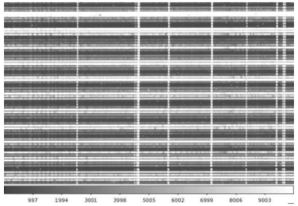
machine learning for stellar spectra

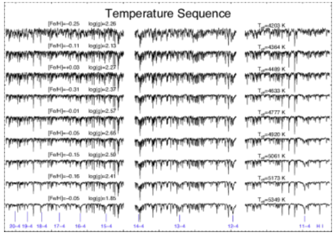
Sébastien Fabbro

NTCO 2017

raw 2D image



calibrated 1D spectra

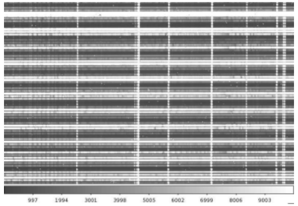


estimated stellar parameters

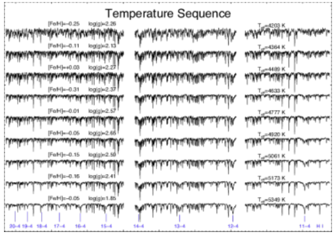
T_{eff}, log(g), [Fe/H], [C/H], [N/H], [O/H], [Na/H], [Mg/H], [Al/H], [Si/H], [S/H], [K/H], [Ca/H], [Ti/H], [V/H], [Mn/H], [Ni/H],...

stellar populations
galactic archeology
near-field cosmology

synthetic 2D image

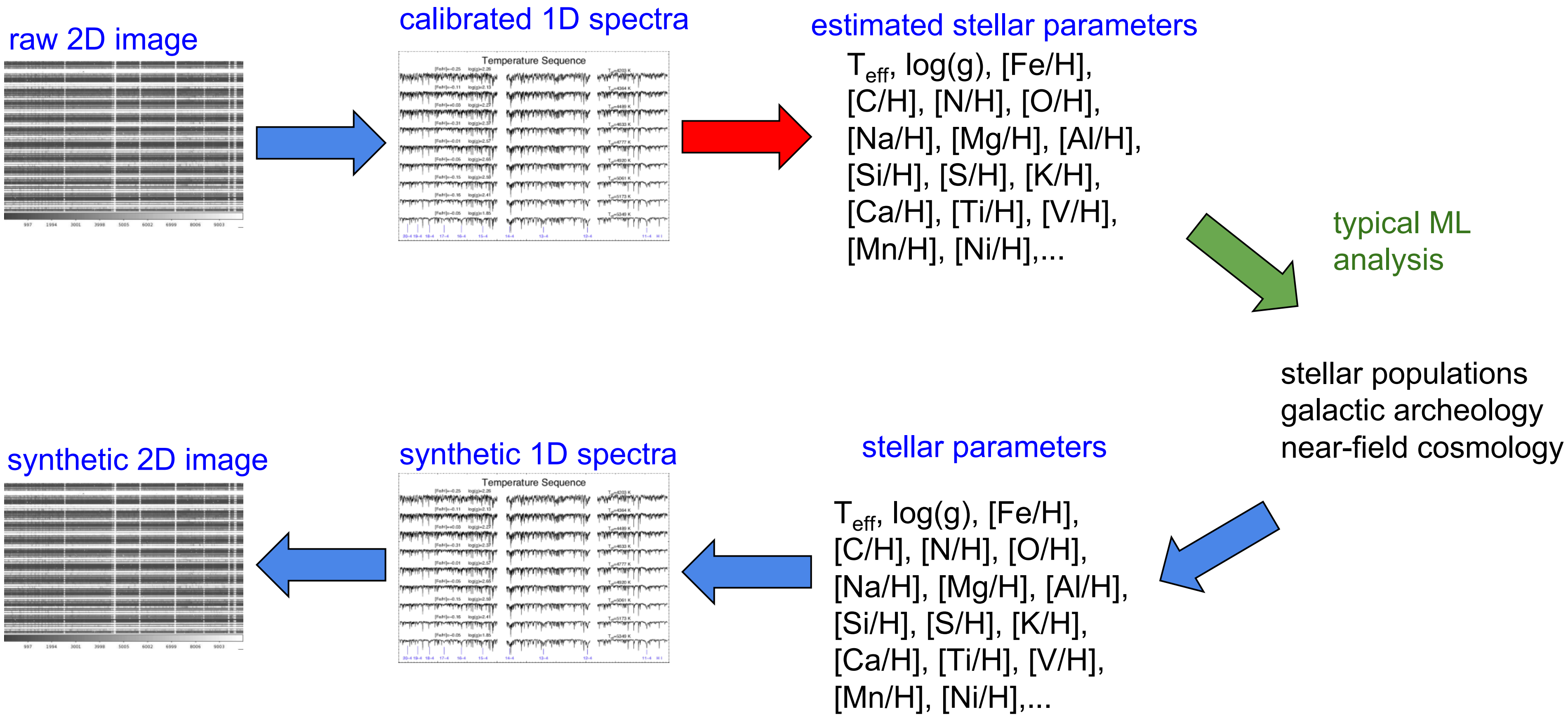


synthetic 1D spectra



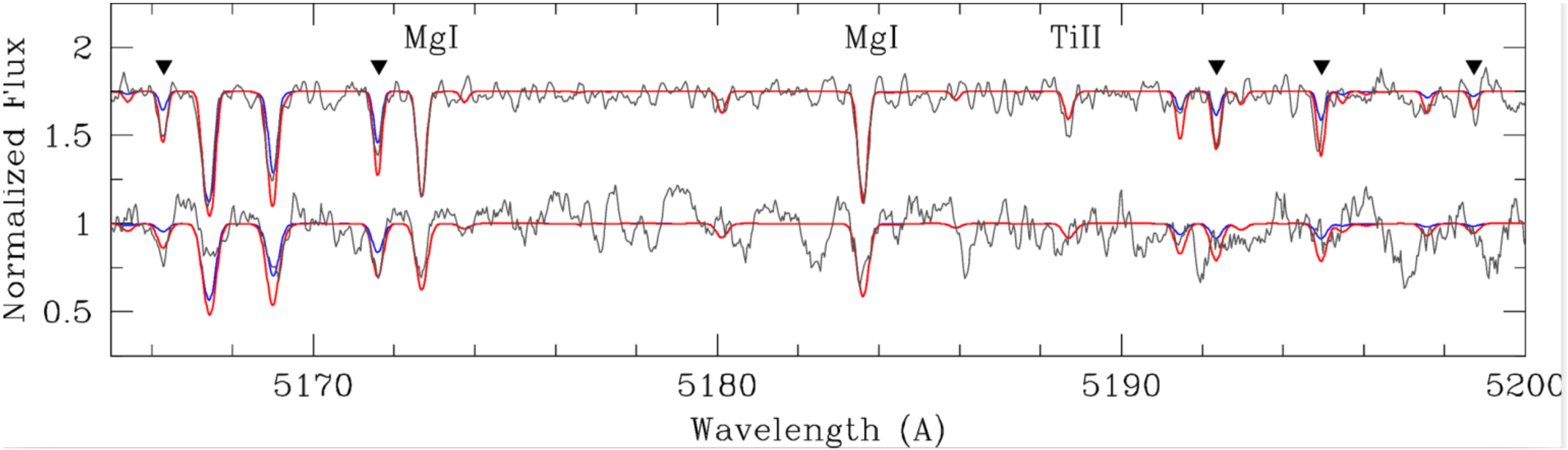
stellar parameters

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],...


raw 2D image
calibrated 1D spectra
Temperature Sequence
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],...
typical ML analysis
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
Temperature Sequence
synthetic 2D image

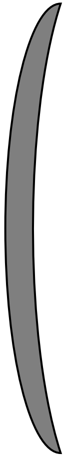
traditional approach

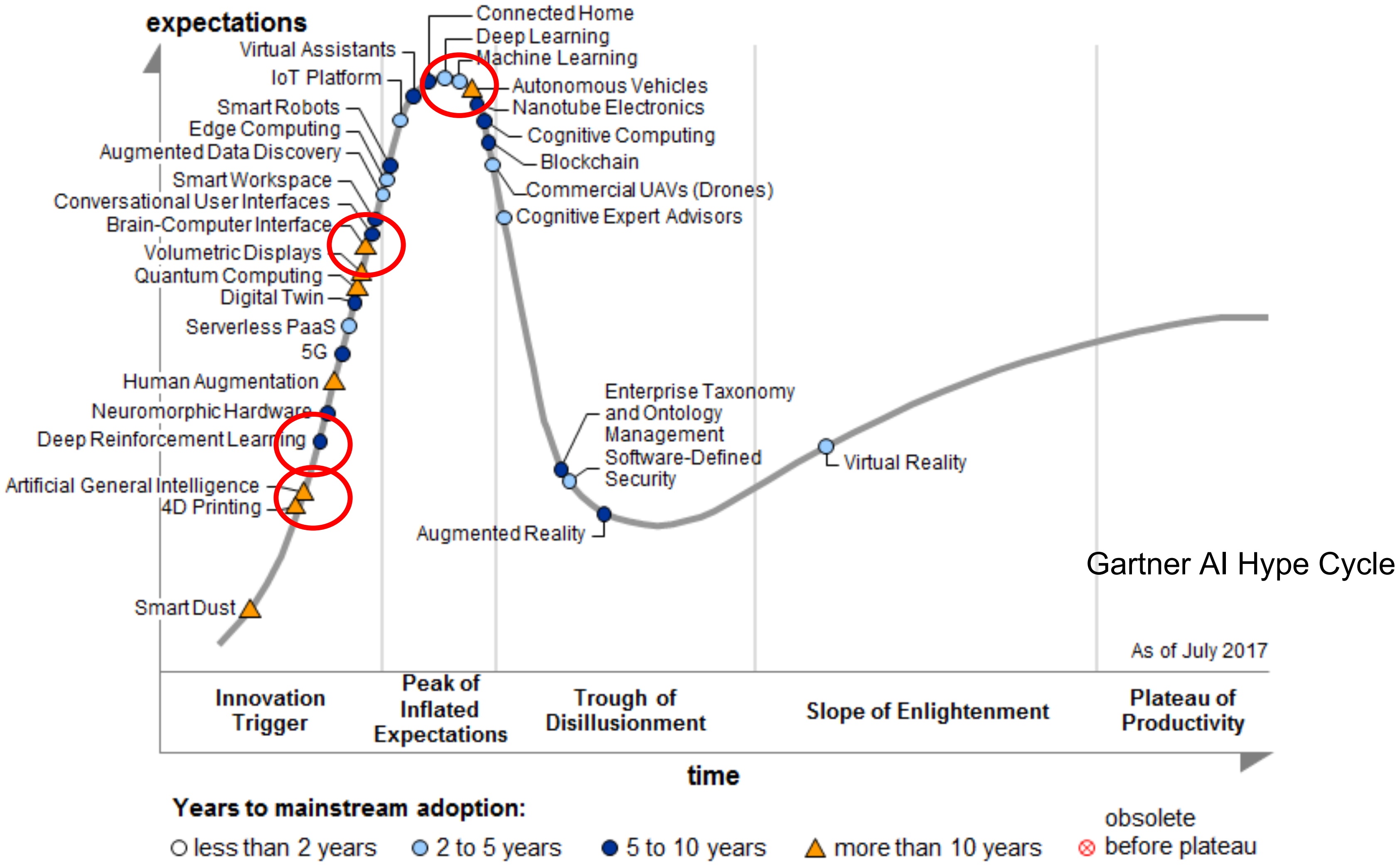
Gemini-GRACES spectroscopy, Venn et. al (2017)



