Mapping the Milky Way Galaxy with Deep Learning

Henry W. Leung

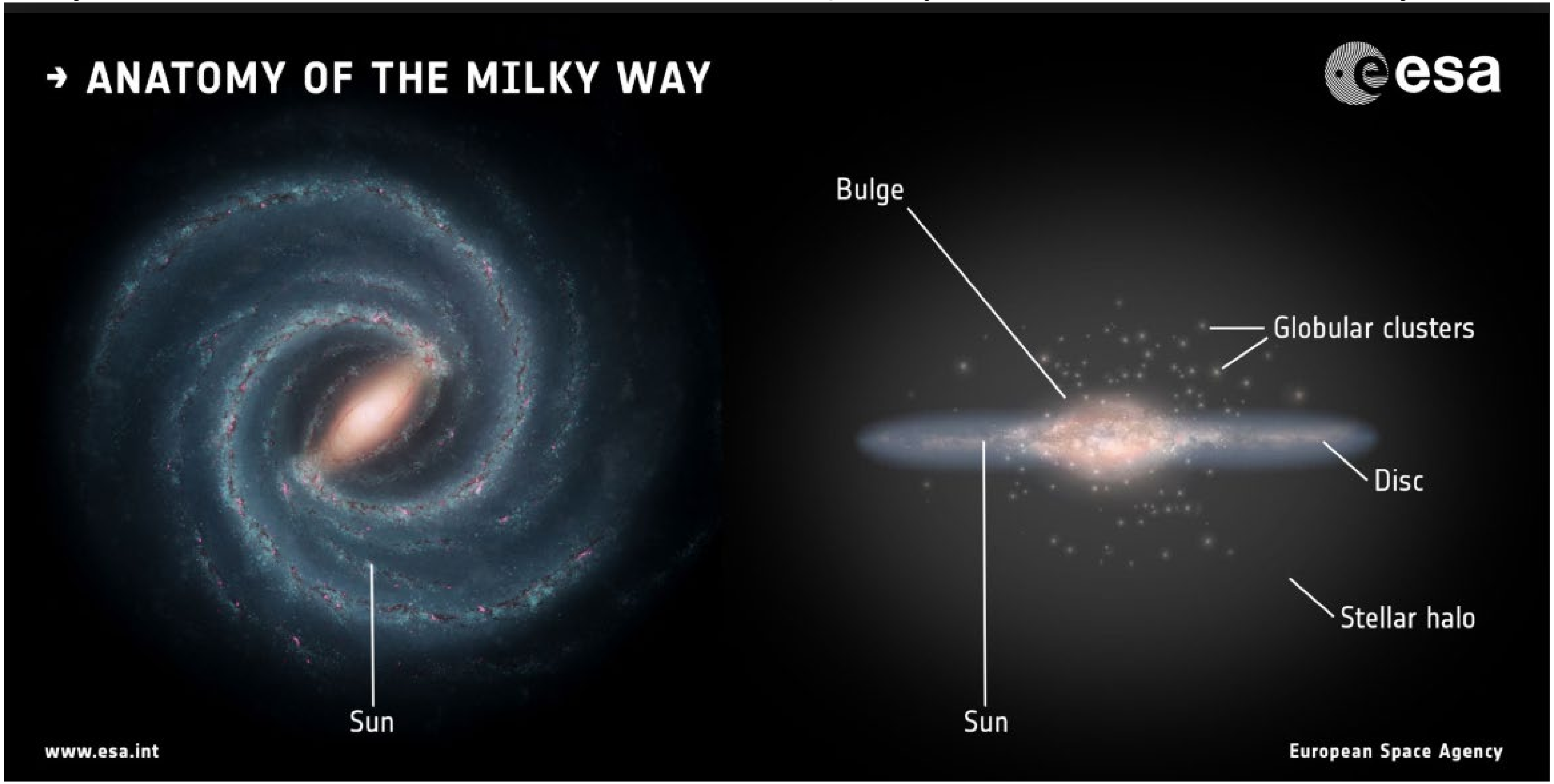
Astronomy & Astrophysics
UNIVERSITY OF TORONTO

Email: henrysky.leung@utoronto.ca

henrysky

Why mapping the Milky Way??

- The only one to be observed in detail (3D position/3D velocity/chemistry)



Chemical & dynamical evolution

Kinematics

- Disk Formation History
- Merger?
- Galaxies interaction

- By high precision astrometric survey

Chemistry

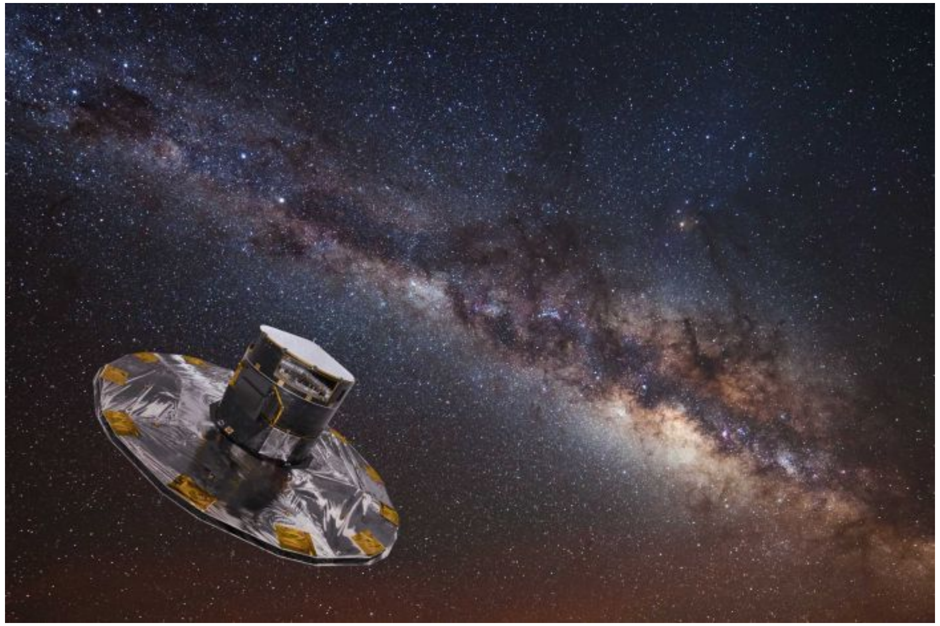
- Star Formation History
- Enrichment history

