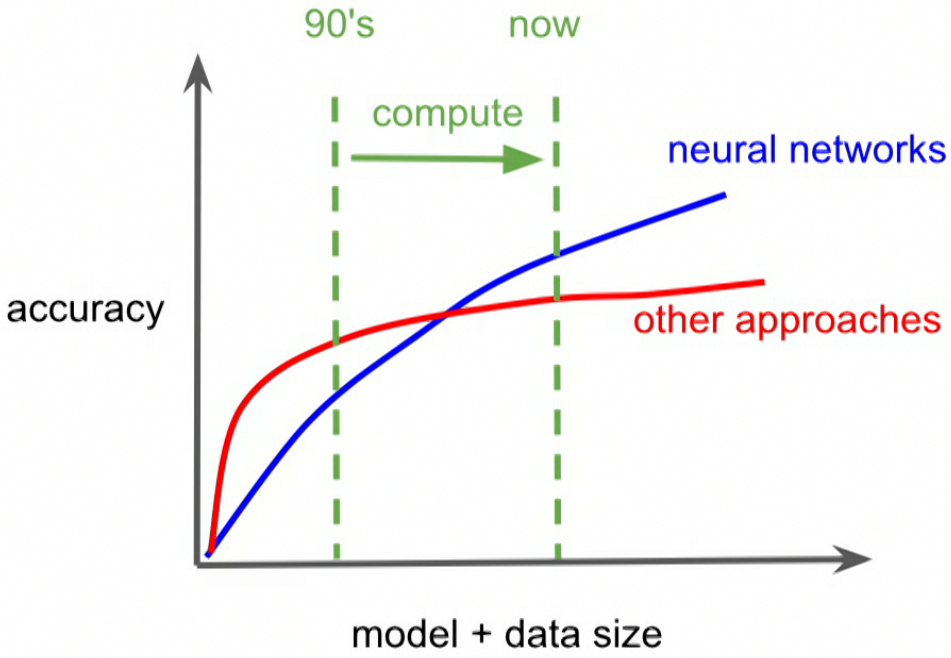
muffin or puppy?



Some Astrophysical Applications of Deep Learning

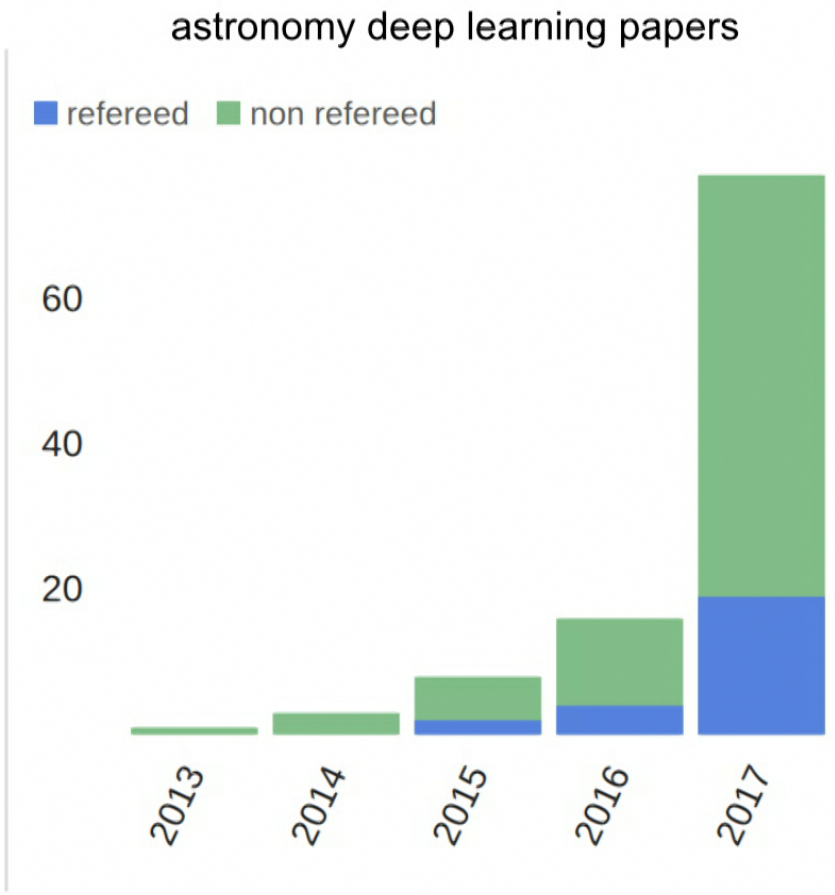
Deep Learning





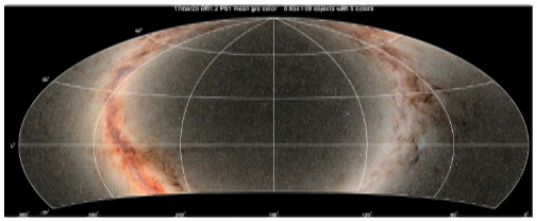
Deep Learning: Do Astronomers Care?

- detection in large data sets
- classify astronomical objects
- emulate time consuming simulations
- automate human error-prone analysis
- powerful well written software
- attract students



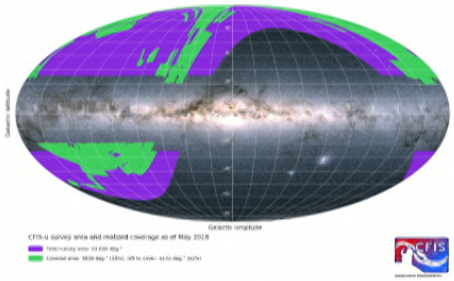
Classification with Canada-France Imaging Survey

PanSTARRS

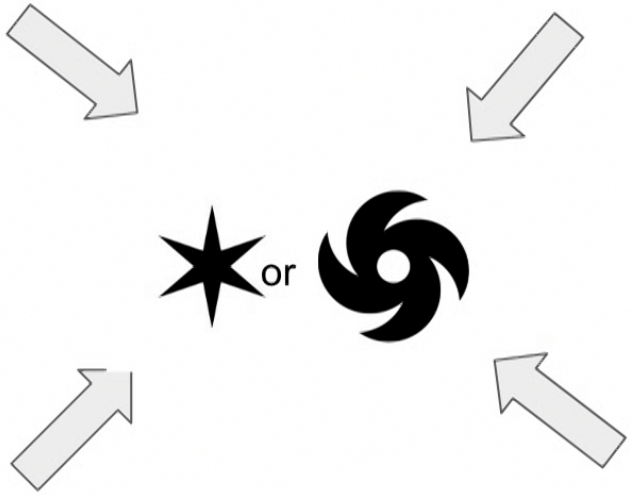


2B objects

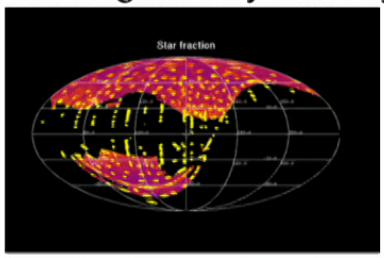
Canada France Imaging Survey



200M objects

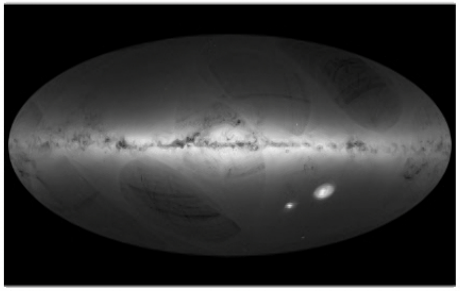


Sloan Digital Sky Survey



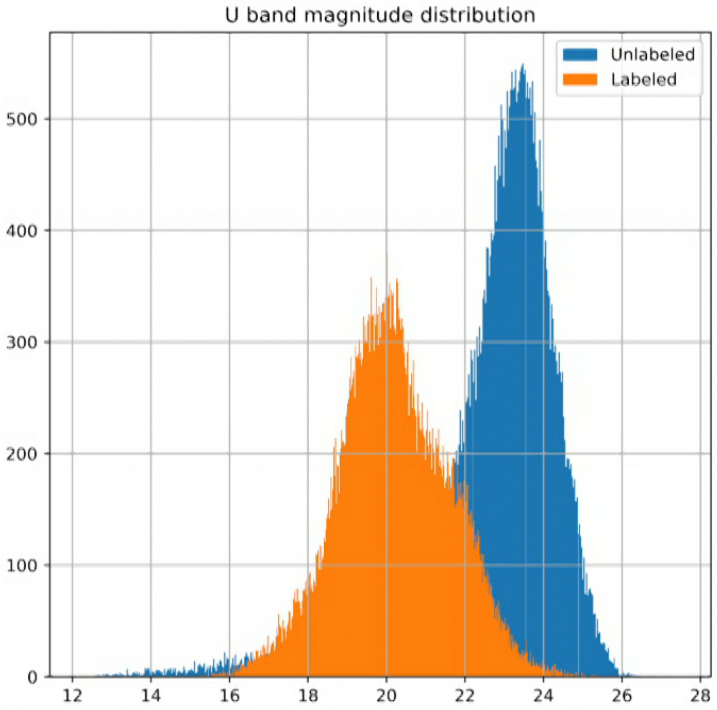
500M objects

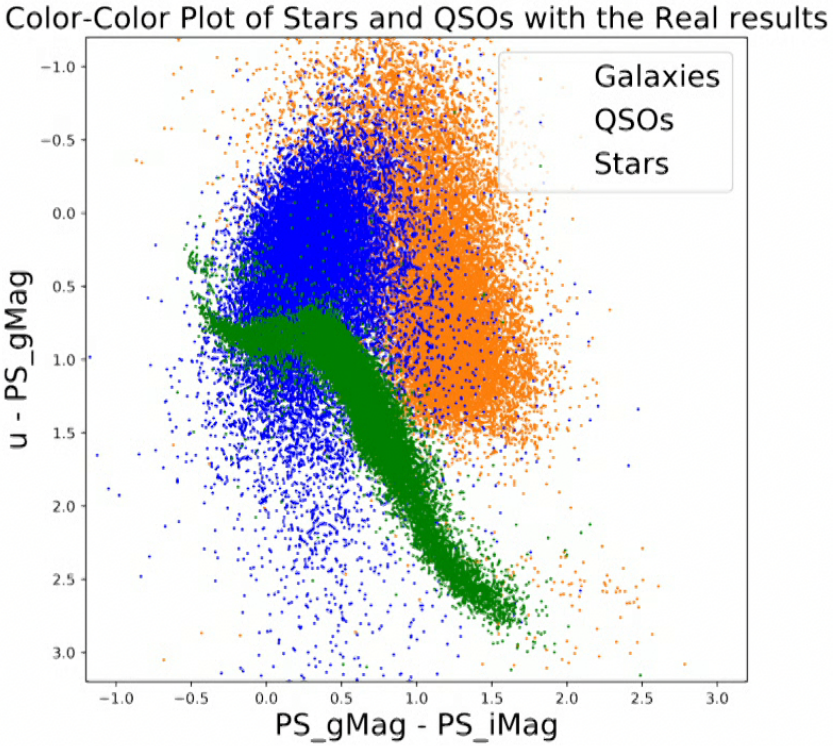
GAIA



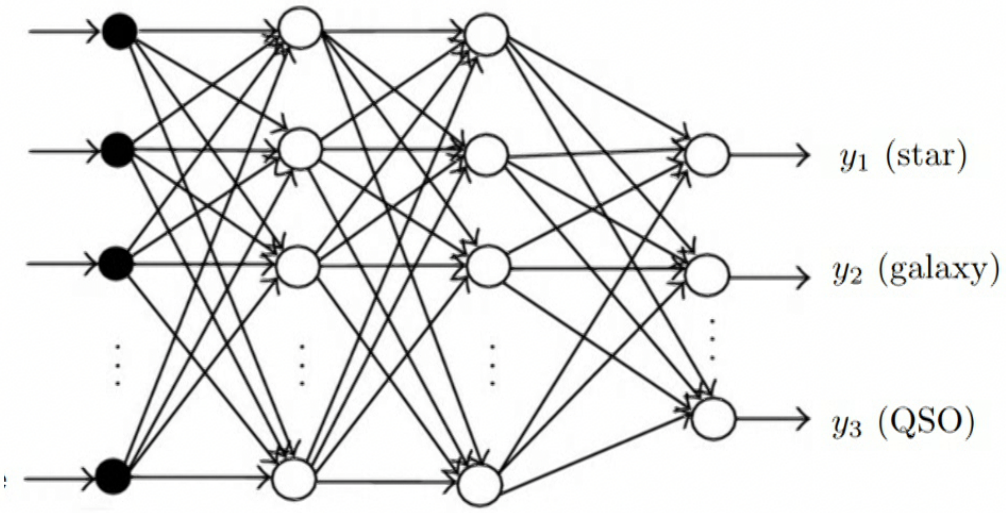
1.7B objects

Classification on Canada-France Imaging Survey

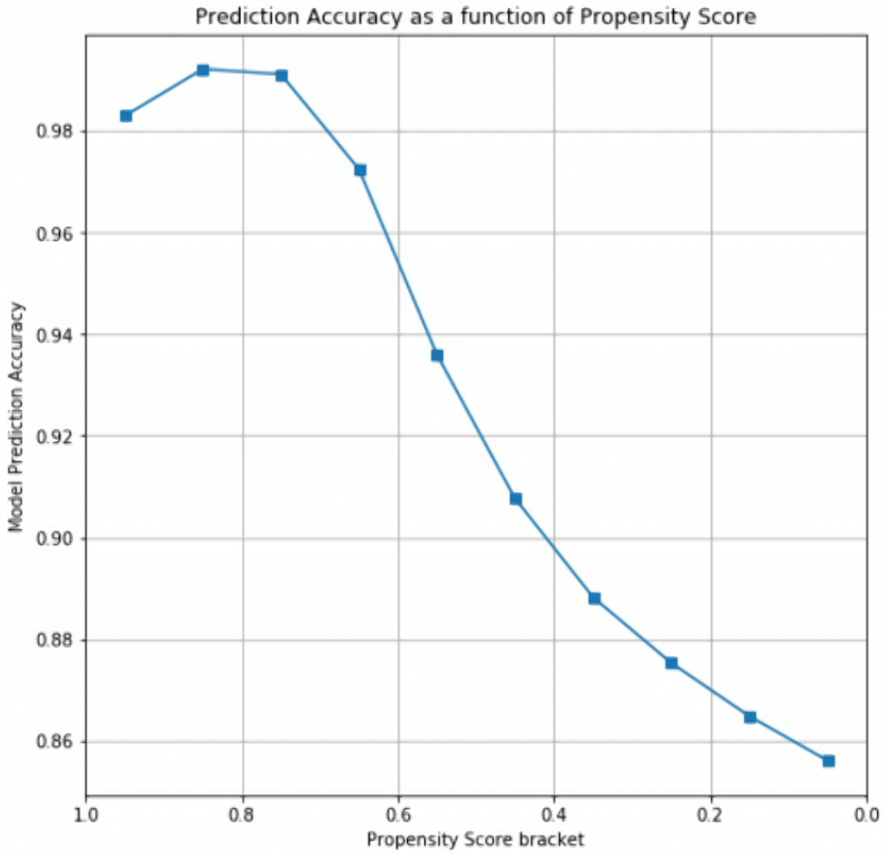




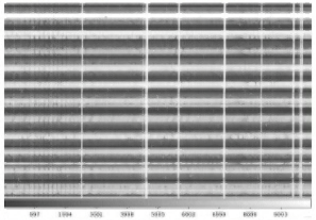
Deep Ensemble Networks



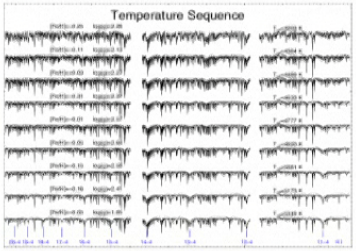
X5 + Monte Carlo input on Features
Propensity scores to match distributions