- physics models methods (EQW, NL fit, projections…)
- "data driven" methods (The Cannon, ANN): need a set of reference stars

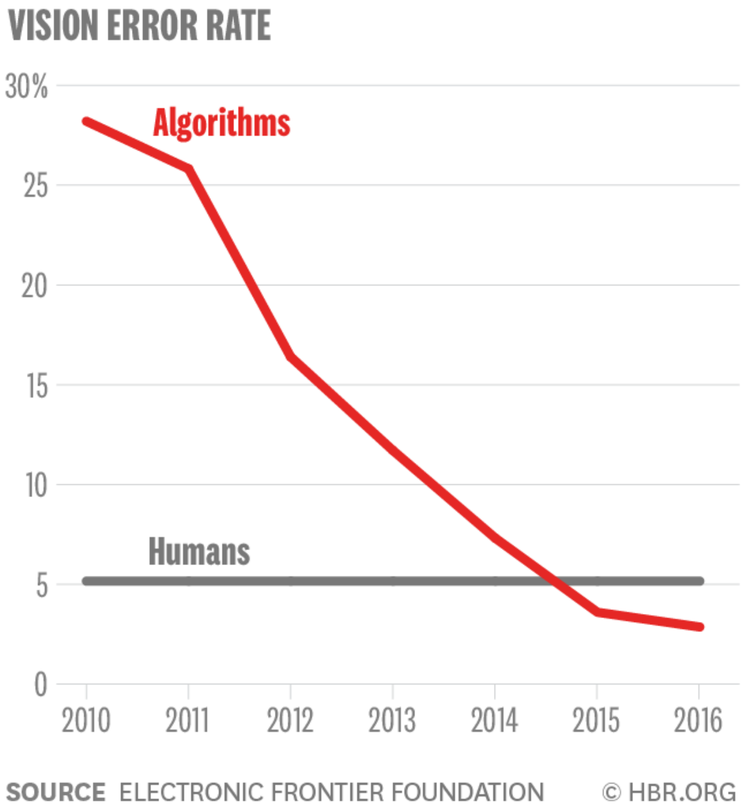
AI





expectations
Connected Home
Deep Learning
Machine Learning
Virtual Assistants
IoT Platform
Autonomous Vehicles
Nanotube Electronics
Smart Robots
Edge Computing
Cognitive Computing
Augmented Data Discovery
Blockchain
Smart Workspace
Commercial UAVs (Drones)
Conversational User Interfaces
Cognitive Expert Advisors
Brain-Computer Interface
Volumetric Displays
Quantum Computing
Digital Twin
Serverless PaaS
5G
Human Augmentation
Neuromorphic Hardware
Deep Reinforcement Learning
Enterprise Taxonomy and Ontology Management
Software-Defined Security
Virtual Reality
Artificial General Intelligence
4D Printing
Augmented Reality
Smart Dust
Gartner AI Hype Cycle
As of July 2017
Innovation Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateau of Productivity
time
Years to mainstream adoption:
less than 2 years
2 to 5 years
5 to 10 years
more than 10 years
obsolete before plateau

recognize puppy from muffins



translate from 100 languages

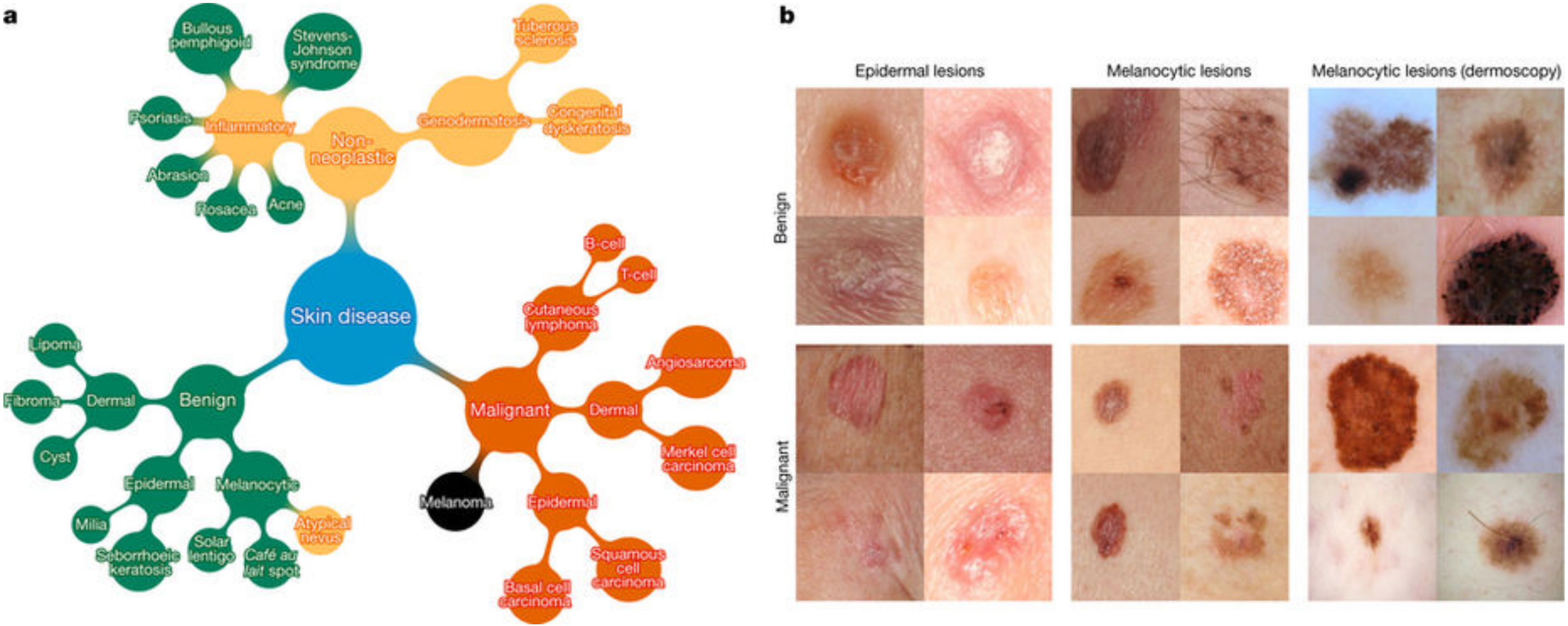
google

restyle paintings

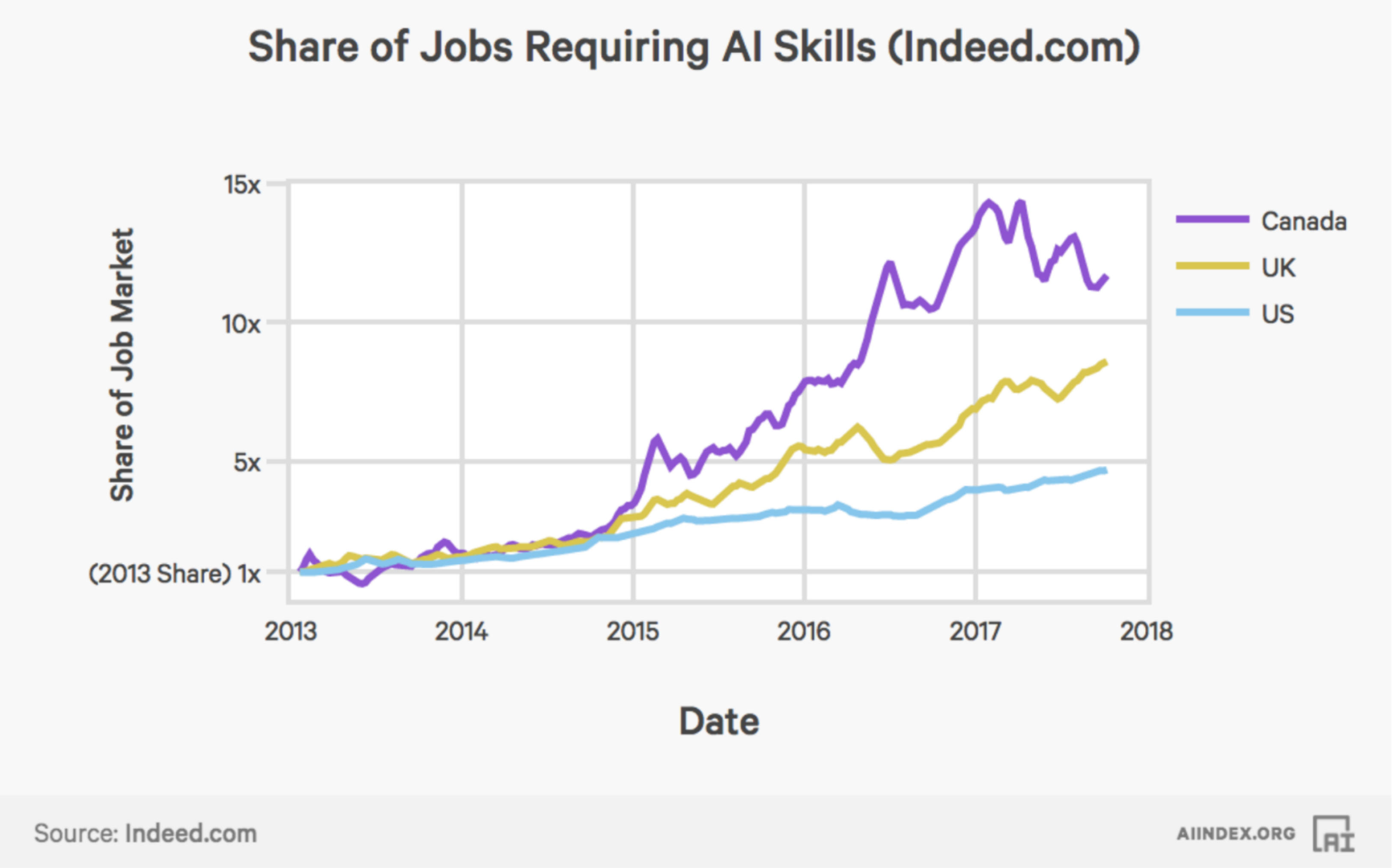
generate new celebrities

Karas et. al (2017)