- By high resolution spectroscopic survey

Telescopes

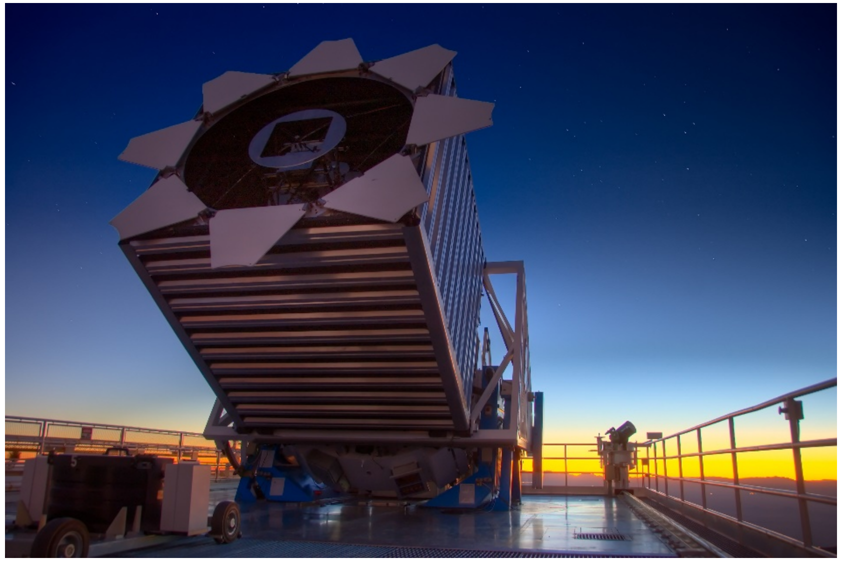
Kinematics

- ESA Gaia mission
- High precision astrometry

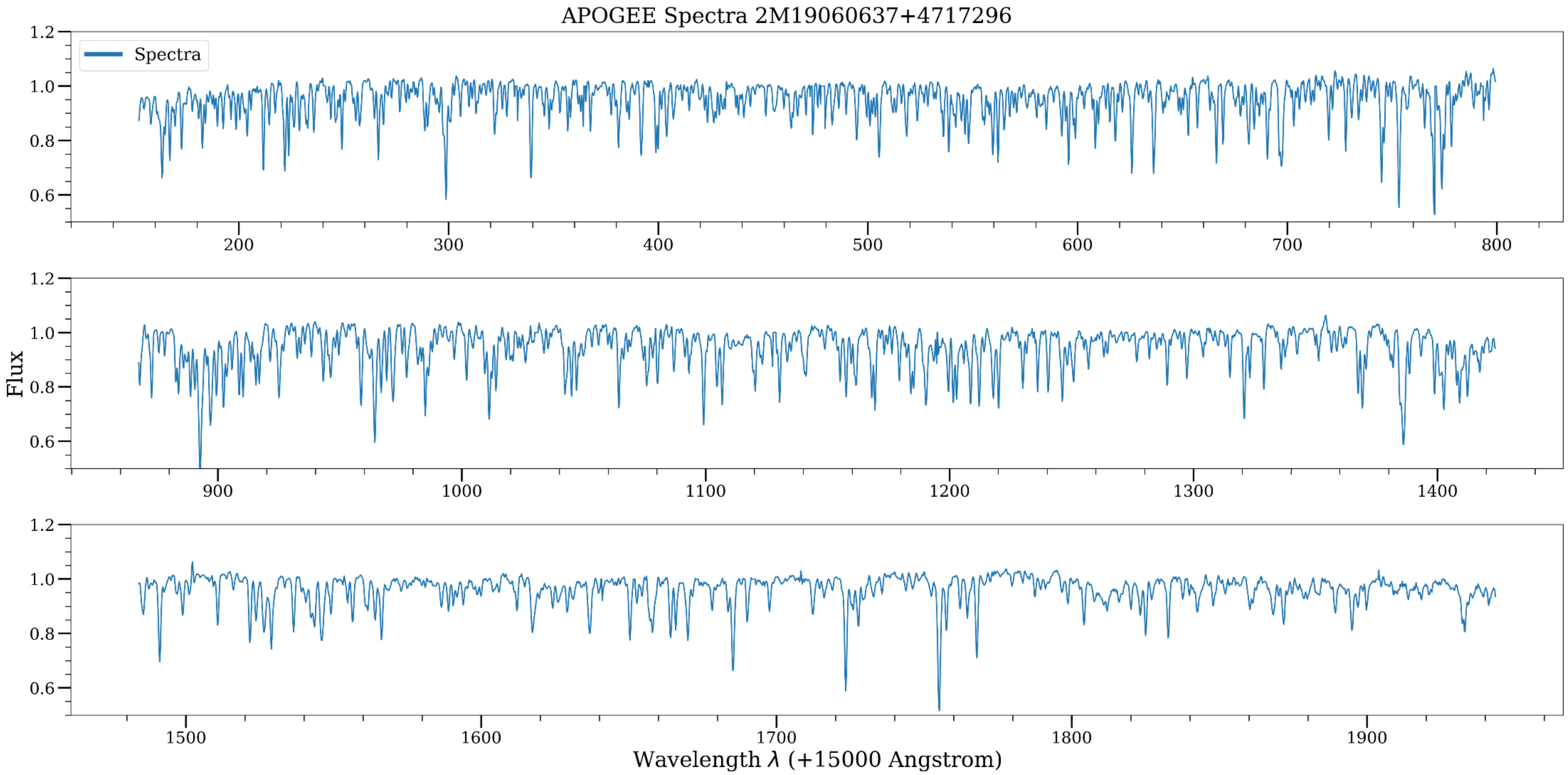


Chemistry

- SDSS APOGEE
- High-SNR, High resolution IR spectrographs



High-res IR spectroscopic data



- (Straightforward): Radial Velocity, Abundances
- (Not Straightforward): Mass?, Age?, Luminosity?, Extinction?
- (Not Possible): Proper motion, Position on the sky