raw 2D image



calibrated 1D spectra



estimated stellar parameters

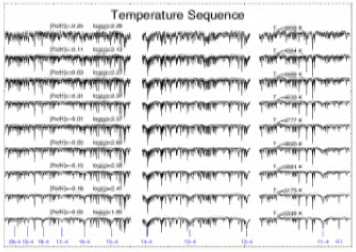
T_{eff}, log(g), [Fe/H],
[C/H], [N/H], [O/H],
[Na/H], [Mg/H], [Al/H],
[Si/H], [S/H], [K/H],
[Ca/H], [Ti/H], [V/H],
[Mn/H], [Ni/H],...

stellar populations
galactic archeology
near-field cosmology

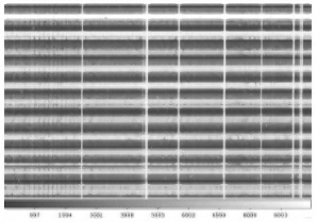
stellar parameters

T_{eff}, log(g), [Fe/H],
[C/H], [N/H], [O/H],
[Na/H], [Mg/H], [Al/H],
[Si/H], [S/H], [K/H],
[Ca/H], [Ti/H], [V/H],
[Mn/H], [Ni/H],...

synthetic 1D spectra

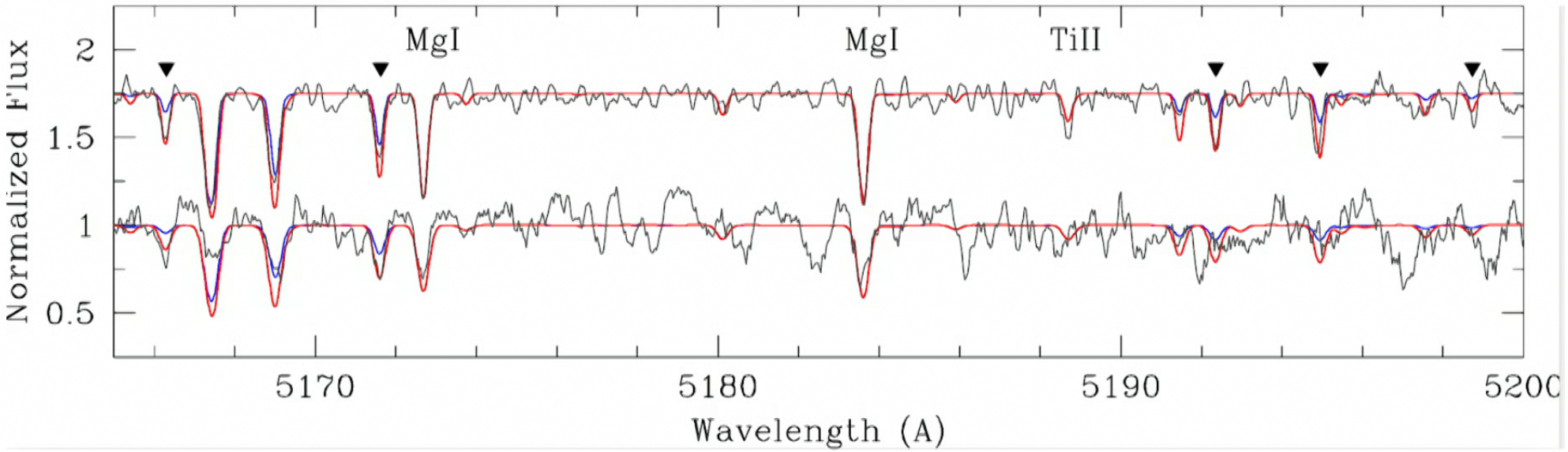


synthetic 2D image



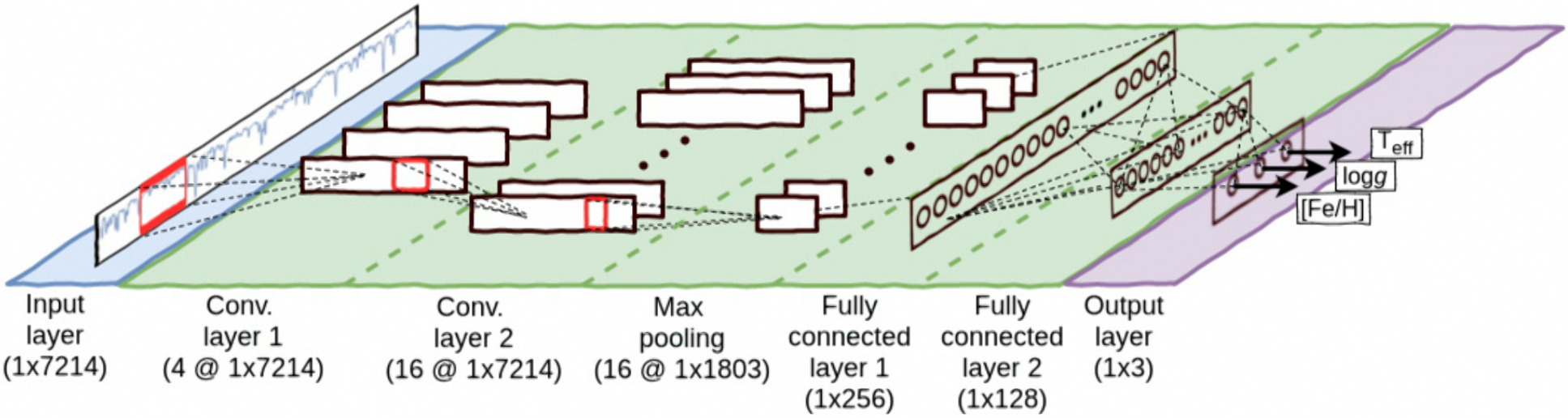
Traditional Approach

Gemini-GRACES spectroscopy, Venn et. al (2017)



- physics driven methods (EQW, NL fit, projections…): very hard to automate
- data driven methods (The Cannon): need a set of reference stars

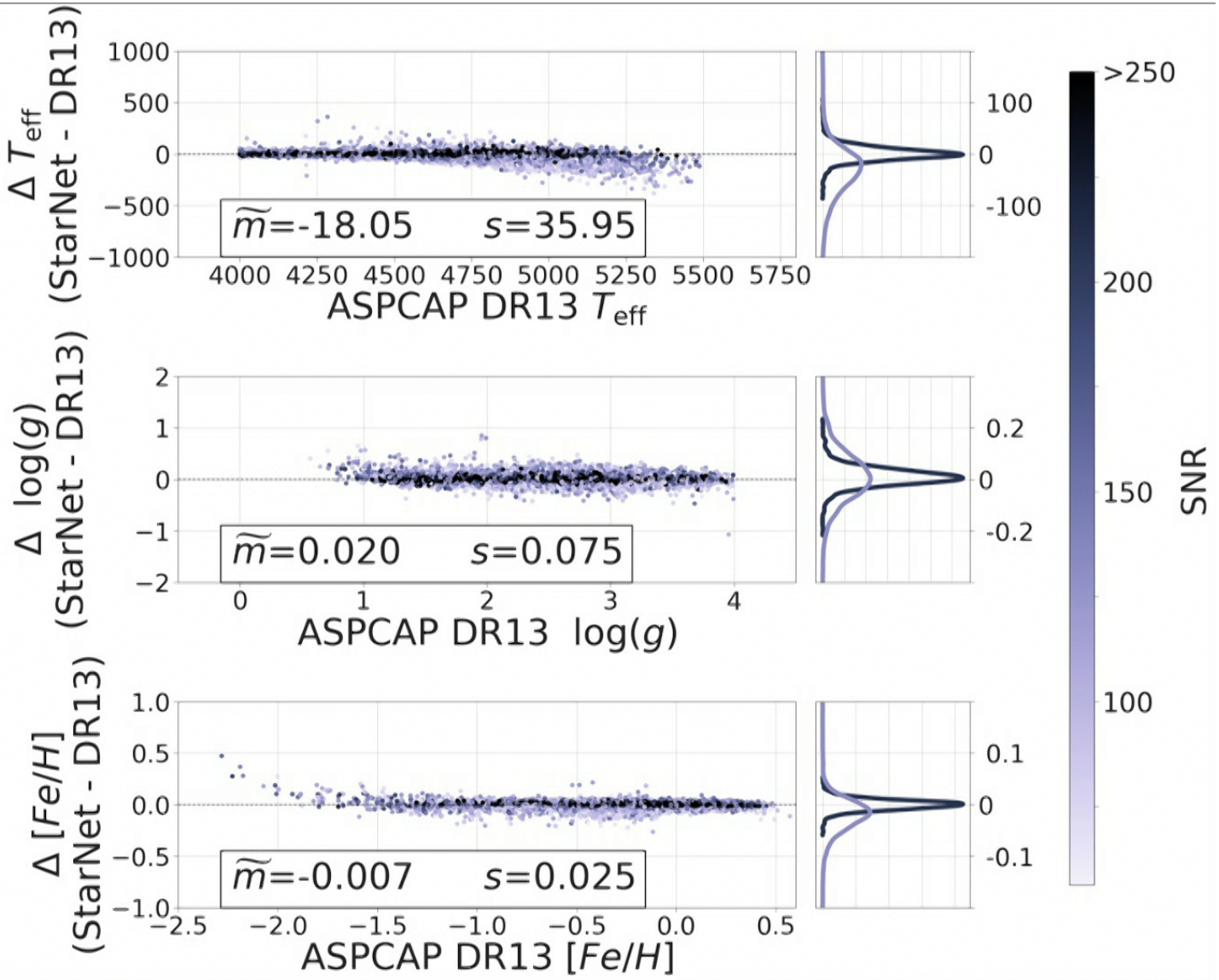
StarNet 1D Convolutional Neural Network



Fabbro et
StarNet
(2018)