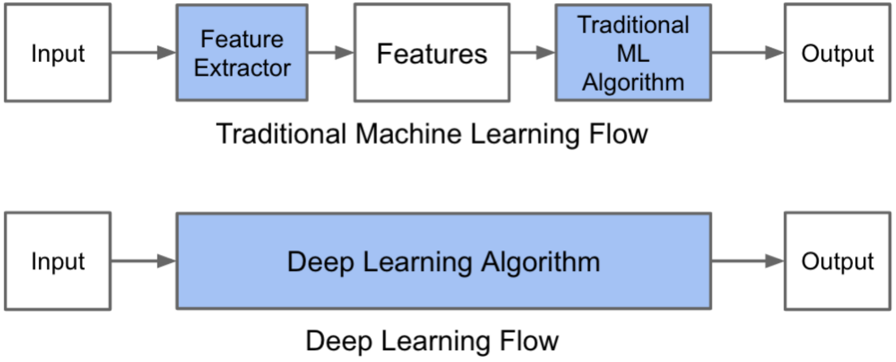
detect skin cancer

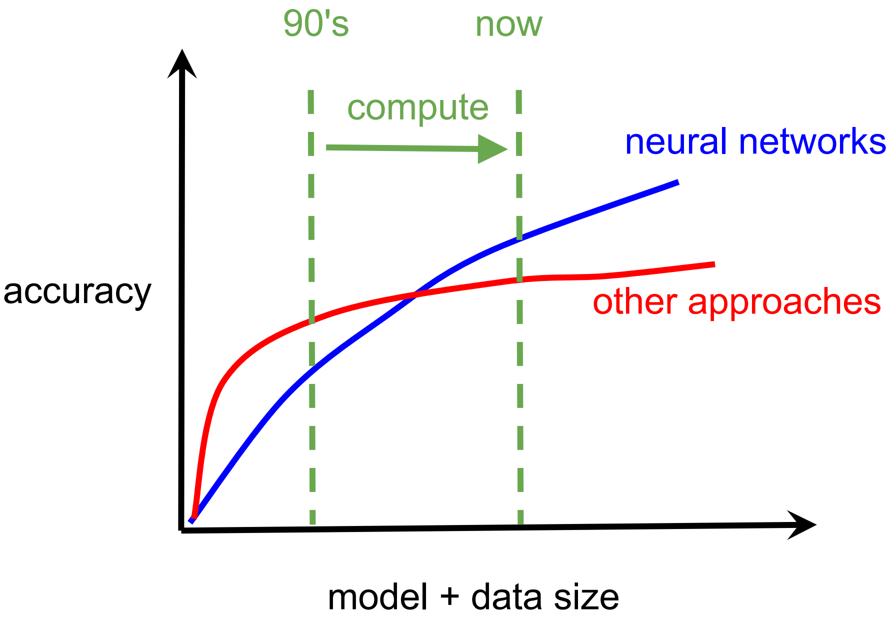


Esteva et. al (2017)


Share of Jobs Requiring AI Skills (Indeed.com)
15x
10x
5x
(2013 Share) 1x
Share of Job Market
2013
2014
2015
2016
2017
2018
Date
Canada
UK
US
Source: Indeed.com
AIINDEX.ORG

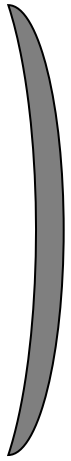
deep learning



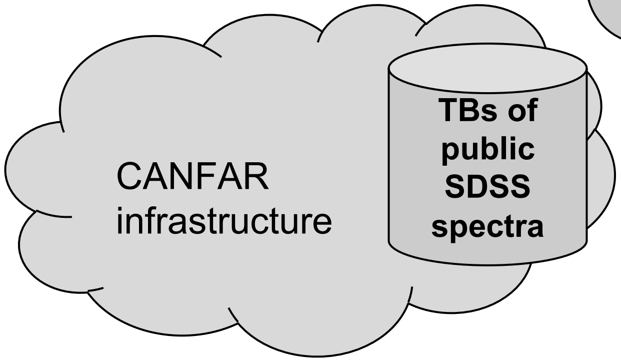


should astronomers care?

- detect complex structures in large data sets
- classify astronomical objects
- learn time consuming simulations
- automate and accelerate manual data analysis tasks
- replace many image processing techniques
- powerful well written software
- prepare students

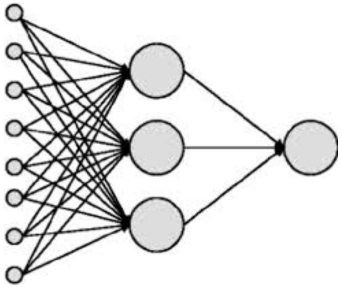


an experiment



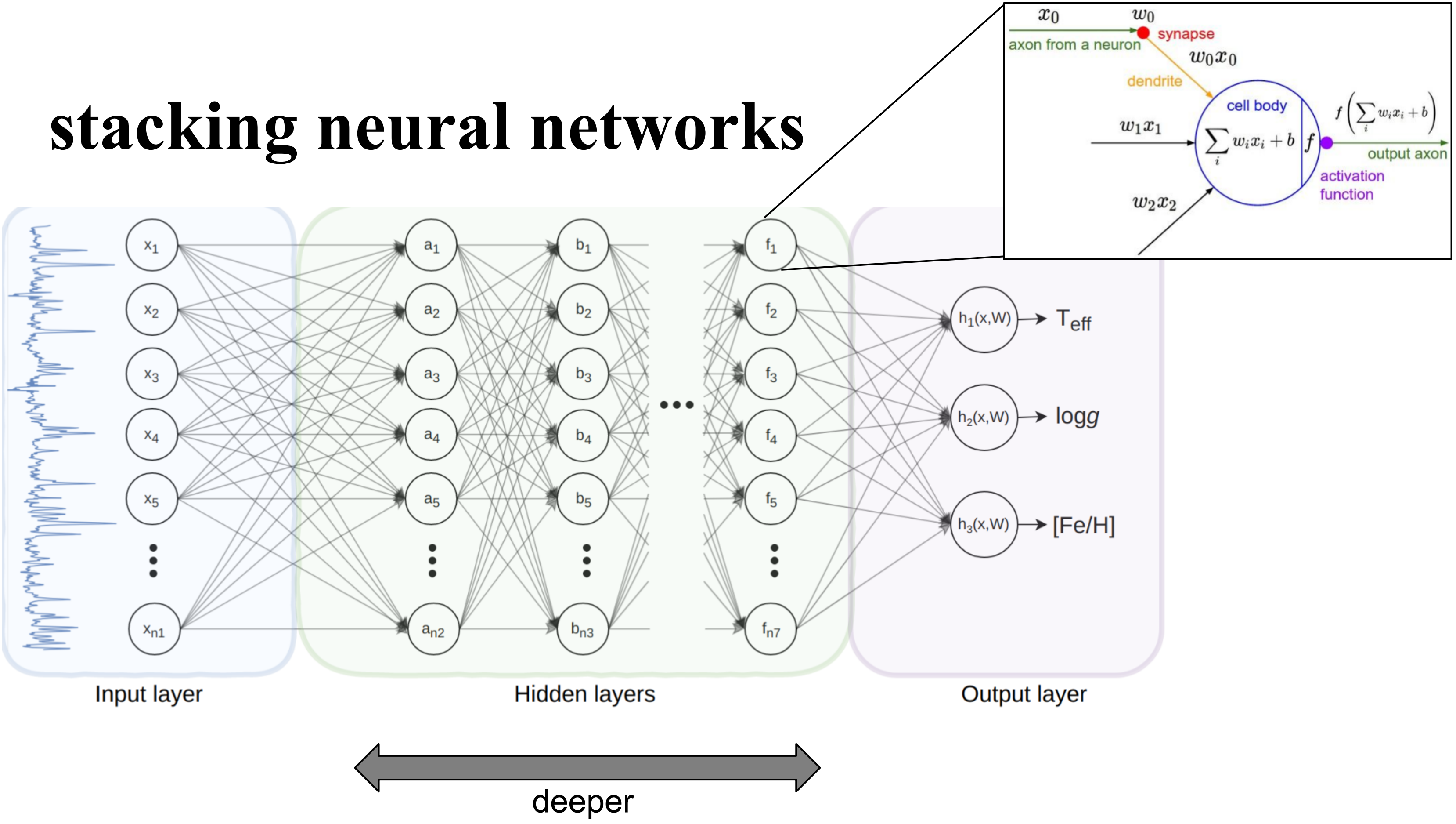

```
google: tensorflow, keras
facebook: caffe2, pytorch
amazon: mxnet, gluon
microsoft: CNTK
```