Difficulties (Both science and method)

Kinematics

- Distances are bad at far away from parallax
- Instrumental bias

Chemistry

- Multiple surveys -> multiple pipeline
- Bad at low SNR
- Pipelines are slow

Machine learning method

- Uncertainties in training data
- Incomplete training data
- Uncertainties in prediction

My undergraduate neural net-related research with Prof. *Jo Bovy* (Toronto)

- Method + Science Papers:
  - Deep learning of multi-element abundances from high-resolution spectroscopic data
    [**Henry W. Leung**, Jo Bovy (2019a) arXiv:1808.04428]
  - Simultaneous calibration of spectro-photometric distances and the Gaia DR2 parallax zero-point offset with deep learning
    [**Henry W. Leung**, J. Bovy (2019b) arXiv:1902.08634]
- Science Papers using NNs data:
  - Dynamical heating across the Milky Way disc using APOGEE and Gaia
    [Ted J. Mackereth, Jo Bovy, **Henry W. Leung**, et al. (2019) arXiv:1901.04502]
  - Life in the fast lane: a direct view of the dynamics, formation, and evolution of the Milky Way's bar
    [Jo Bovy, **Henry W. Leung**, Jason A.S. Hunt, et al. (2019) arXiv:1905.11404]

Email: henrysky.leung@utoronto.ca henrysky Website: henrysky.github.io

astroNN (https://github.com/henrysky/astroNN)

- A python (≥3.6) package
- Well tested (86% code coverage)
- Well documented (https://astronn.rtfd.io/)
- Compatible with TensorFlow ≥ 1.13.2 & TensorFlow ≥ 2.0.0b0
- In active development (Both method & science)
- Easy to use & few handy functions
- Galaxy10 dataset (from SDSS/GalaxyZoo, counterpart of MNIST)
- Open Science (Software code & papers code)

Now go into the details of methods and sciences

Stellar parameters & Chemical abundances

- Develops a Python package called astroNN
- Bayesian NNs with Dropout VI (Y. Gal, et al. 2015) w/ modified loss function (**H. W. Leung**, J. Bovy 2019)
- Structures NNs to reflect physical knowledge
- Infer stellar parameters and chemical abundances precisely at low SNR, account for training data uncertainty & uncertainty in prediction

Bayesian NNs with Dropout Variation Inference

- Probabilistic weights with prior distribution
- Weights’ true PDF: VI by approximating vs MCMC by sampling
  - MCMC is impossible as we have millions to billions weights
- Dropout approx. true PDFs as a product of Bernoulli distributions
- Probably one of the easiest Bayesian NNs to be done that works

Modified loss function & uncertainties

- Taking training labels uncertainty into account along with *predictive uncertainty,* along with incomplete labels
- Prediction uncertainty = sum of predictive and model uncertainty

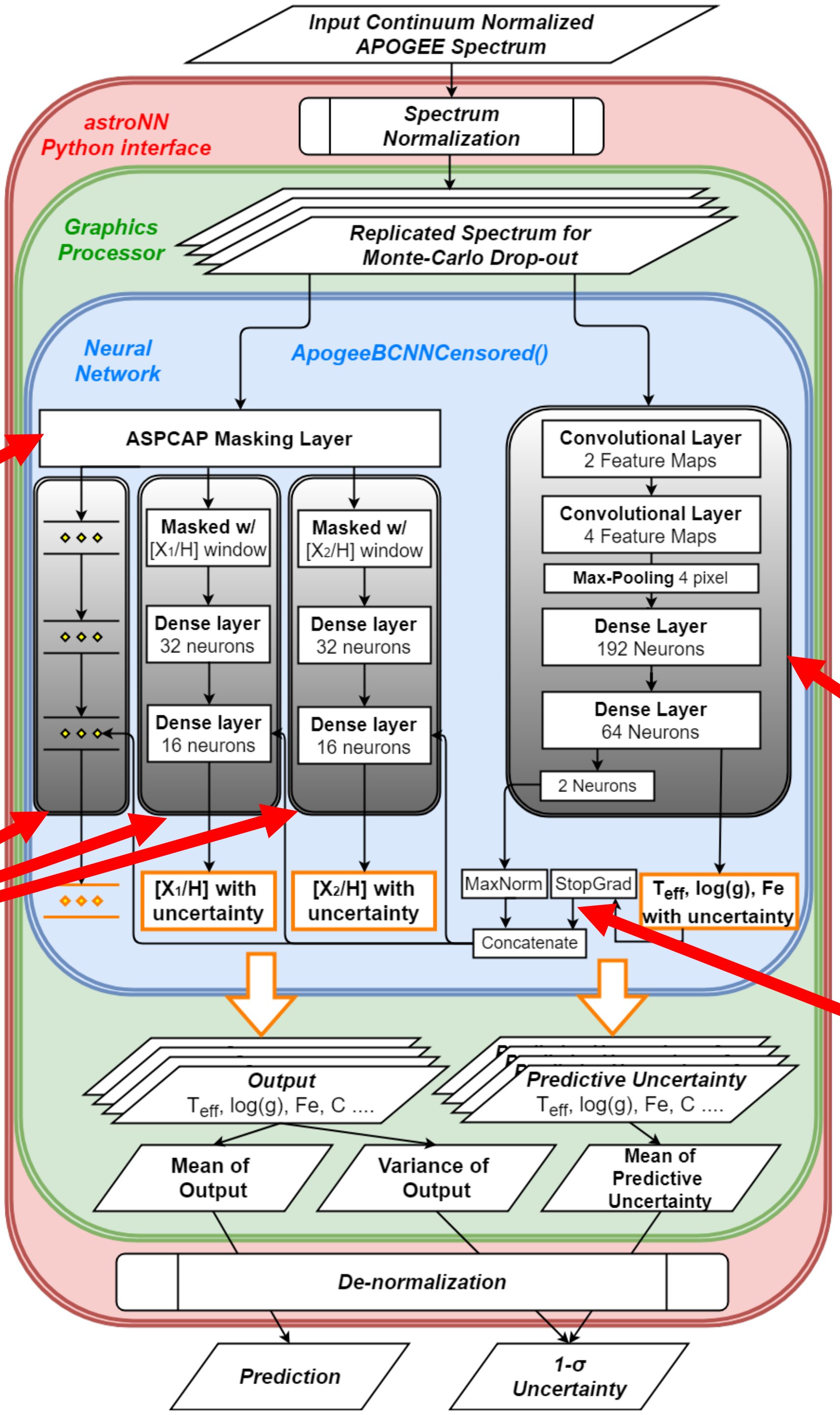
$$J(y_i, \hat{y}_i) = \begin{cases} \frac{1}{2}(\hat{y}_i - y_i)^2 e^{-s_i} + \frac{1}{2}(s_i) & \text{for } y_i \neq \texttt{MAGIC NUM} \\ 0 & \text{for } y_i = \texttt{MAGIC NUM} \end{cases}$$