train with 41,000 APOGEE spectra predict on the rest of APOGEE

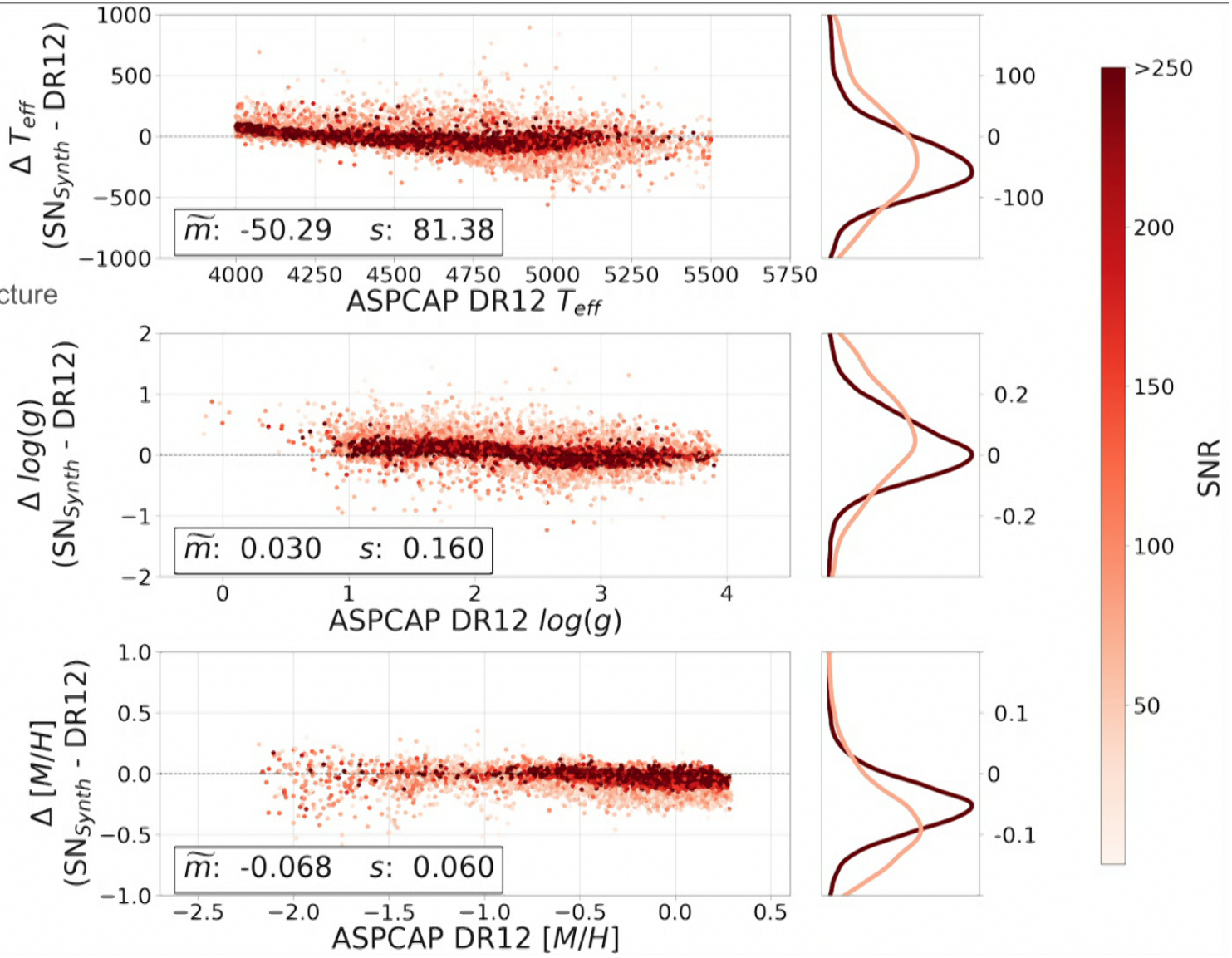
15k stars with 47k visits



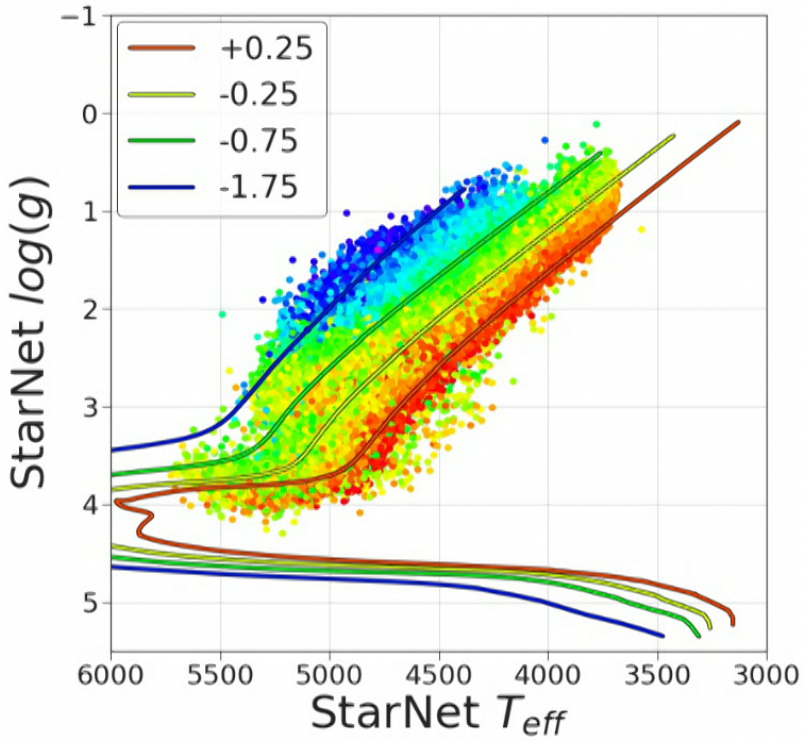
train on synth predict on APOGEE

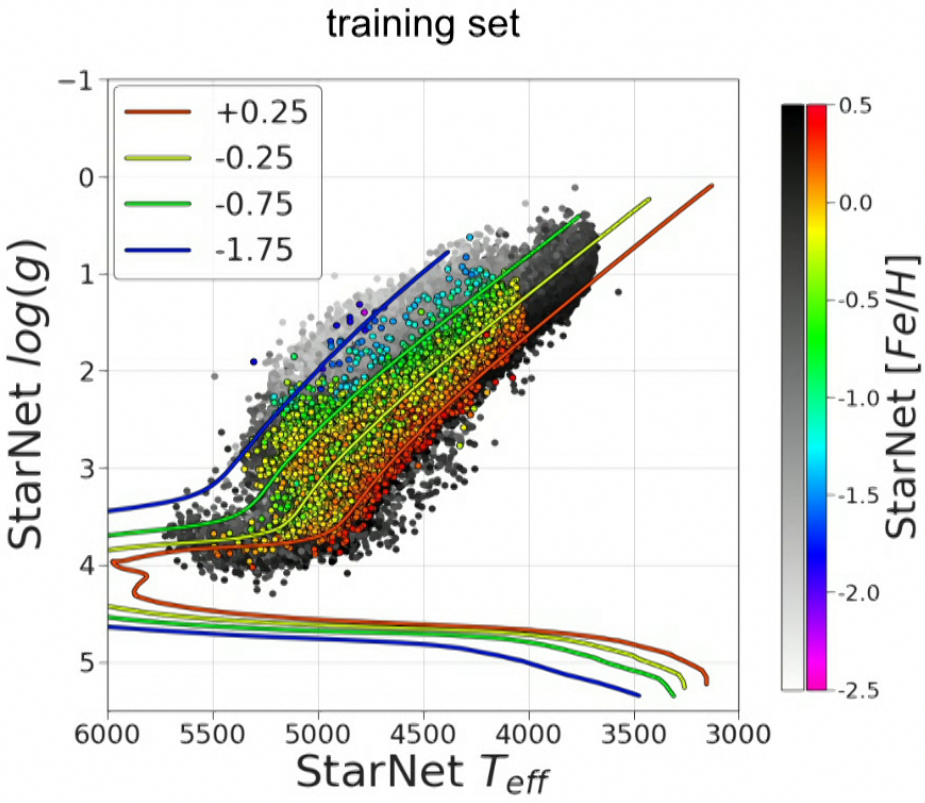
no modification to StarNet architecture

same 224k stars as simulations

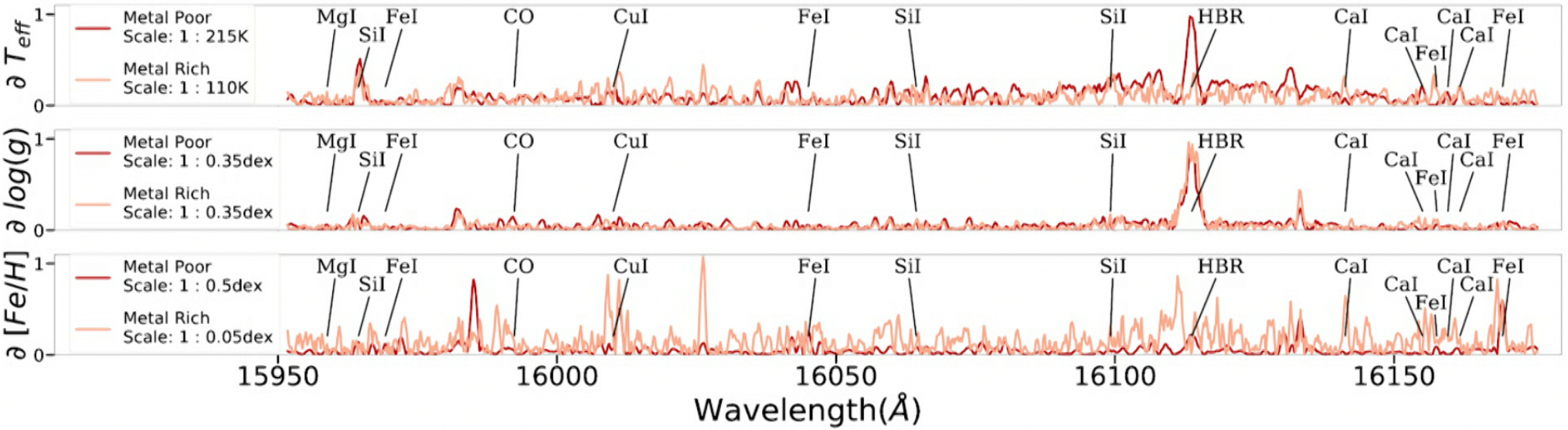


StarNet Fidelity Check

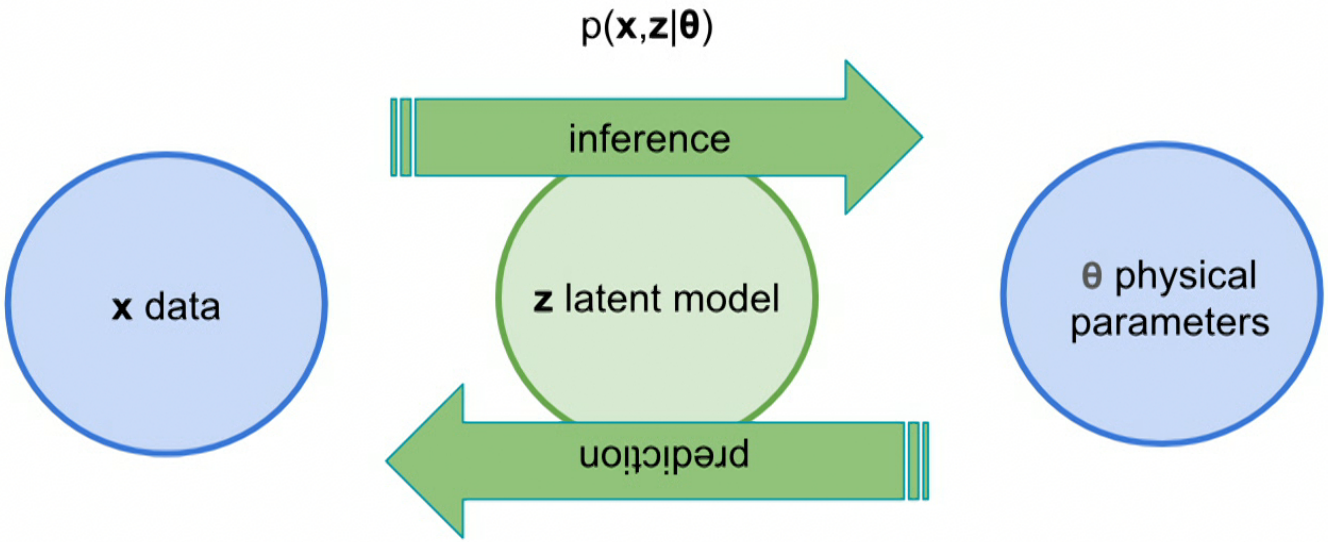




Stellar Spectrum Information Synthetic Modeling



Machine Learning + Statistical Modeling

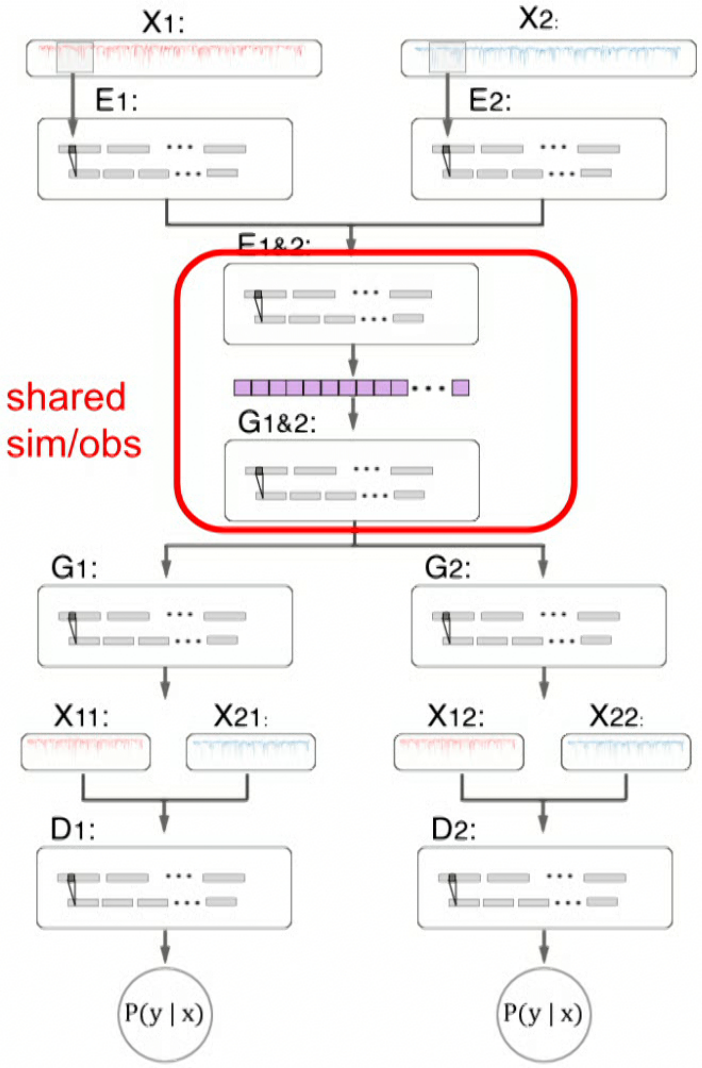


generative implicit models, deep probabilistic programming

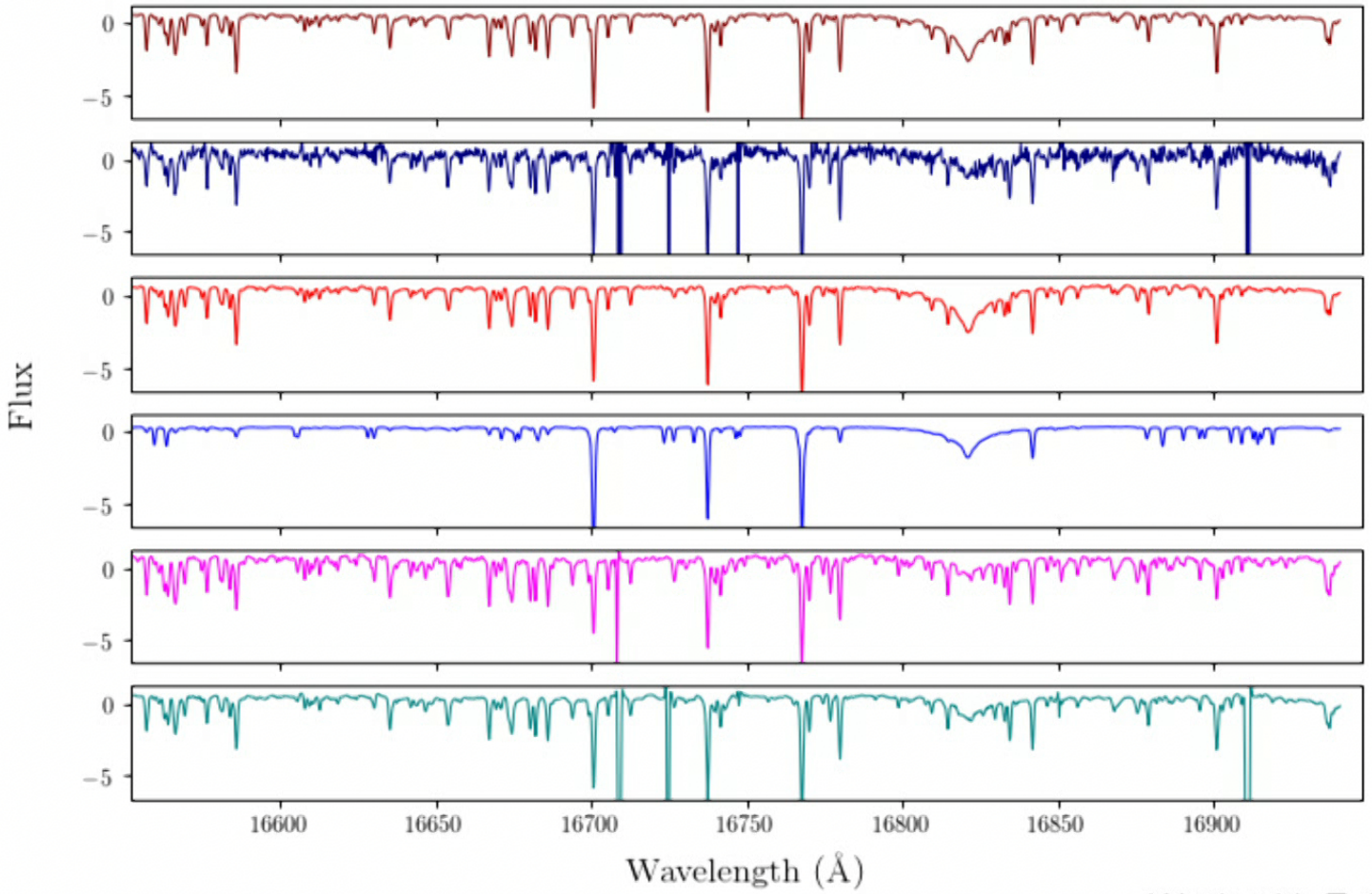
Unsupervised Spectra to Spectra



Liu, Breuel, Kauz (2017)



Mixing Models and APOGEE data with GAN!



X1 simulated

X2 observed

X11 ML simulated

X21 ML observed+simulated

X12 ML simulated+observed

X22 ML observed