many choice of open DL frameworks

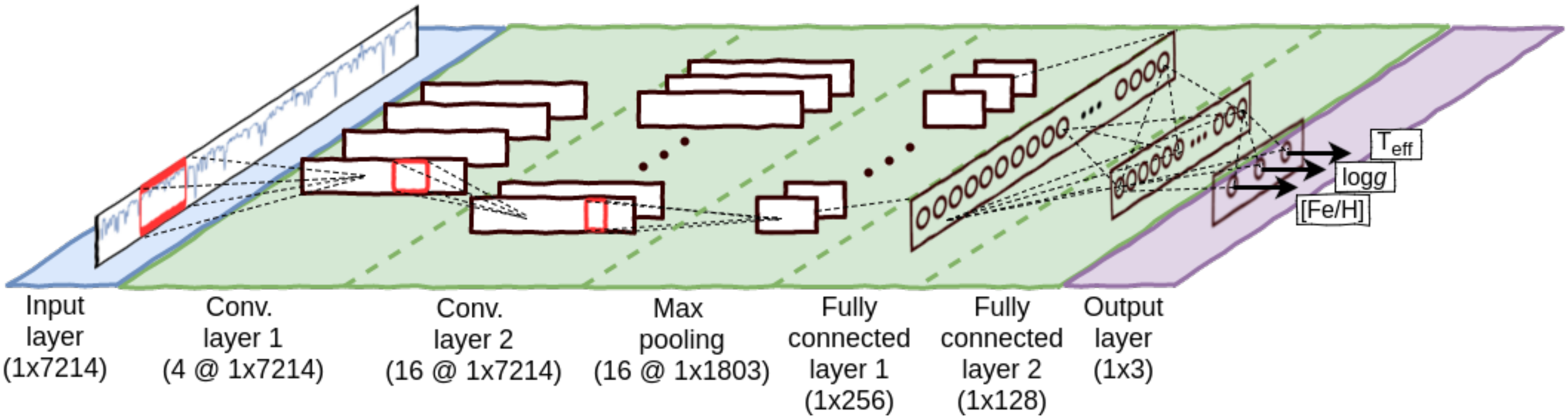


distant past ML knowledge

stacking neural networks

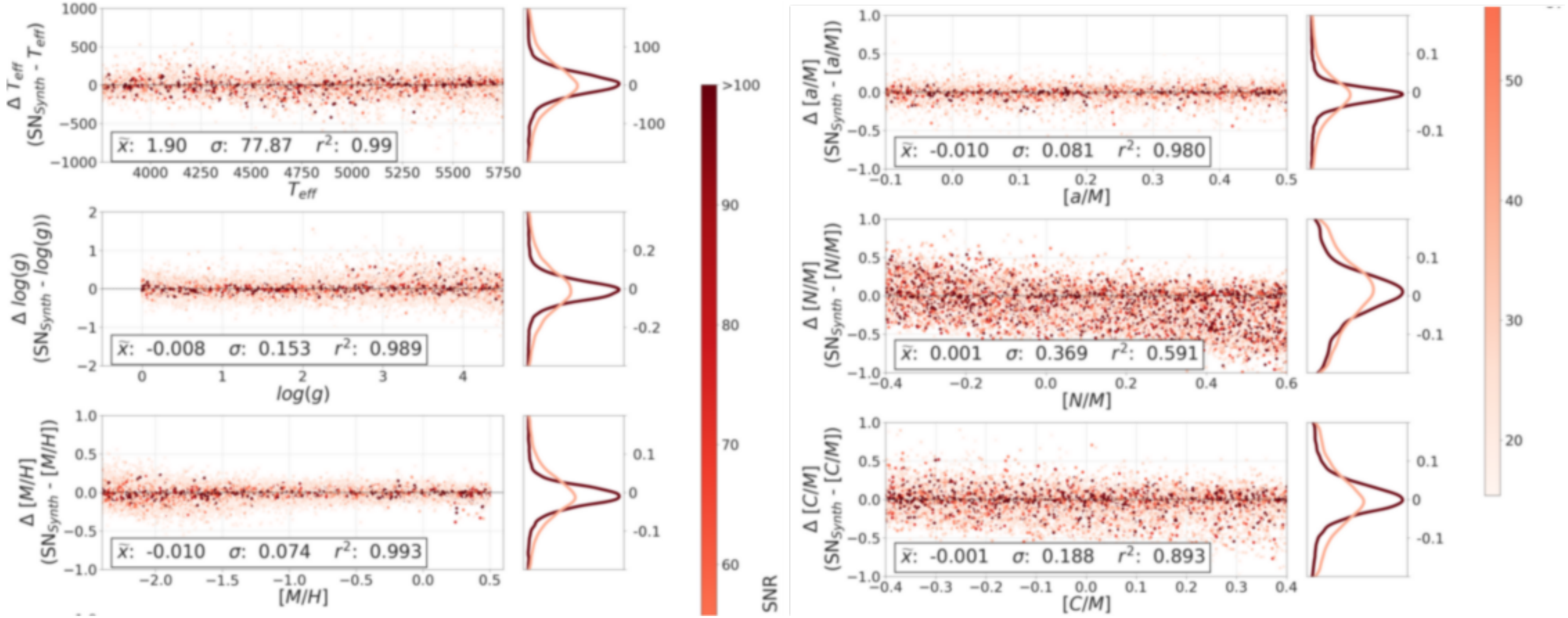


StarNet



T_{eff}
log$g$
[Fe/H]
Input layer (1x7214)
Conv. layer 1 (4 @ 1x7214)
Conv. layer 2 (16 @ 1x7214)
Max pooling (16 @ 1x1803)
Fully connected layer 1 (1x256)
Fully connected layer 2 (1x128)
Output layer (1x3)