,where $s_i = \ln\left[\sigma^2_{\text{known},i} + \sigma^2_{\text{predictive},i}\right]$

Magic Number for incomplete labels

Then $$J(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{D}\sum_{i=1}^{D} J(y_i, \hat{y}_i)\right)\mathcal{F}_{\text{correction},i}$$,where $\mathcal{F}_{\text{correction},i} = \frac{D}{D_i}$

Henry W. Leung, Jo Bovy 2019



Censoring on elements windows

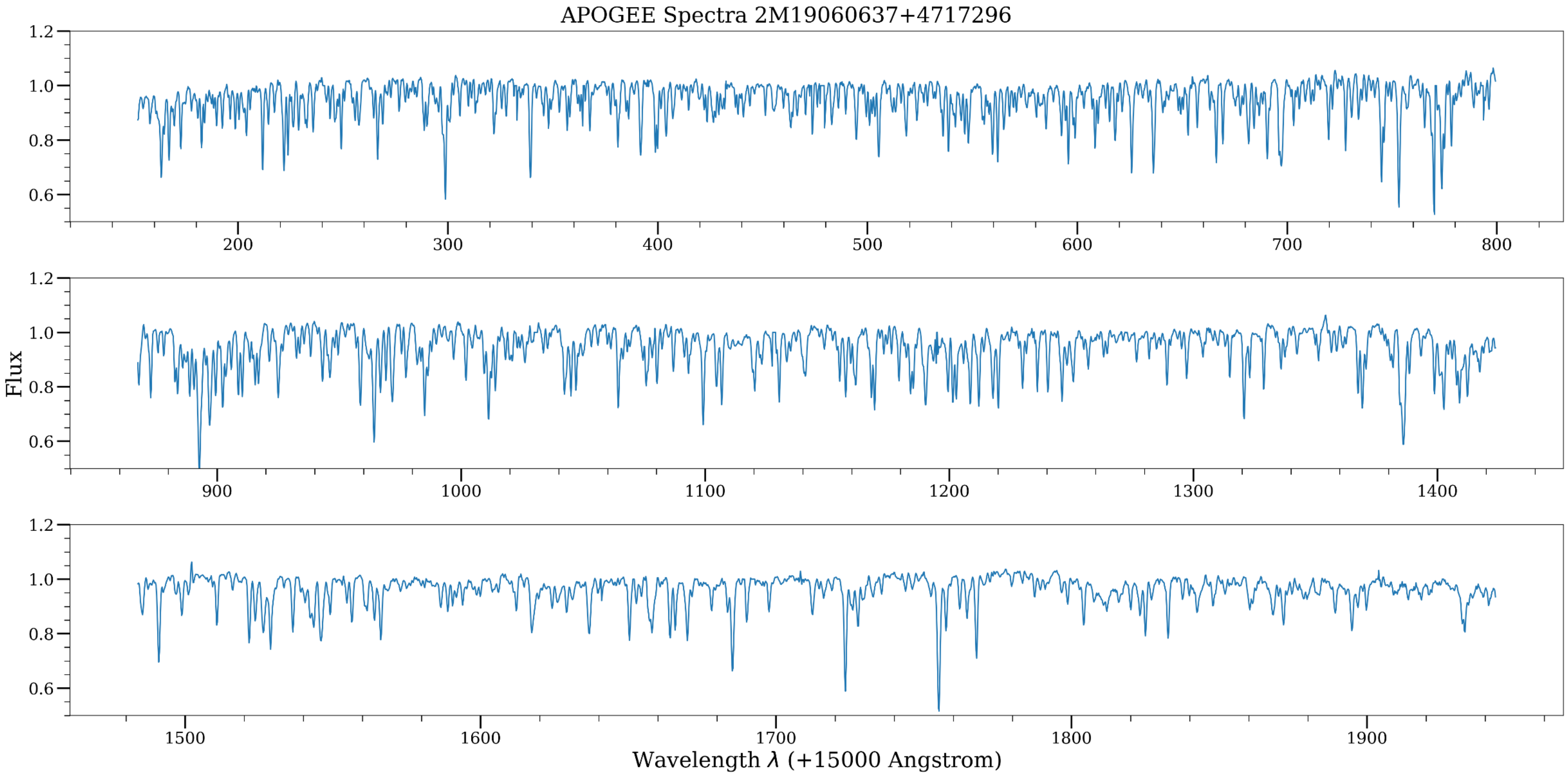
These small models produced elemental abundances prediction