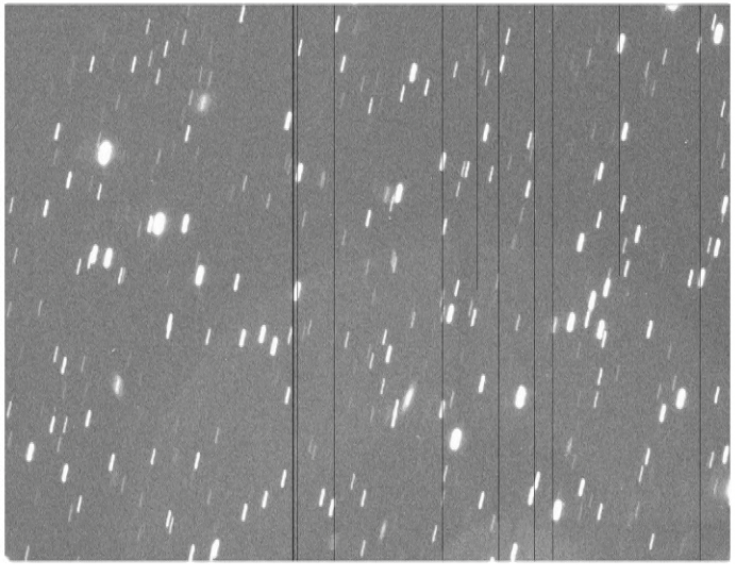
Work with T.O'Briain - StarNet and K.M. Yi

CADC Image Quality Assessment with CNN

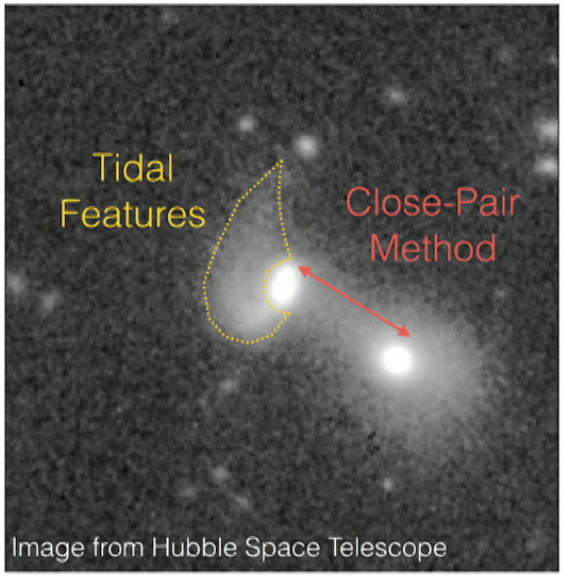
A few millions exposures to assess
Manual label of 130k images
Training on CNN

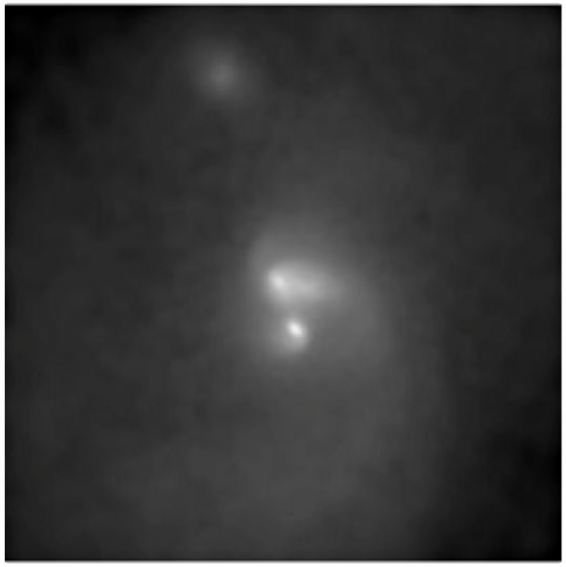
So far: 97% accuracy…

Work in progress with H.Teimorinia and K.Rolston



Finding Galaxy Mergers in CFIS with ML





H. Teimoorinia with
CFIS/UVic

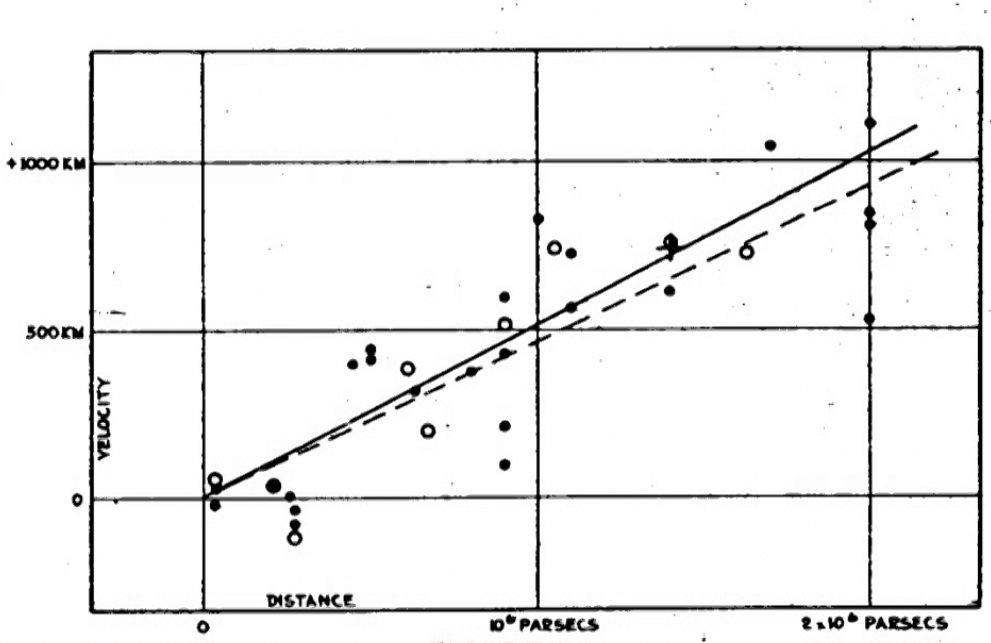
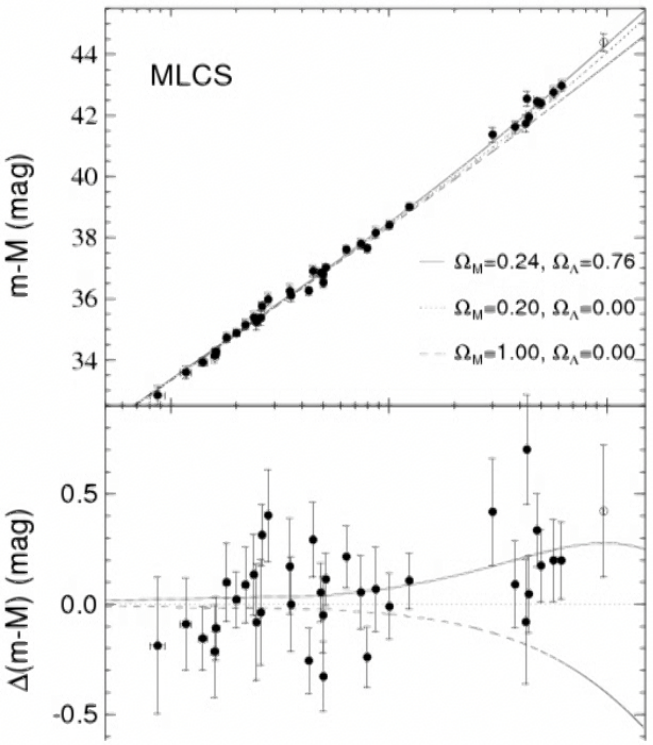


FIGURE 1

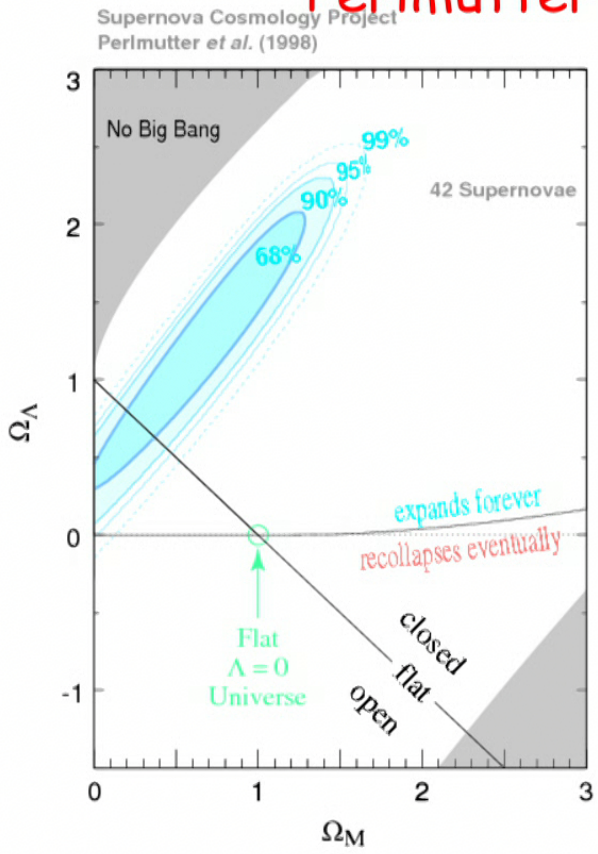
Processing for Supernova Distances

1998: Accelerating Universe



Riess et High-z Search Team

Perlmutter et SCP



1998

Distant supernovae Ia dimmer than expected
Universe expansion accelerates

Now

Acceleration confirmed with other observations
What is the cause of the accelerated expansion?
(a $10B question)