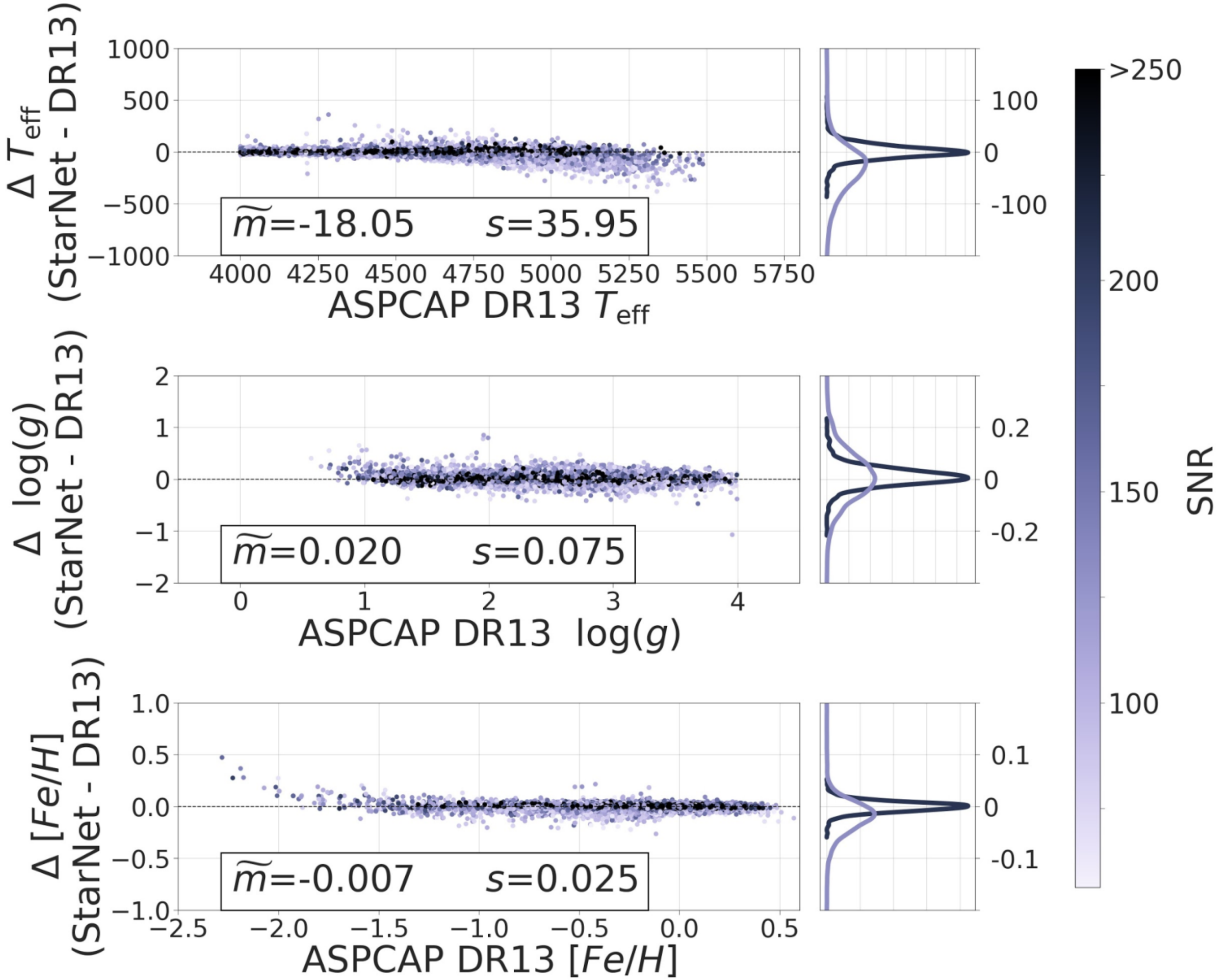
starnet trained on synthetic spectra



starnet trained on APOGEE spectra

15k stars with 47k visits

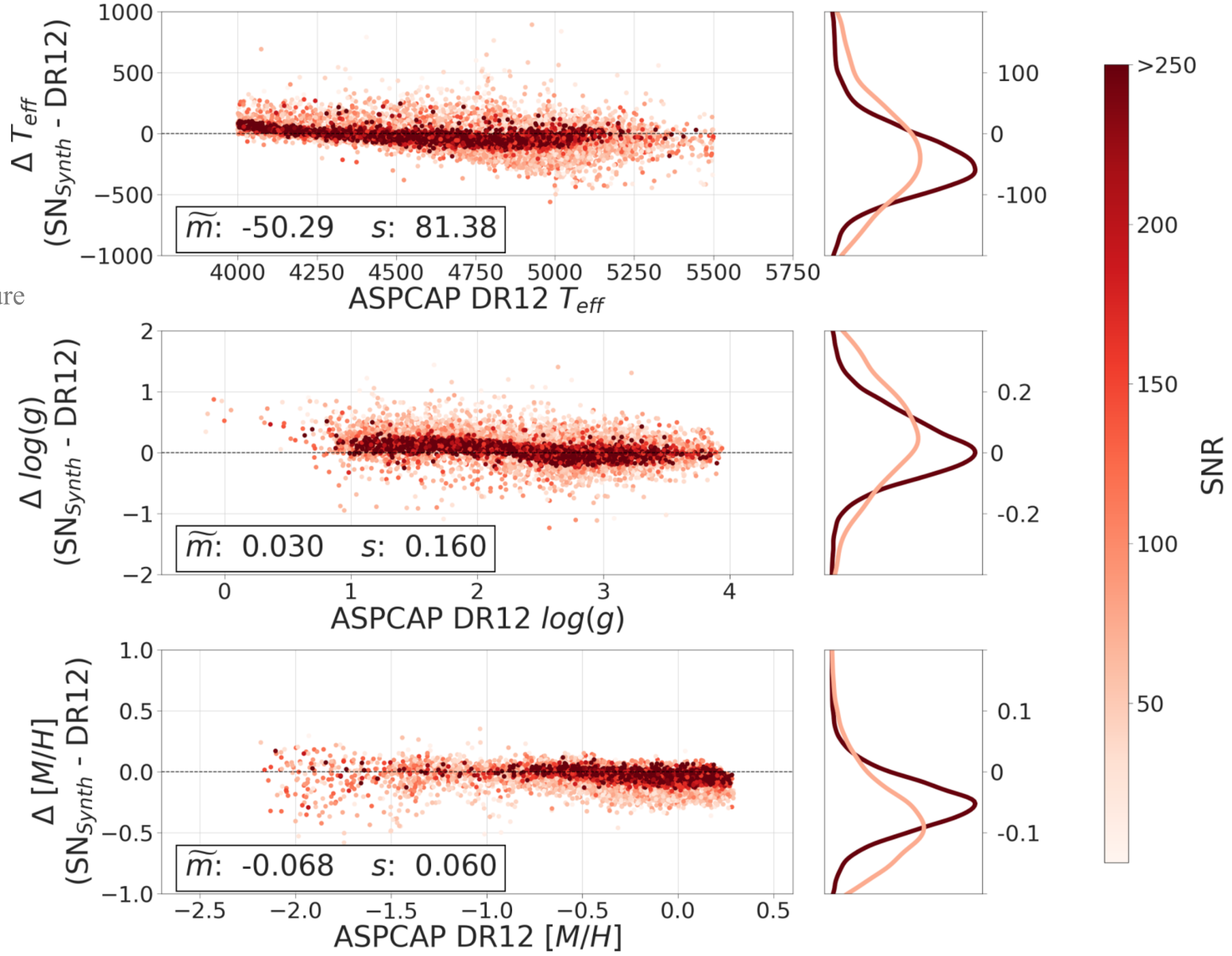
official pipeline parameters



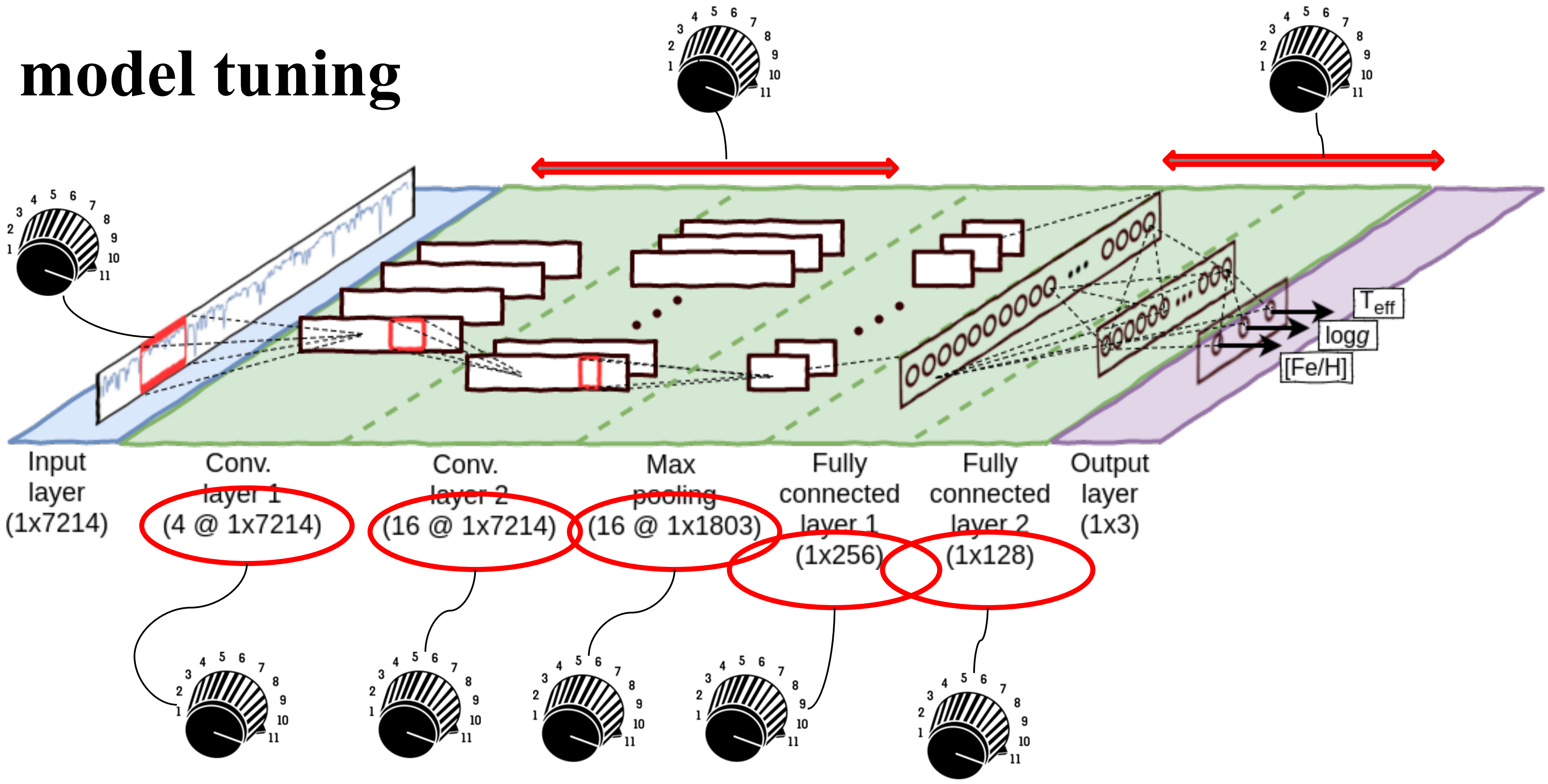
trained on synth
test on APOGEE

no modification to StarNet architecture

same 224k stars as simulations



model tuning



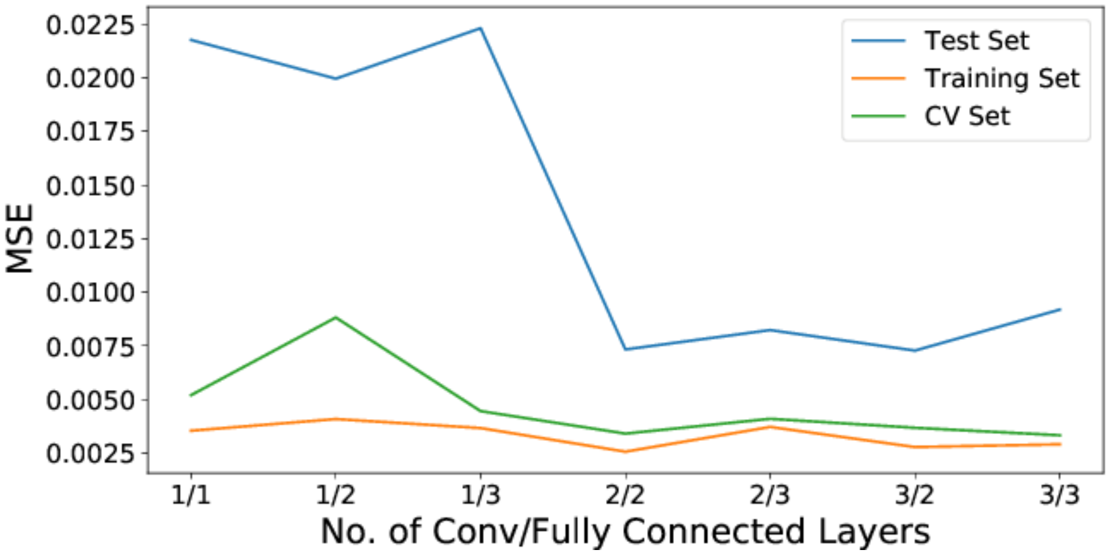
meta learning starnet

13 hyperparameters

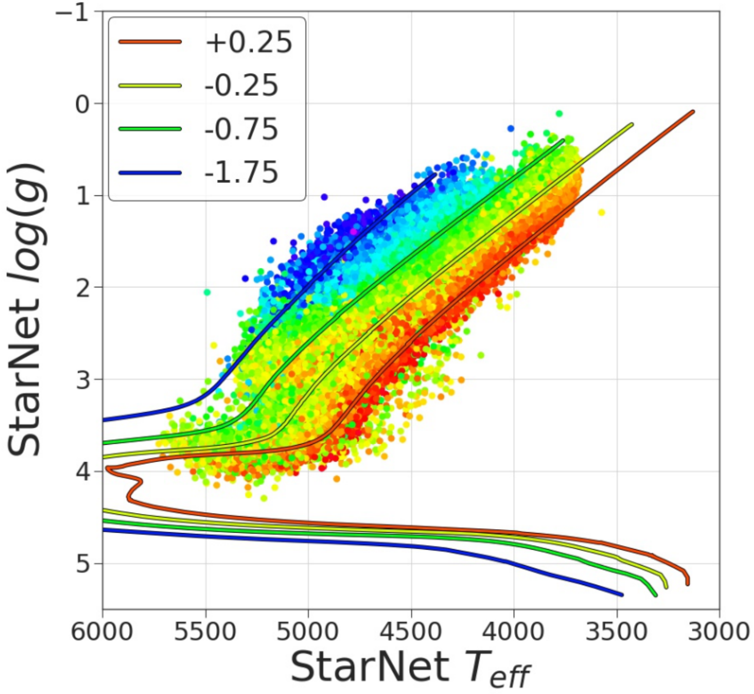
30mn per training (no GPU)