NN model on whole spectra -> T_{eff}, log(g), Fe/H

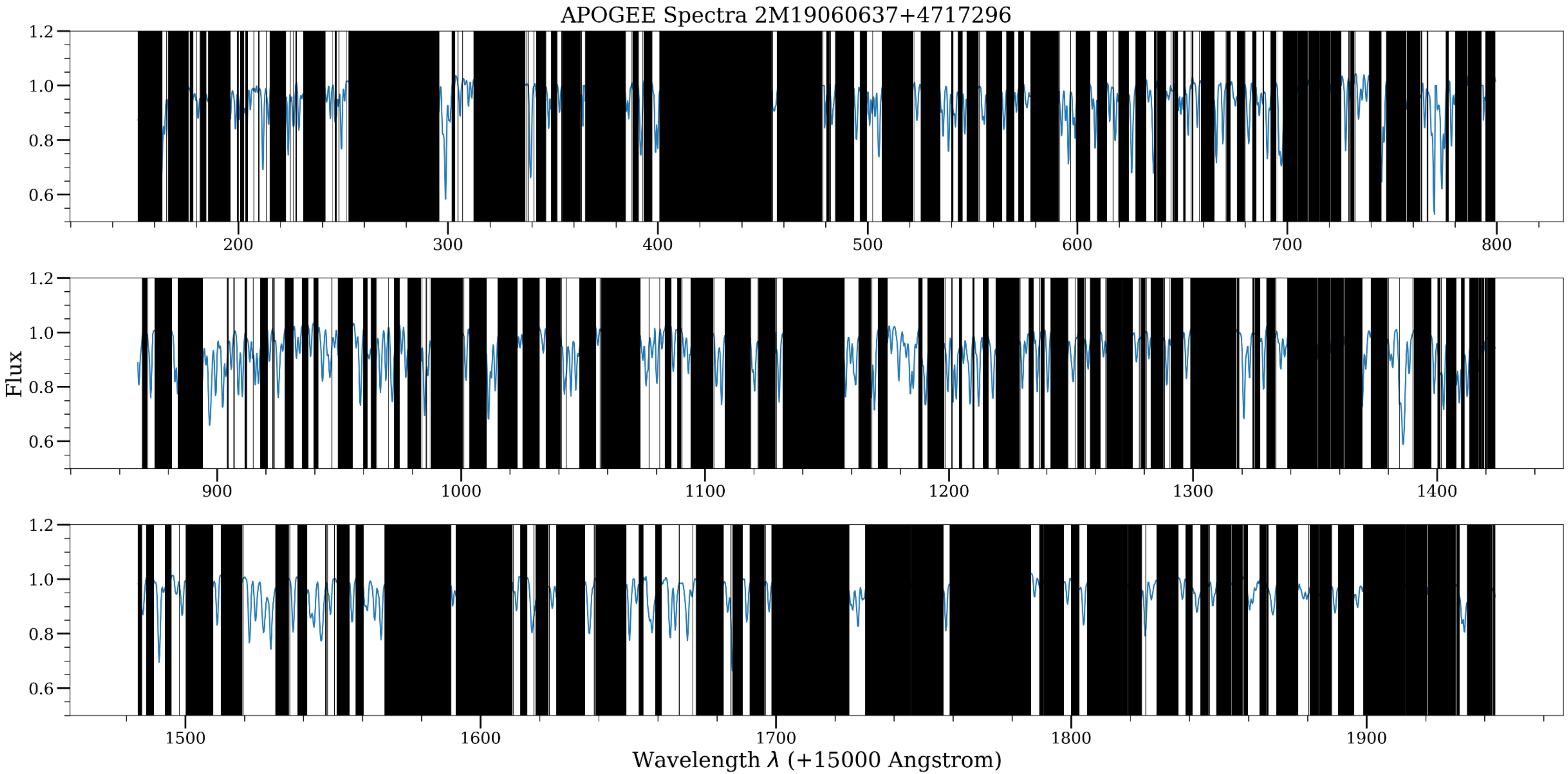
- Allow abundances dependent on T_{eff}, log(g), Fe/H but no gradient passes thru
- Allow limited attention to the whole spectra