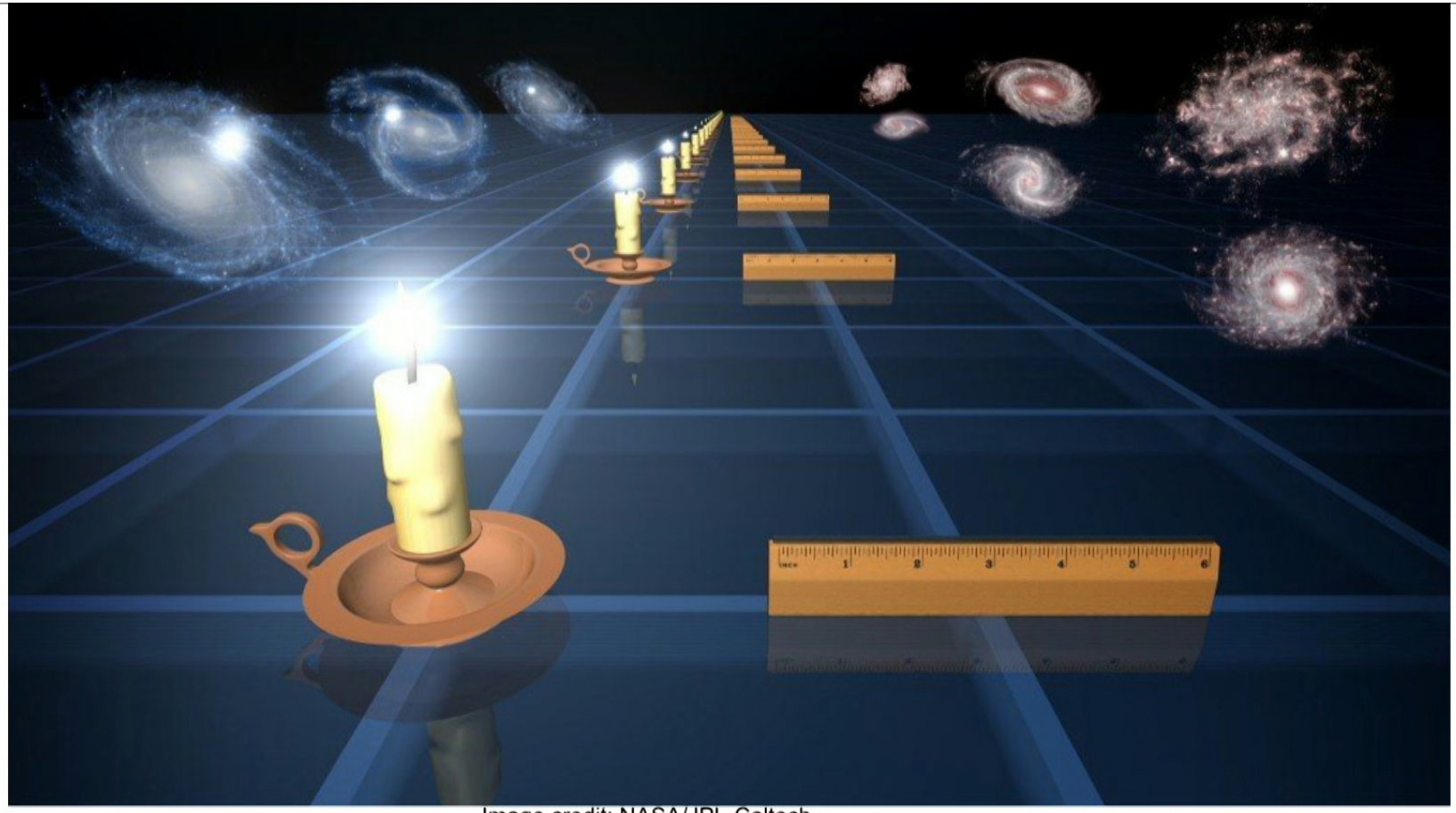


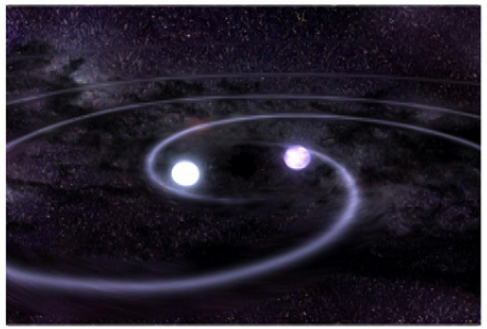
Image credit: NASA/JPL-Caltech

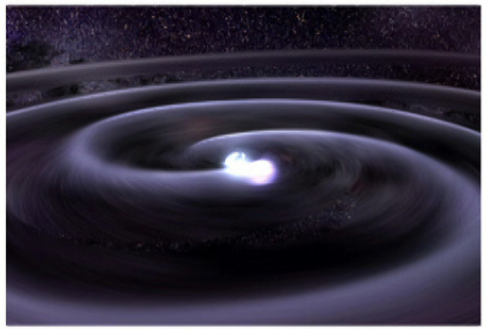
Standard candles
Supernovae Ia

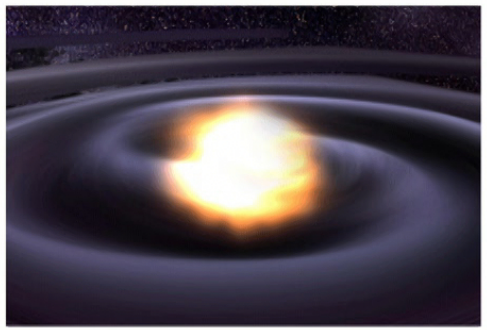
Standard rulers
Baryon Acoustic Oscillations

SNIa









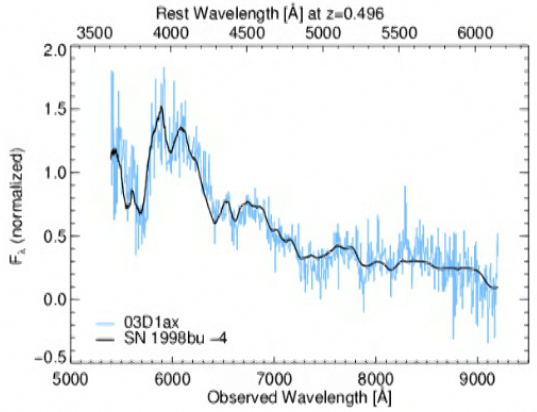
- Very bright ~ host galaxy
- Fast, reach max in ~ 20 days
- Rare ~ 1 / galaxy / millennium
- Luminosity dispersion: 40%

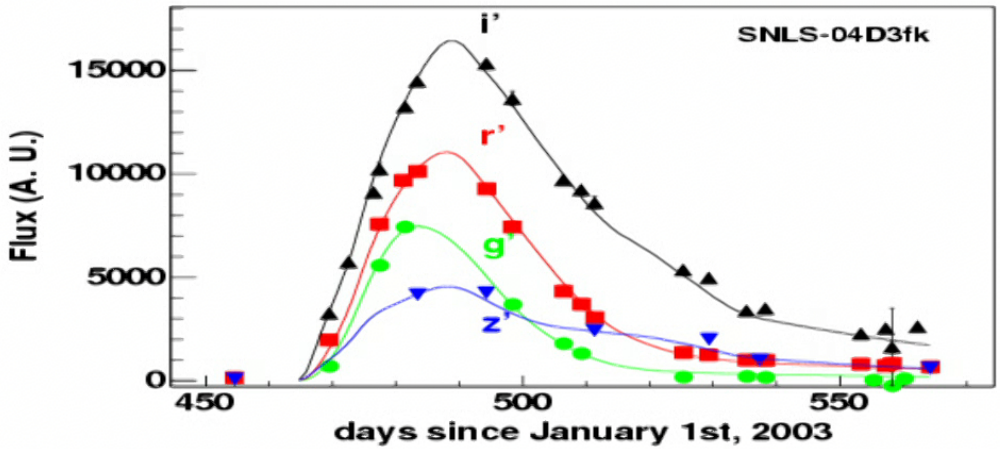
Luminosity Distance: Measure

Spectroscopy

$$d_L^2(z) = \frac{\mathcal{L}_{\rm int}}{4\pi F_{\rm obs}}$$

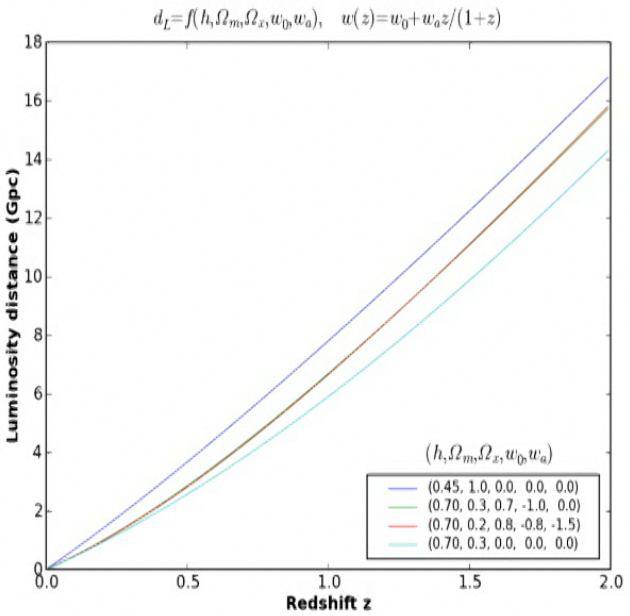
Photometry

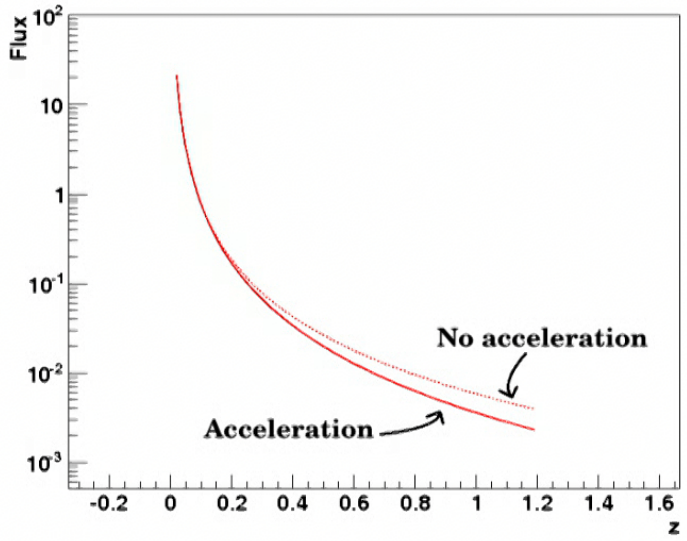




Luminosity Distance: Predict

$$d_L(z) = (1+z)\frac{c}{H_0}\int dz \left(\Omega_m(1+z)^3 + \Omega_x \exp\left(\int_0^z dz' 3\frac{1+w(z')}{1+z'}\right)\right)^{-1/2}$$

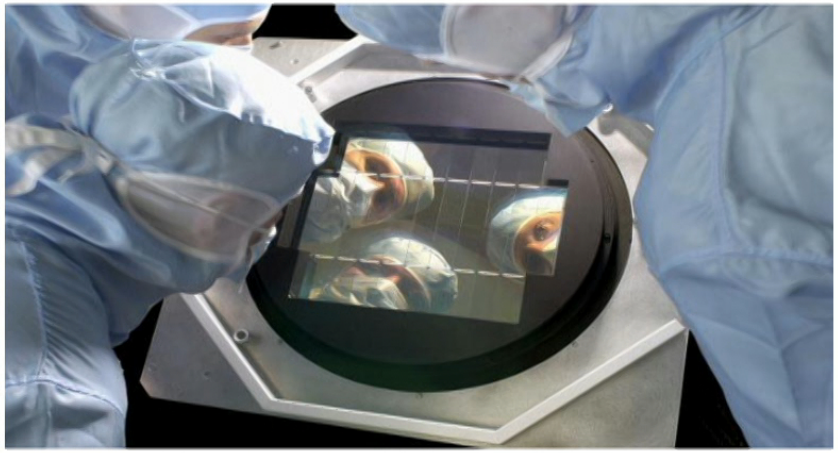




SuperNova Legacy Survey

- 0.2 < z < 0.9
- 4 deg²
- ~10,000 light curves (u*g'r'i'z')
- ~1,000 spectra

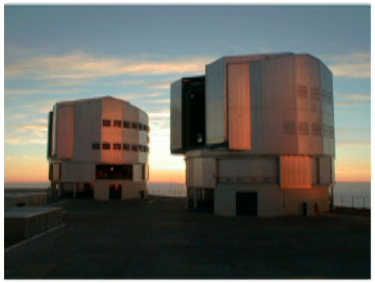
CFHT 3.6m, seeing~0.8"
Megacam – 320Mpixels
Field of view 1 deg²



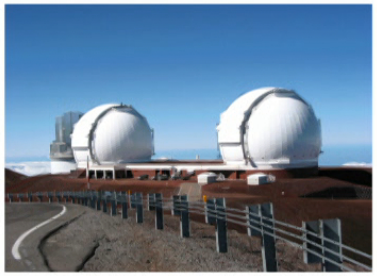
SNLS Telescopes



CFHT (250h/year)
Detection / Photometry
4 bands griz







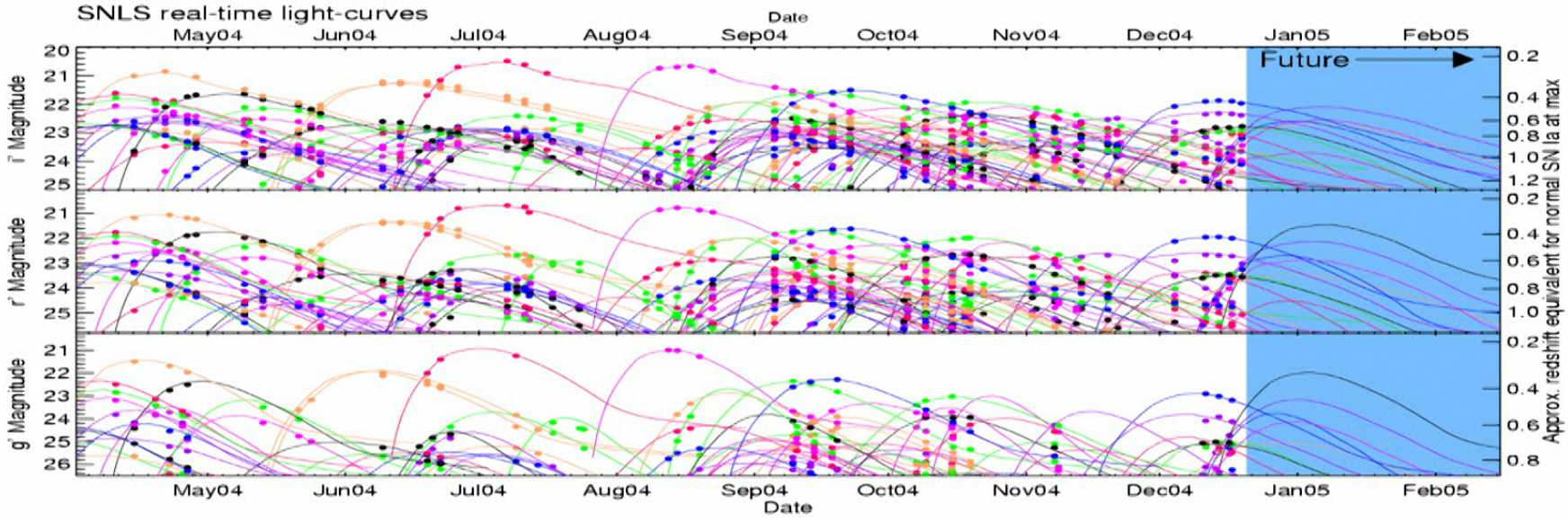
VLT (120 h/year)
SN types/redshifts
$0.3 < z < 0.7$