Tree-structured Parzen
Estimator Bayes Optimizer

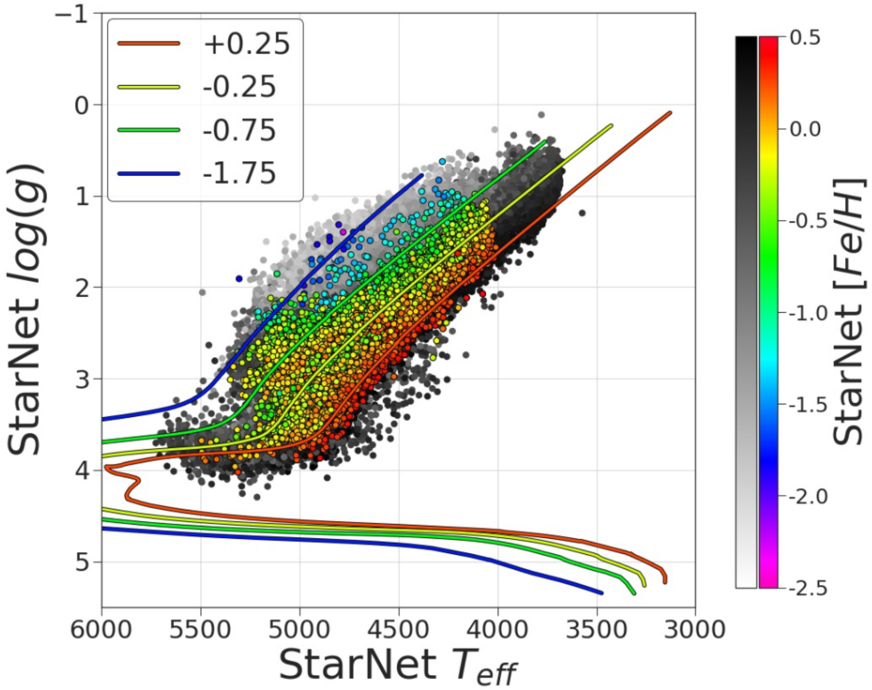
(smartish gridding)