Henry W. Leung, Jo Bovy 2019

An example of censoring with [Fe] windows

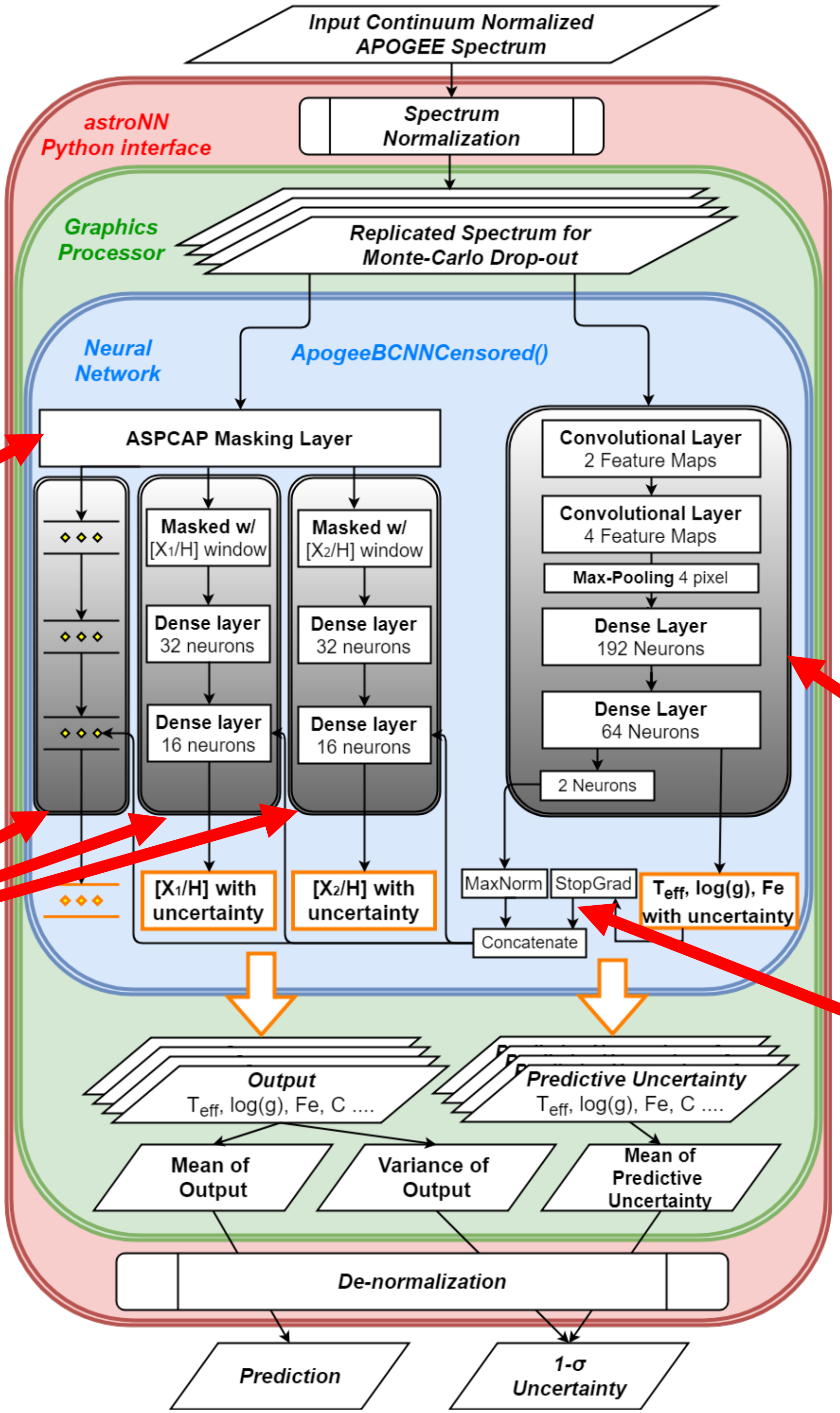


An example of censoring with [Fe] windows



Censoring on elements windows

These small models produced elemental abundances prediction



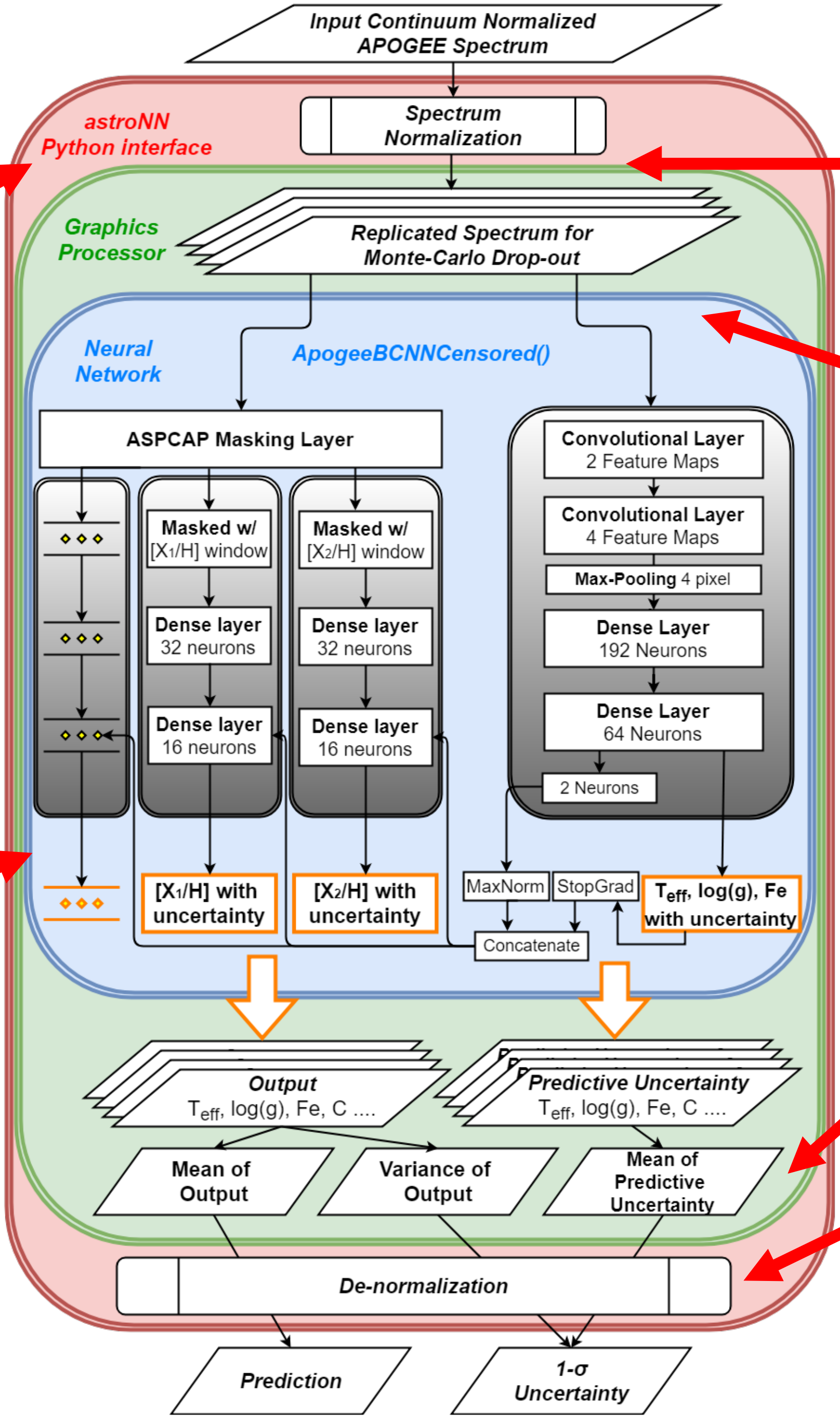
NN model on whole spectra -> T_{eff}, log(g), Fe/H