Gemini (120h/year)
SN types/redshifts
$0.6 < z < 0.9$

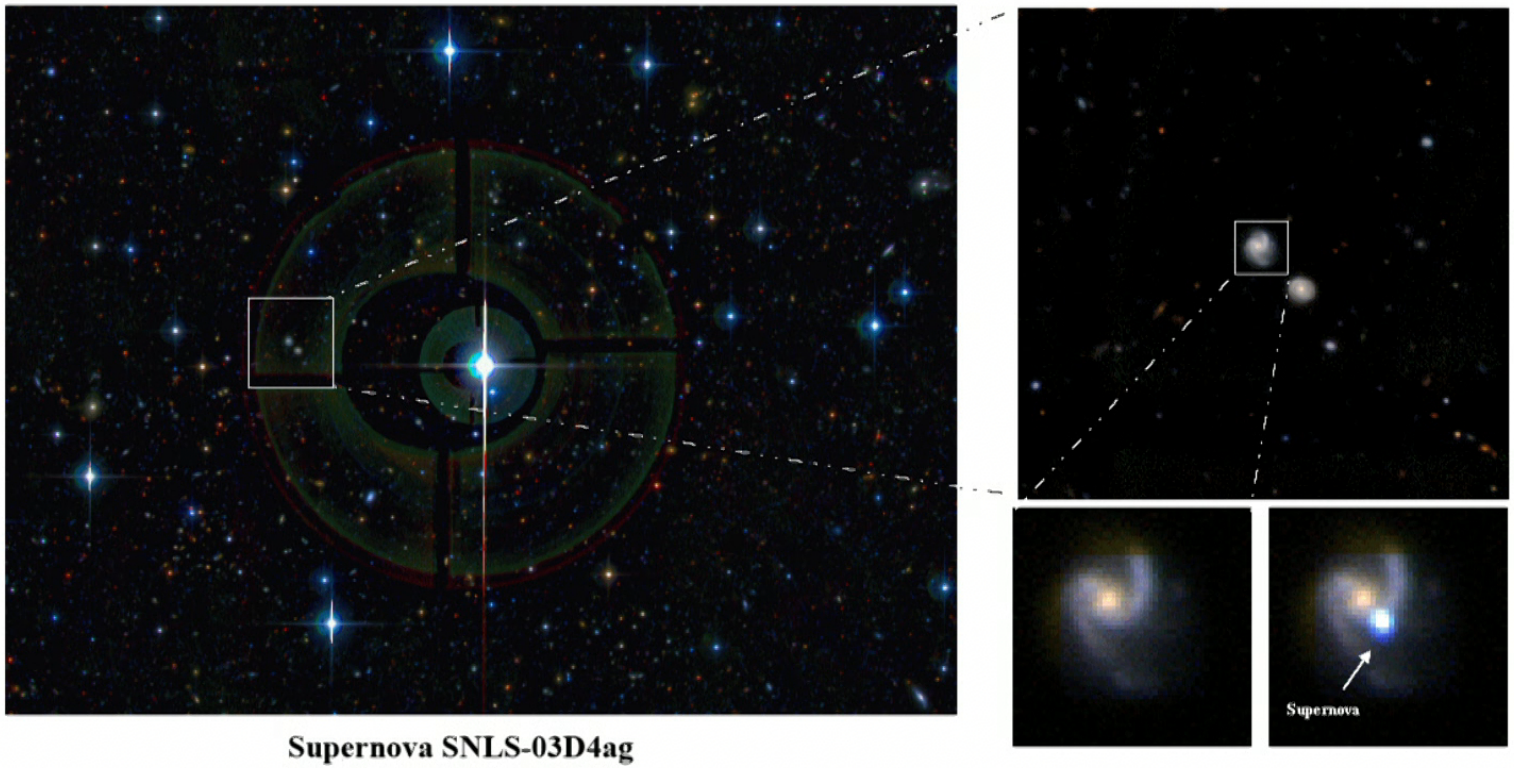
Keck (30h/yr)
- SN types/redshifts
- SN intensivestudy
- SNII

SNLS Imaging Strategy

. New run of 3-4 nights each ~20 nights
. 2 fields in four bands (griz) in field visibility (~6month)