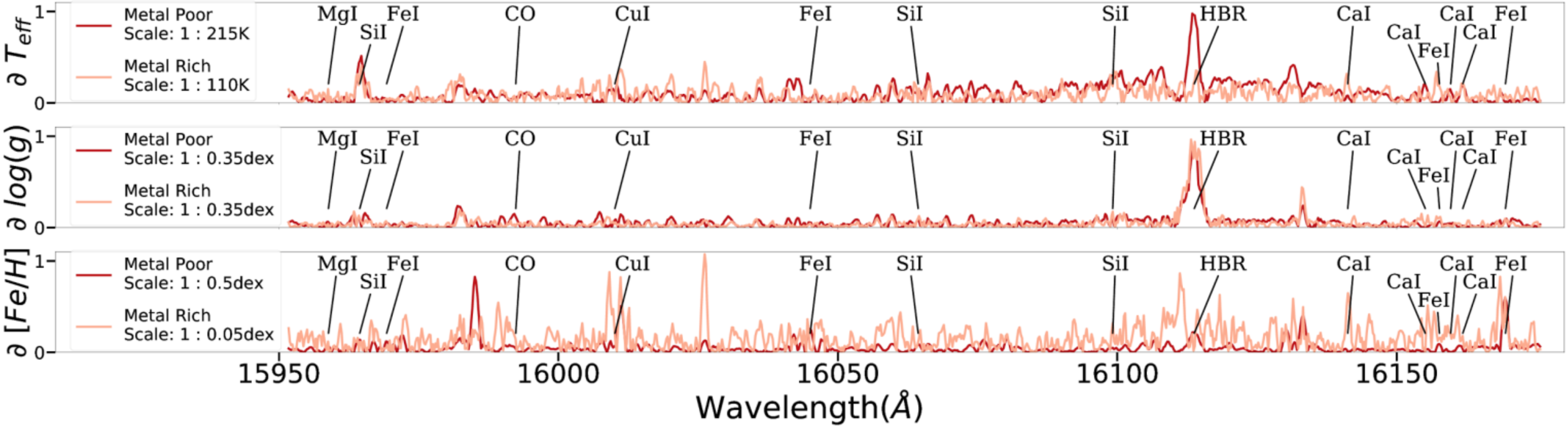
realistic parameters



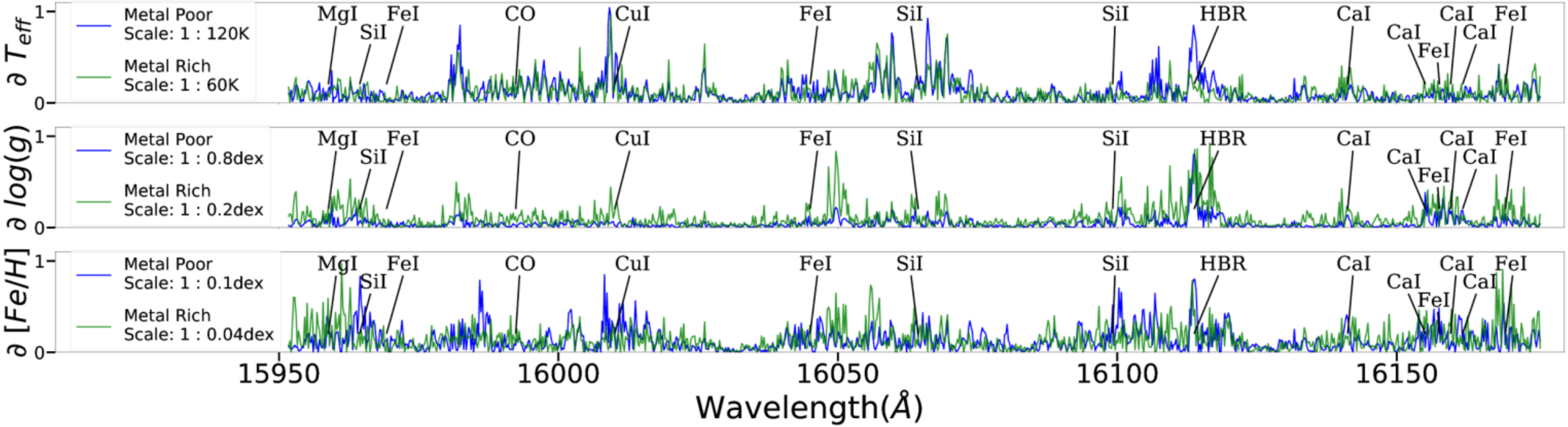
training set



where to look in a spectrum - synthetic

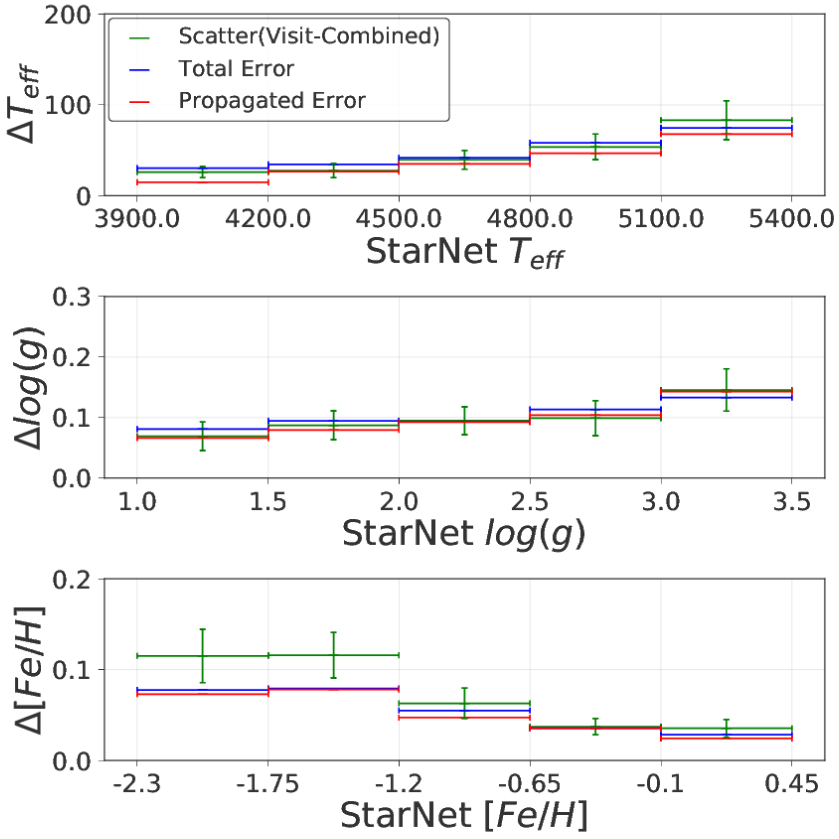


where to look in a spectrum - real

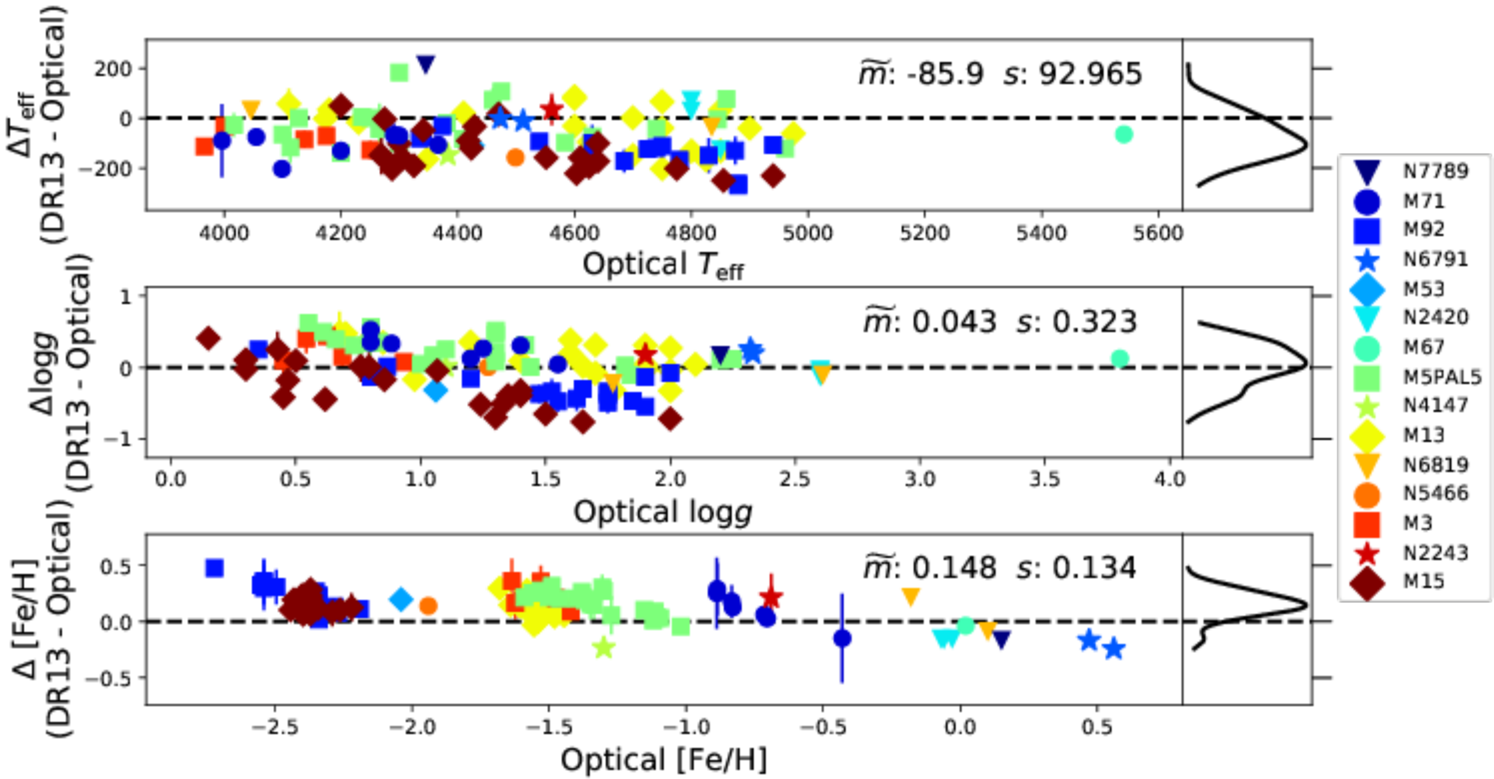


uncertainties

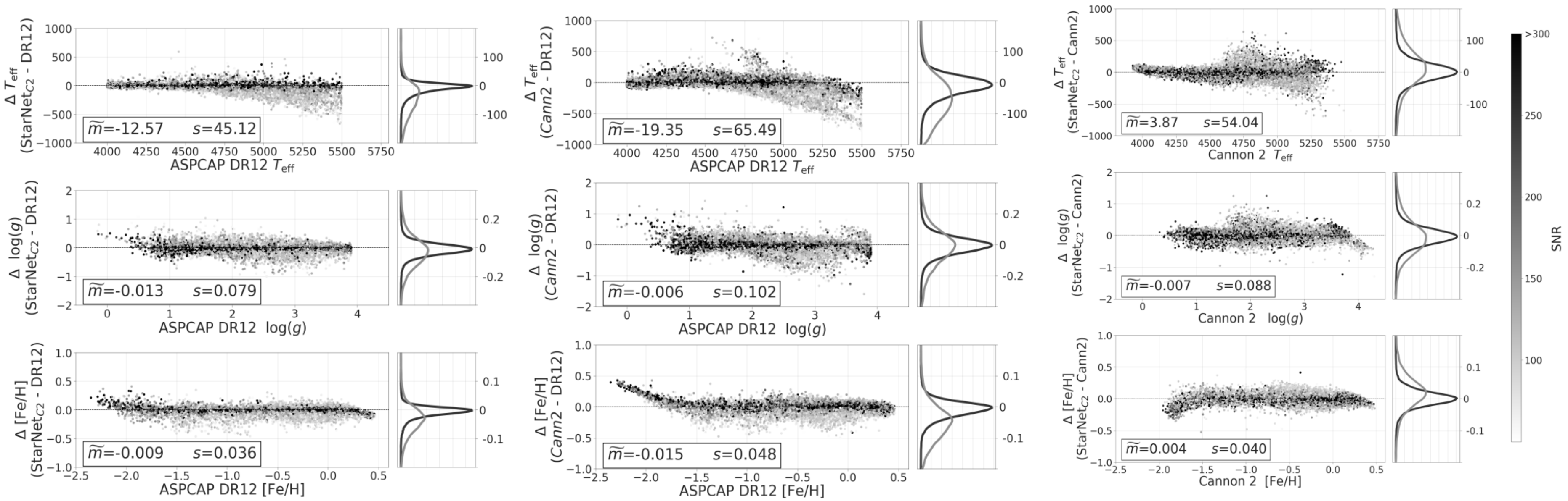
propagated from error spectra onto the starnet model with intrinsic scatter.



calibration uncertainties



comparison with official APOGEE methods



StarNet vs. ASCAP

The Cannon 2 vs. ASCAP

StarNet vs. The Cannon 2

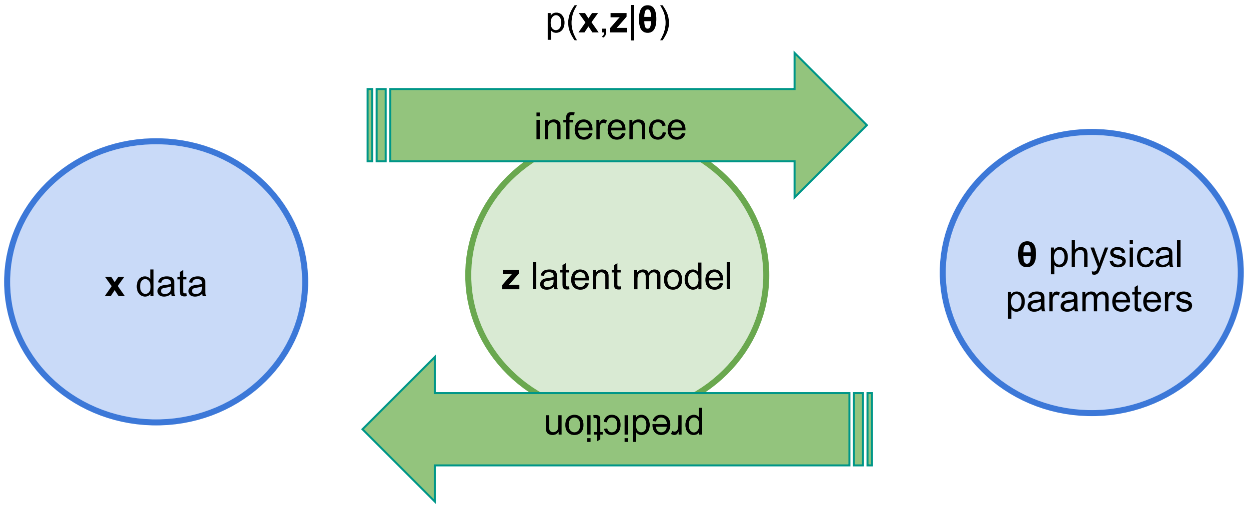
machine learning limitations

- ~~dealing with uncertainties~~
- ~~dealing with heteroscedastic data~~
- ~~dealing with missing d~~
- interpretable mod
- generalisability
- reduce training set size

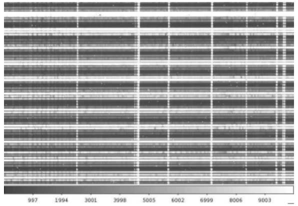
machine learning limitations

- dealing with uncertainties
- dealing with heteroscedastic data
- dealing with missing data
- interpretable models
- generalisability
- reduce training set size

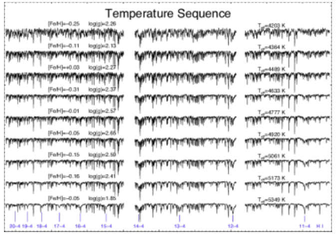
machine learning meets statistics



generative implicit models, deep probabilistic programming 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],...

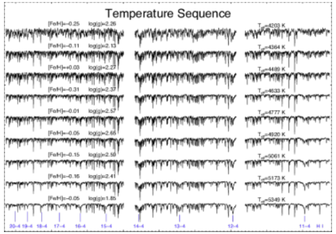
NTCO 2018?

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],...