- Allow abundances dependent on T_{eff}, log(g), Fe/H but no gradient passes thru
- Allow limited attention to the whole spectra

Henry W. Leung, Jo Bovy 2019

Simple Python interface by astroNN

Model architectures are predefined as class in astroNN



Only one spectrum need for its MC inference, to minimize system to GPU bandwidth usage, no looping

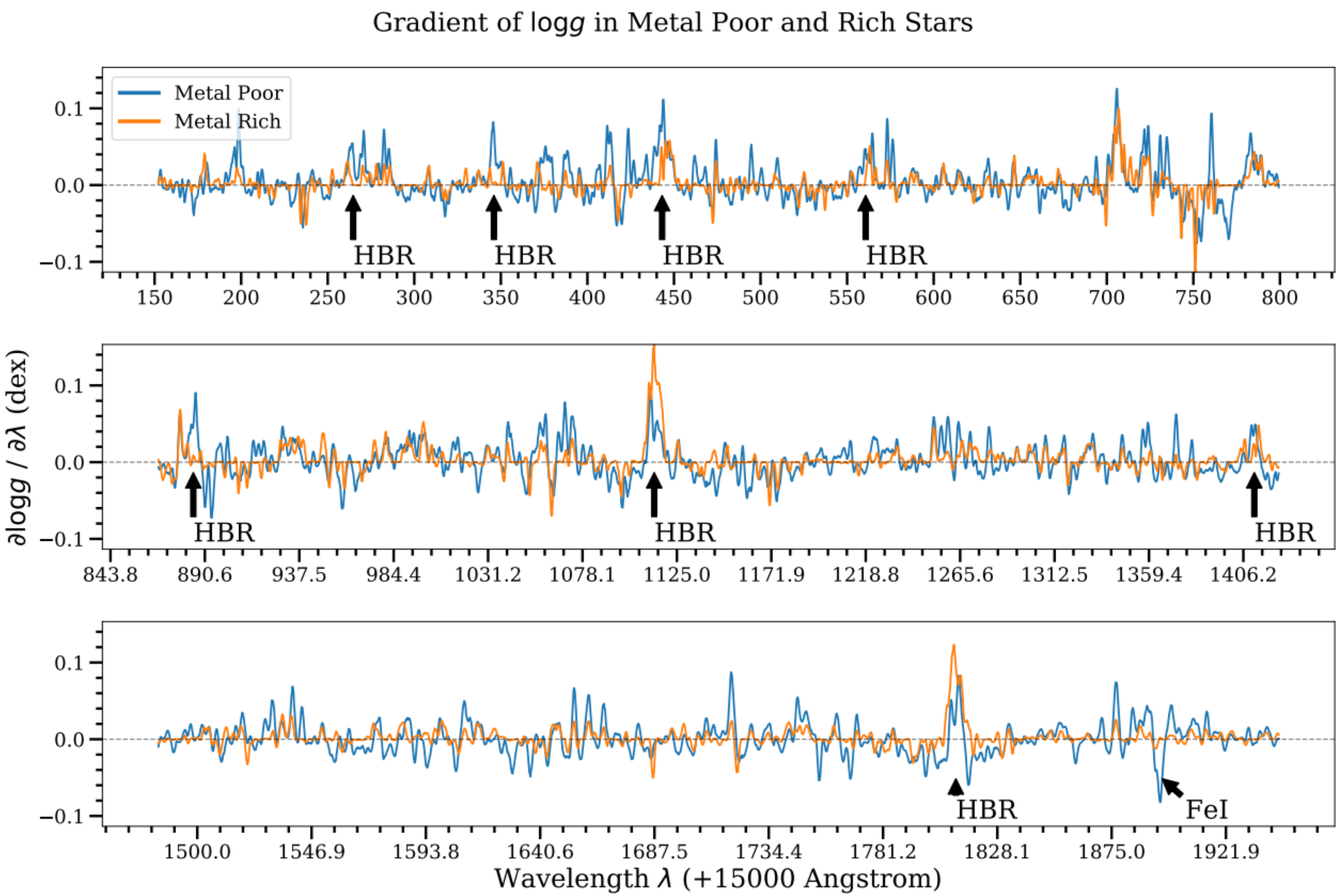
Fast Monte-Carlo inference on GPU, utilizing fast RNN code in TensorFlow to hack it