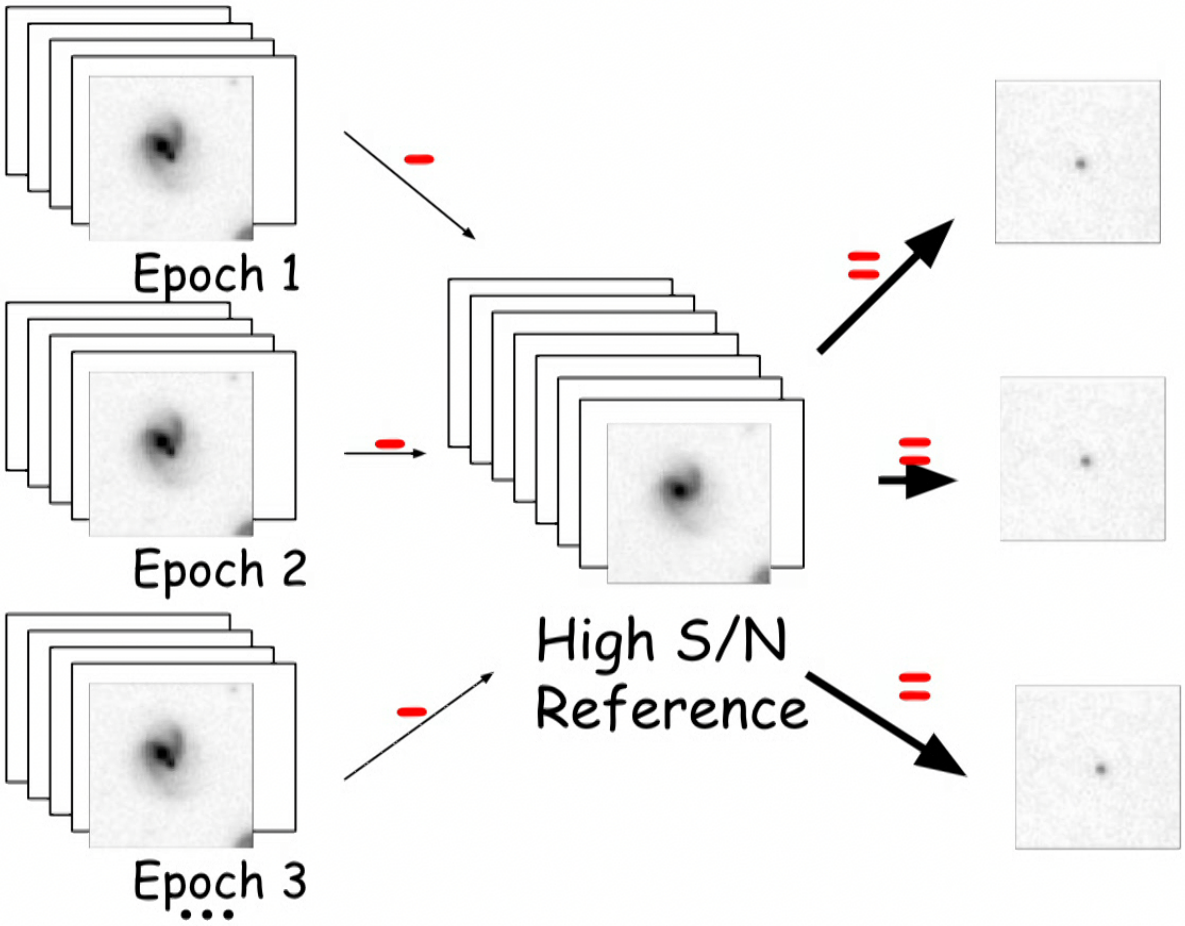


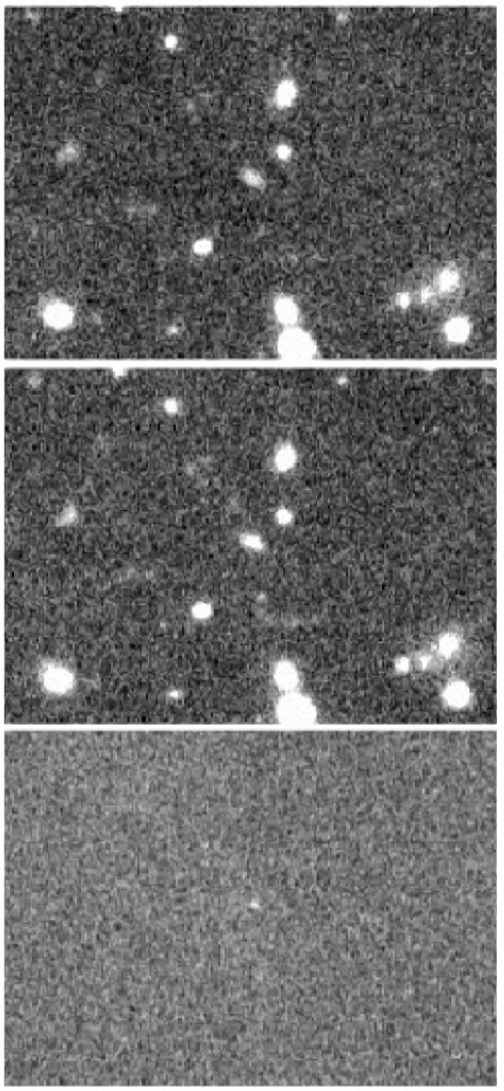
Finding Supernovae



Supernova SNLS-03D4ag

Online Detection



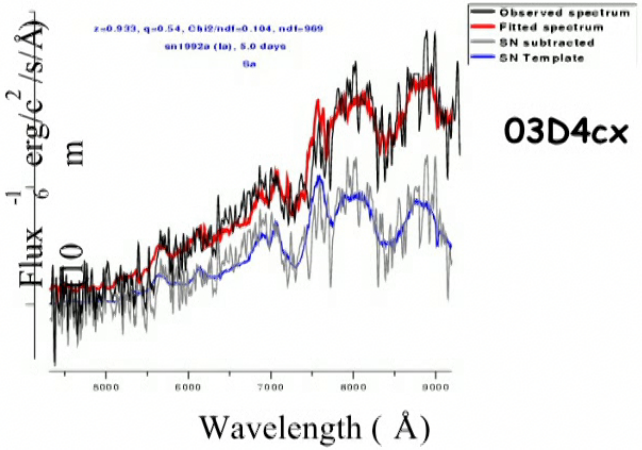


SN Spectroscopy

Redshift: 0.932 ±0.007

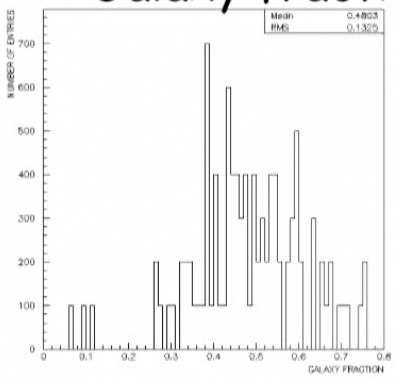
Galaxy fraction: 48 ±13 %

SN age: 2 ±2 days

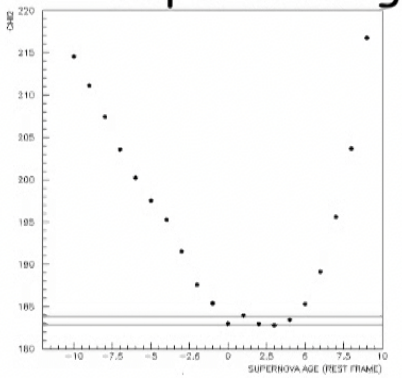


03D4cx

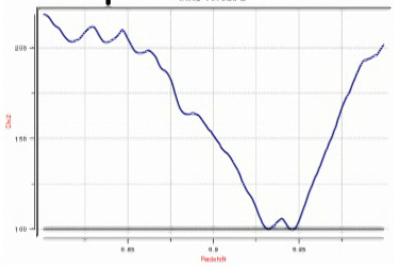
Galaxy fraction



Supernova age

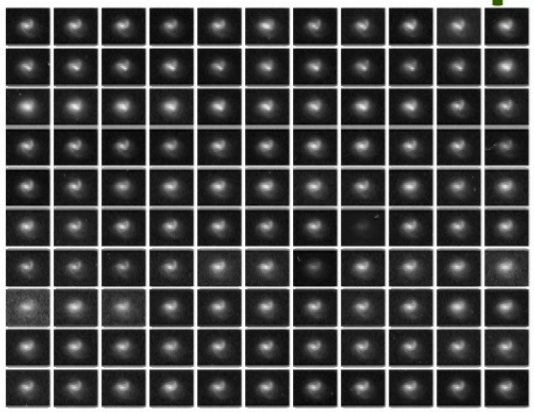


Supernova redshift

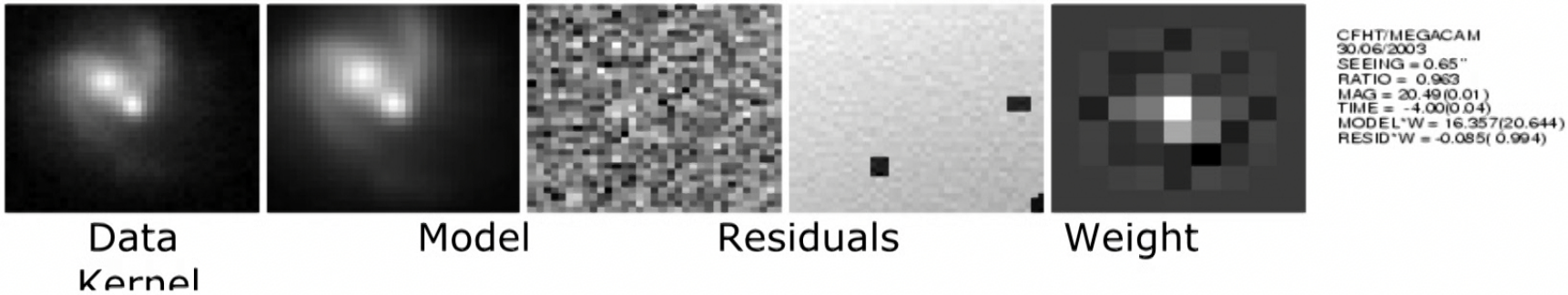


Precise Photometry After the Fact

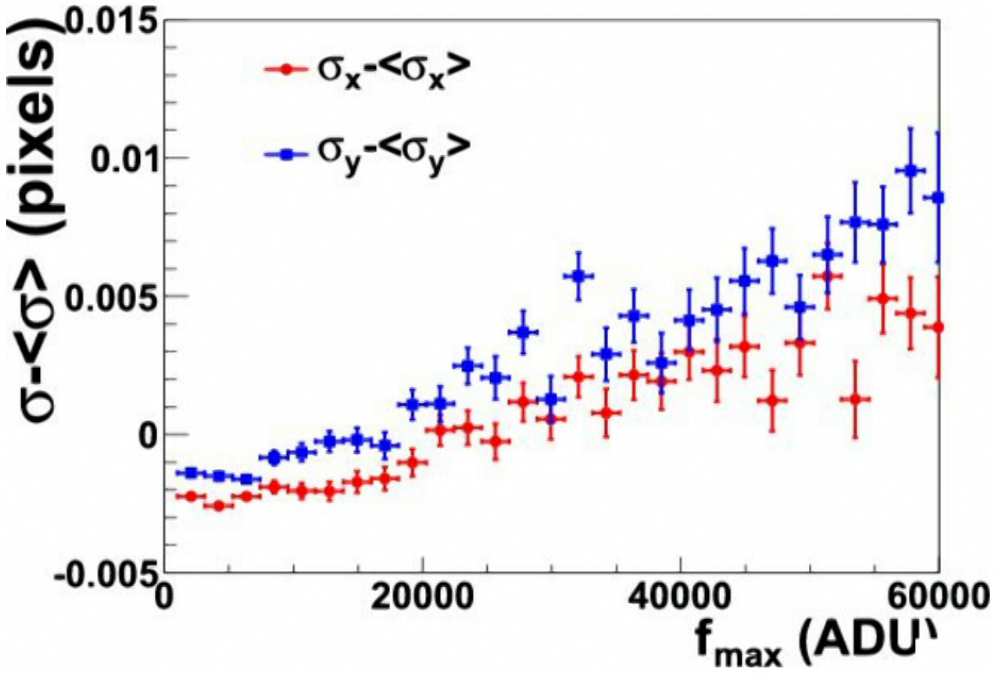
Fabbro et. SCP (2001)
Astier et. SNLS (2013)



- Image sequences
- Multi filters
- Better calibration
- Forward modeling

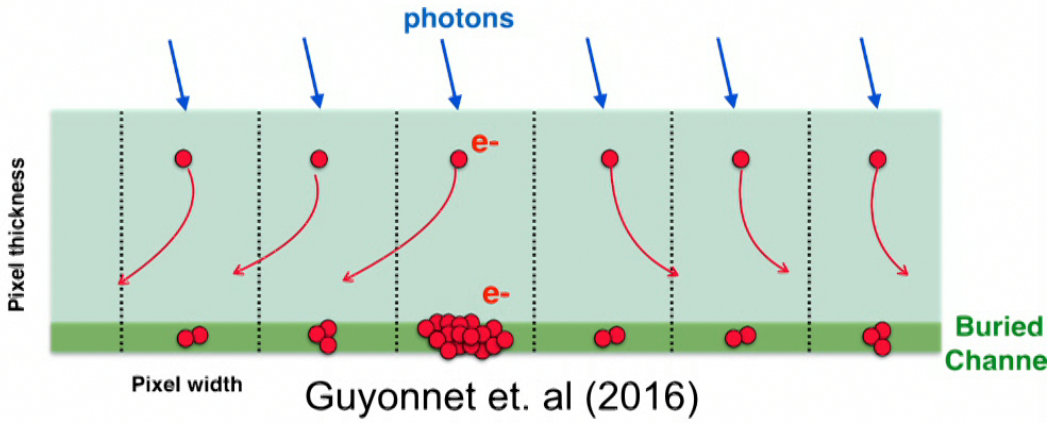


$$M_{i,p} = \left[f_i \times \phi_i(\vec{x}_p - T_i[\vec{x}_{obj}]) + G\left(T_i^{-1}(\vec{x}_p)\right) \otimes K_i + S_i \right] R_i$$



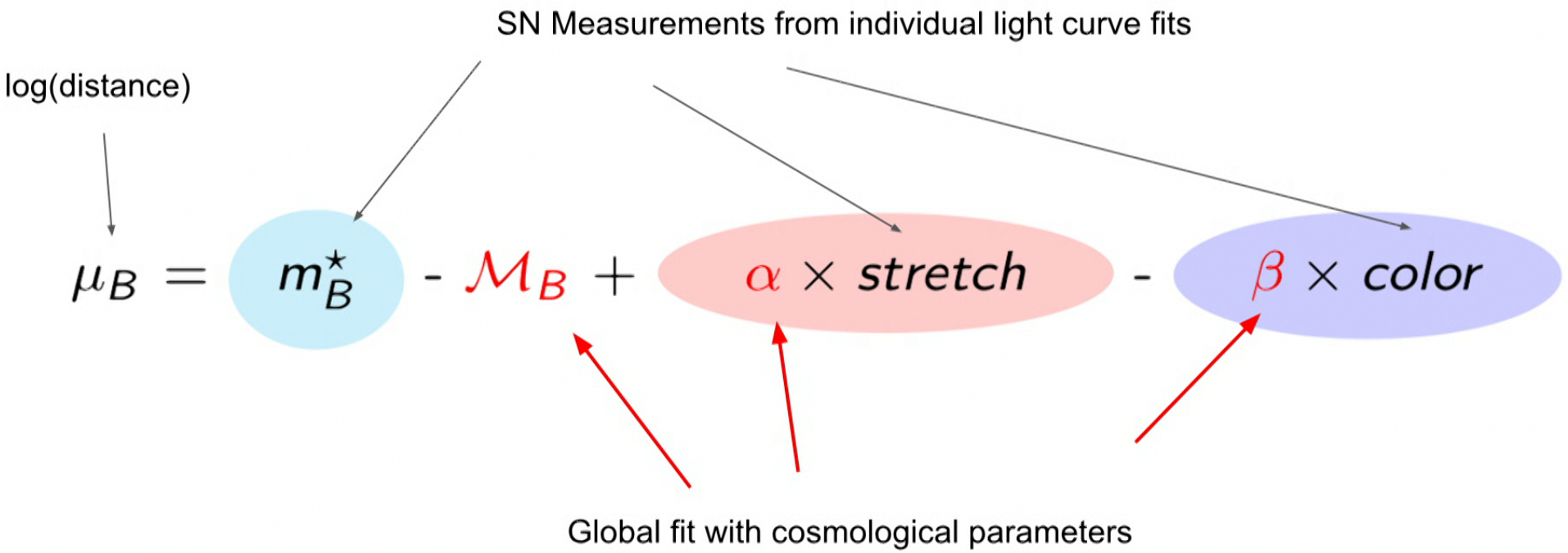
Astier et. SNLS (2013)
CFHT Megacam
SNLS Data
Antilogus et. al (2013)

Pushing CCD limits
ex: "brighter/fatter" effect

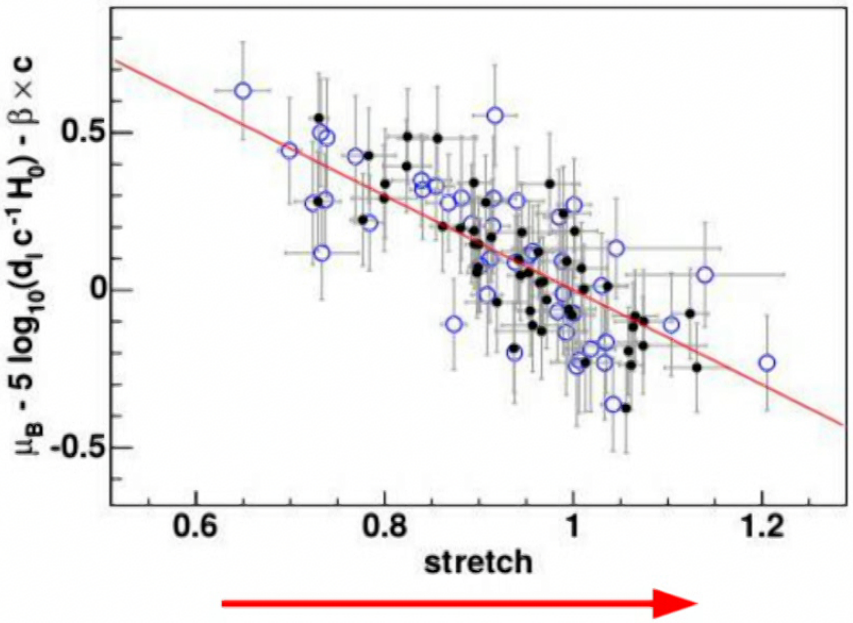


Guyonnet et. al (2016)