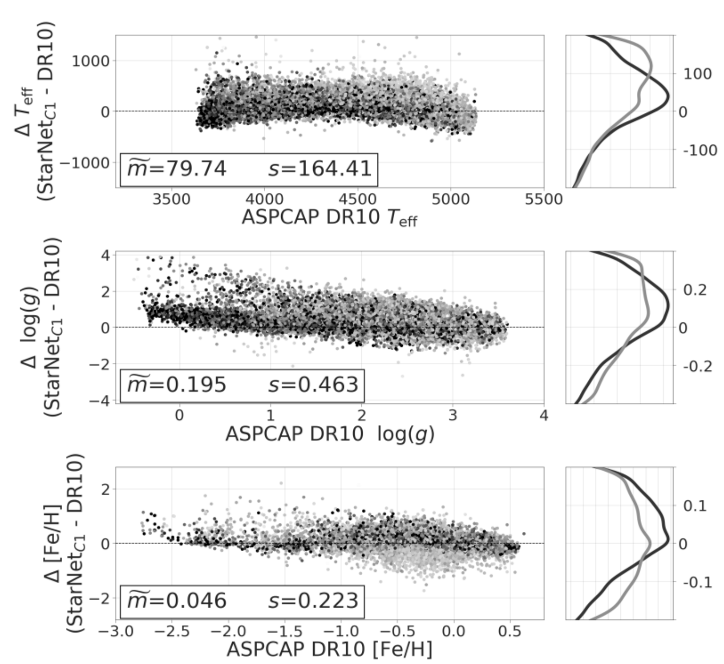
synthetic 1D spectra

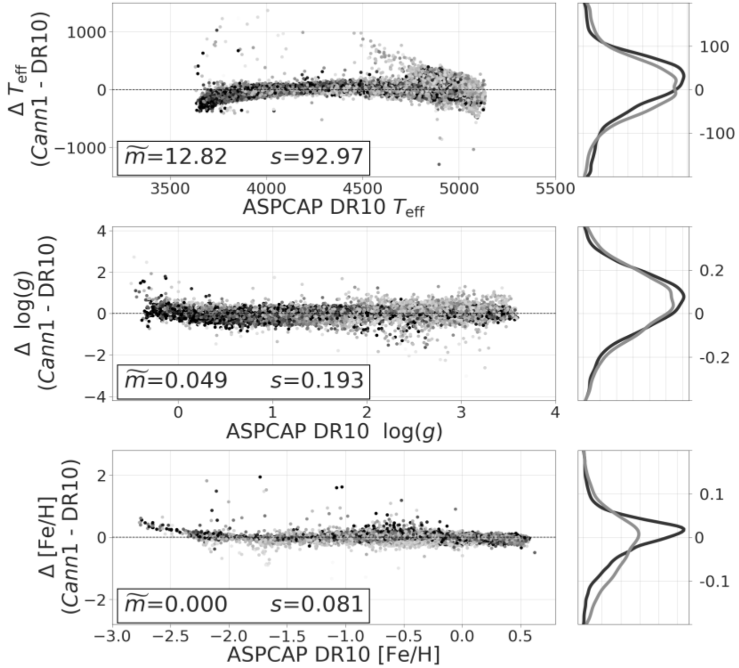


end

extra slides

small training set





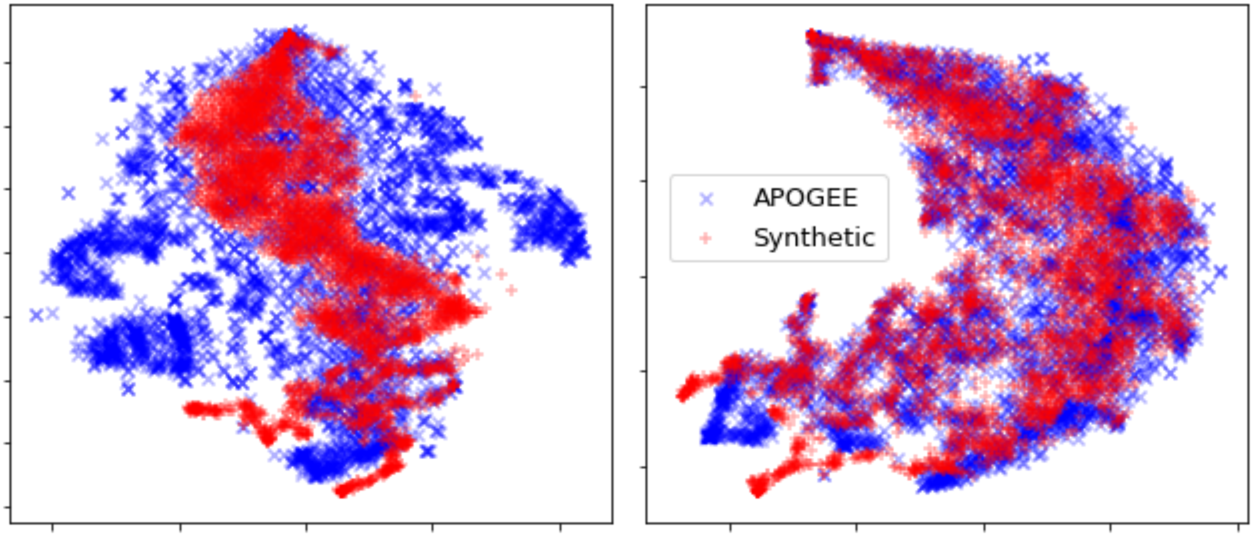
StarNet

The Cannon Ness (2015)

542 DR10 stars

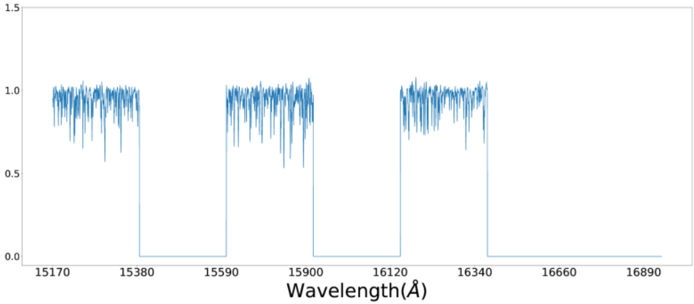
synthetic gap?

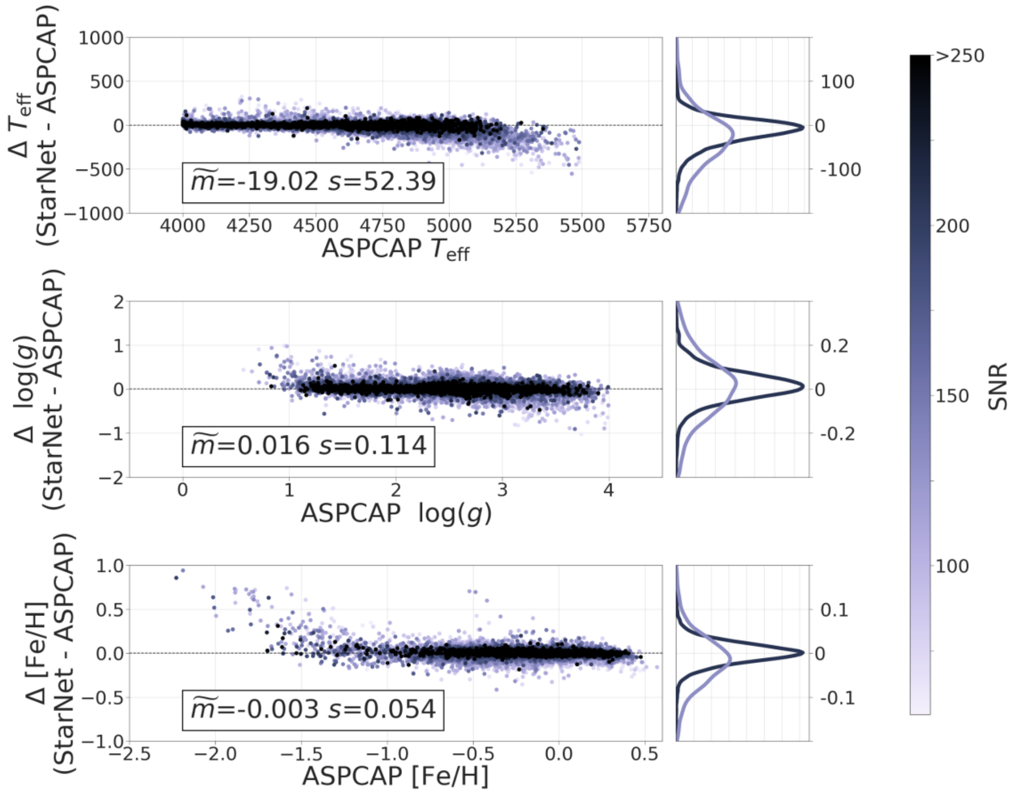
t-SNE before and after zero-data - interpolation



missing data?

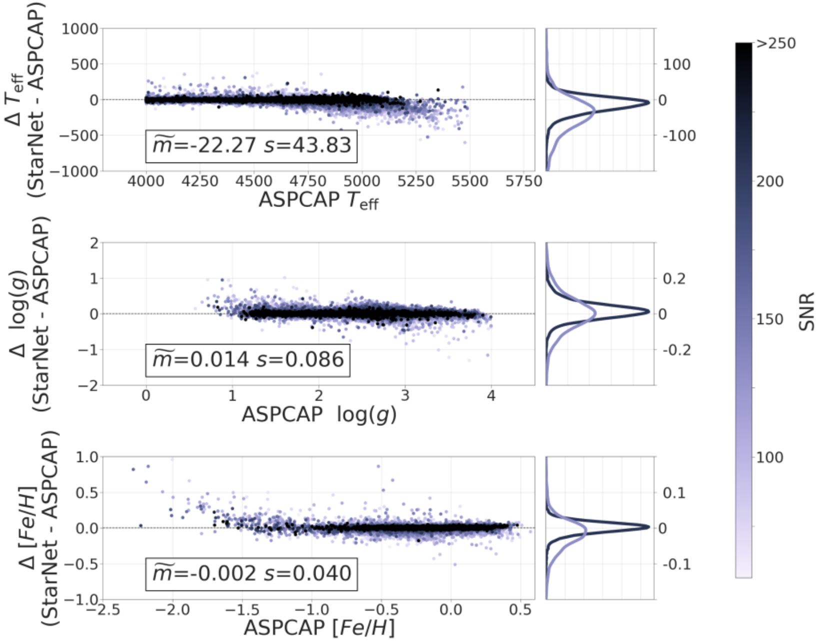
training and testing using random parts of the spectra



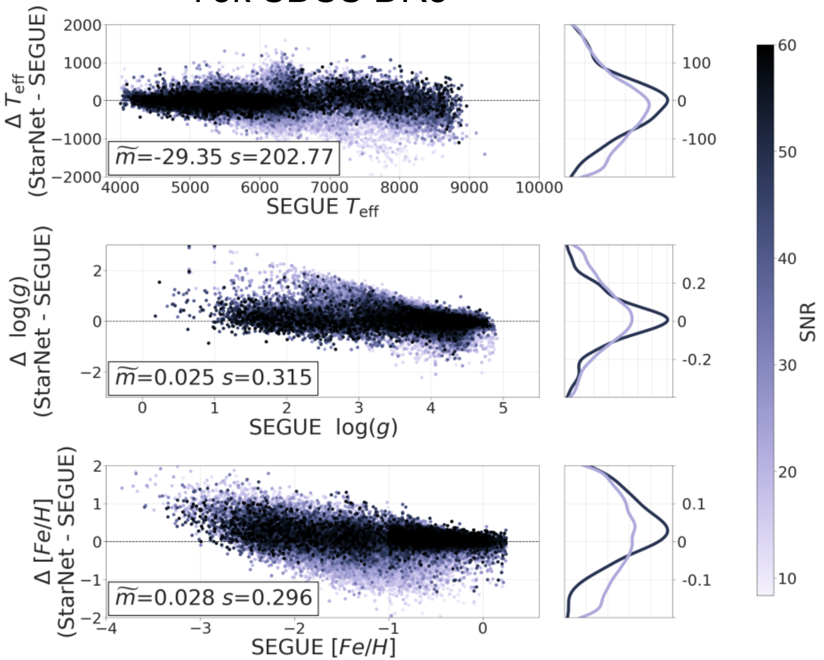


multi-surveys

21k APOGEE DR13



75k SDSS DR8



simultaneous training on 41k IR with APSCAP + 77k optical spectra with SEGUE

exploring deep learning architectures

Neural Network	no. of Filters in Conv. Layers	no. of Nodes in Fully Connected Layers	r^2	T_{eff}		$log(g)$		$[Fe/H]$	
				$\widetilde{x}$	σ	$\widetilde{x}$	σ	$\widetilde{x}$	σ
Shallow NN	0	512, 128	0.9667	-7.2	72.3	0.003	0.142	-0.008	0.057
Deep NN	0	2048, 1024, 512, 256, 128, 32	0.9642	-10.2	73.9	0.027	0.144	-0.011	0.057
Shallow CNN	16	128	0.9573	-1.2	82.9	0.017	0.155	-0.008	0.066
StarNet CNN	4, 16	256, 128	0.9749	-12.1	63.5	0.005	0.108	-0.014	0.049
Deep CNN	16, 32, 32, 64, 64	1024, 512, 256	0.9737	-8.7	70.2	0.003	0.105	-0.010	0.053