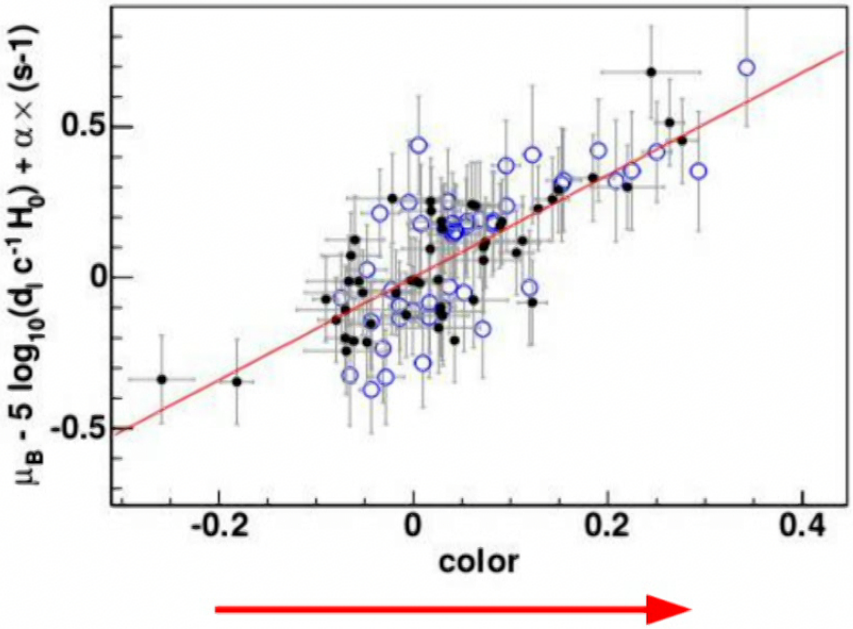
Back to Supernova Distance Estimation



Residuals to Hubble Diagram without Corrections

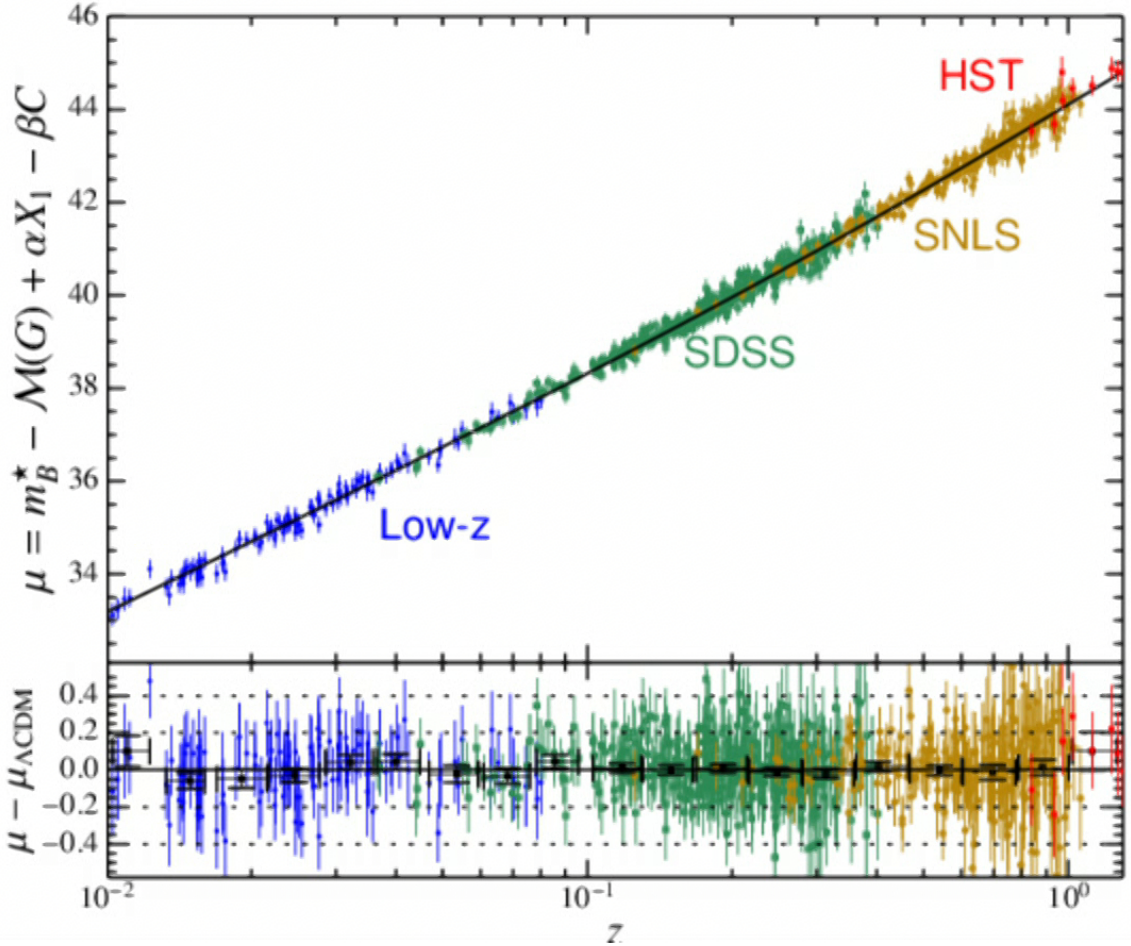


Width of the
Light curve



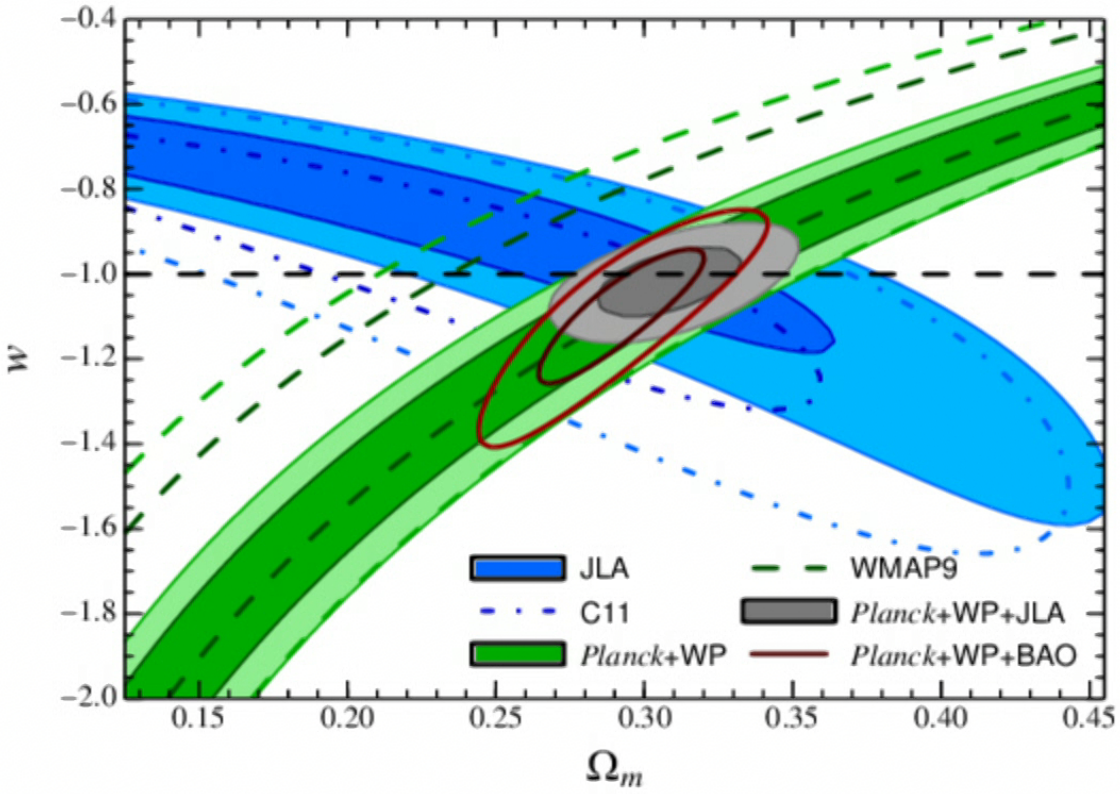
Colour : « B-V »
Green flux/Blue flux

Joint Lightcurve Analysis Distances



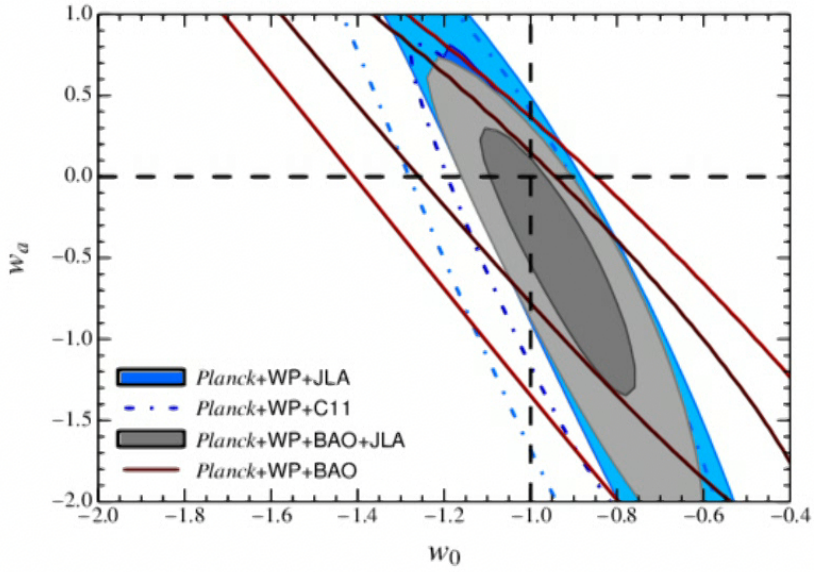
- 740 SNIa:
  - 118 from 5 nearby surveys
  - 366 SDSS
  - 242 SNLS
  - 14 HST
- joint calibration effort
- stat error ~ syst. error.

Betoule et SNLS/SDSS (2014)



w = −1.027 ± 0.055
Ωm = 0.303 ± 0.012

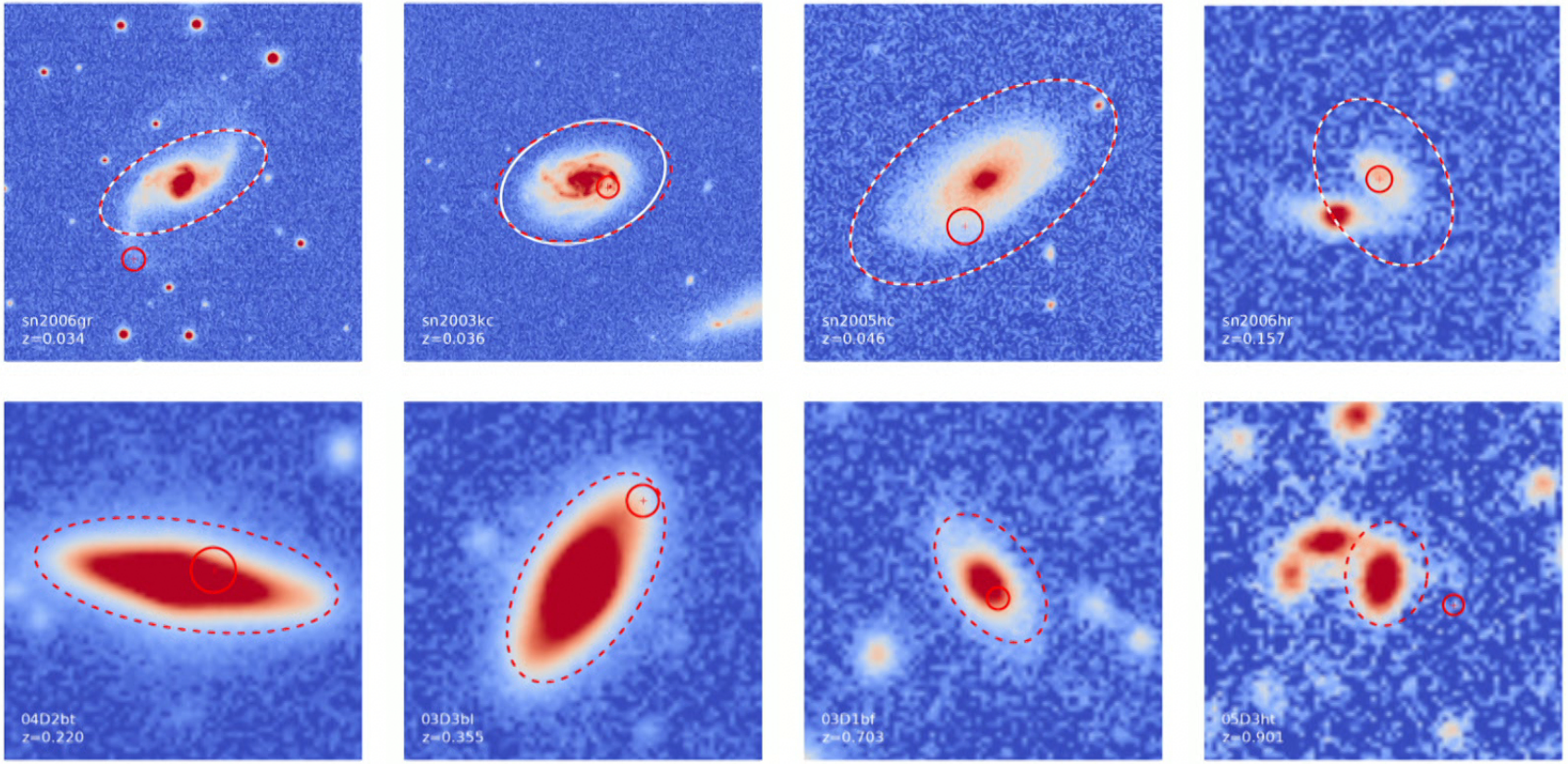
w(z) = w

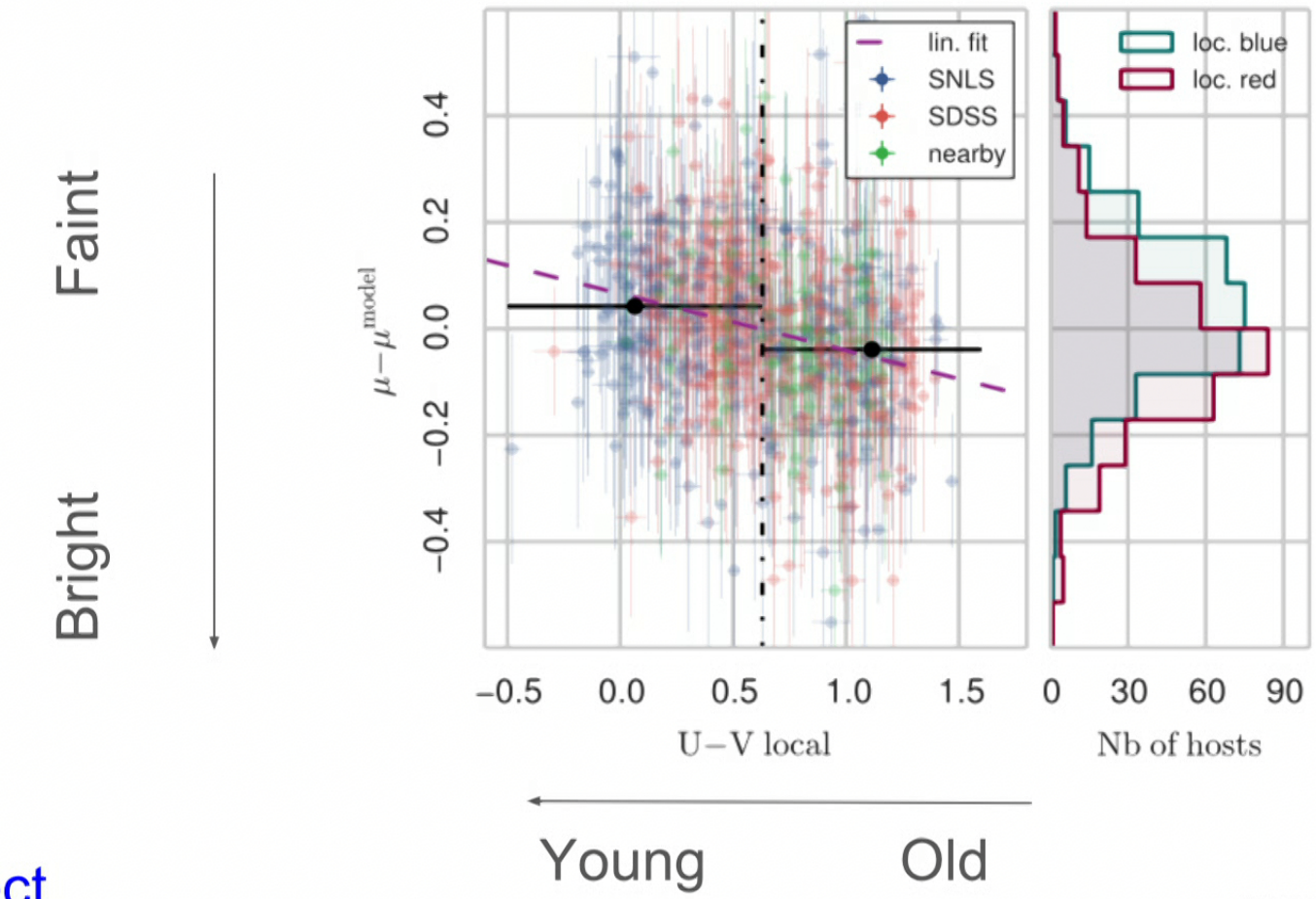


w0 = −0.957 ± 0.124
wa = −0.336 ± 0.552

$w(z) = w_0 + w_a\, z/(1+z)$

SN Environment as a 3rd parameter?





882 Supernovae

7σ effect
impact on DE δω~1%

Roman et SNLS (2018)

Beyond Dark Energy with Supernovae

- Hubble Constant from nearby SN
- Dark matter from lensed SN
- Testing gravitational theories with low-redshift peculiar motions
- SN progenitors
- Nuclear astrophysics / stellar evolution
- Star formation studies with SN rates
- Time dilation from SN light curves