Monte-Carlo result processed on GPU and then sent back all at once

De-normalizing automatically with training values

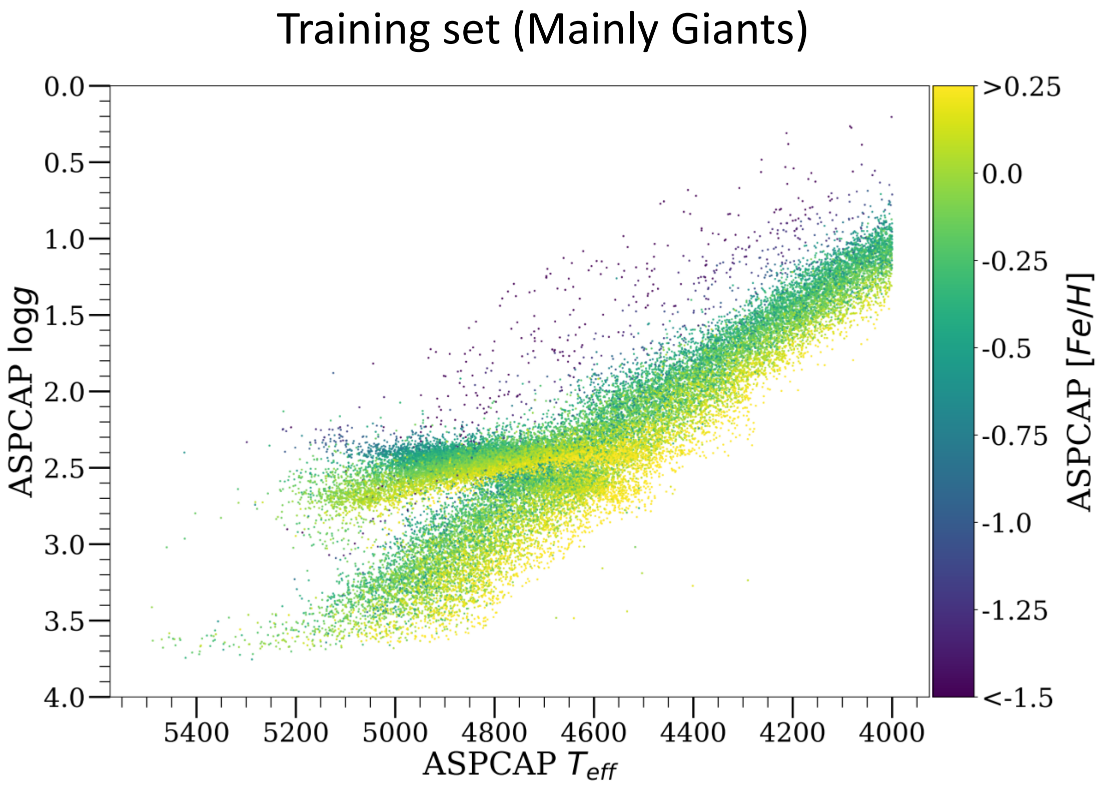
Henry W. Leung, Jo Bovy 2019

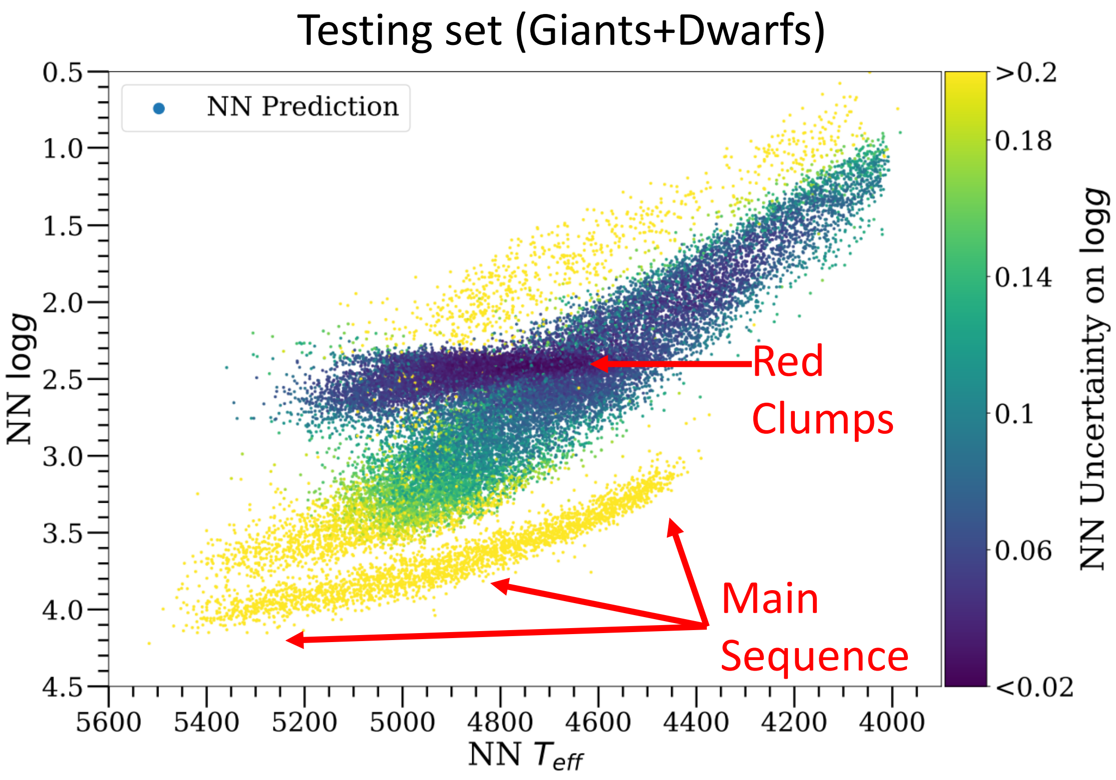
Where do the model pay attention to?



Henry W. Leung, Jo Bovy 2019

Stellar parameters & Chemical abundances

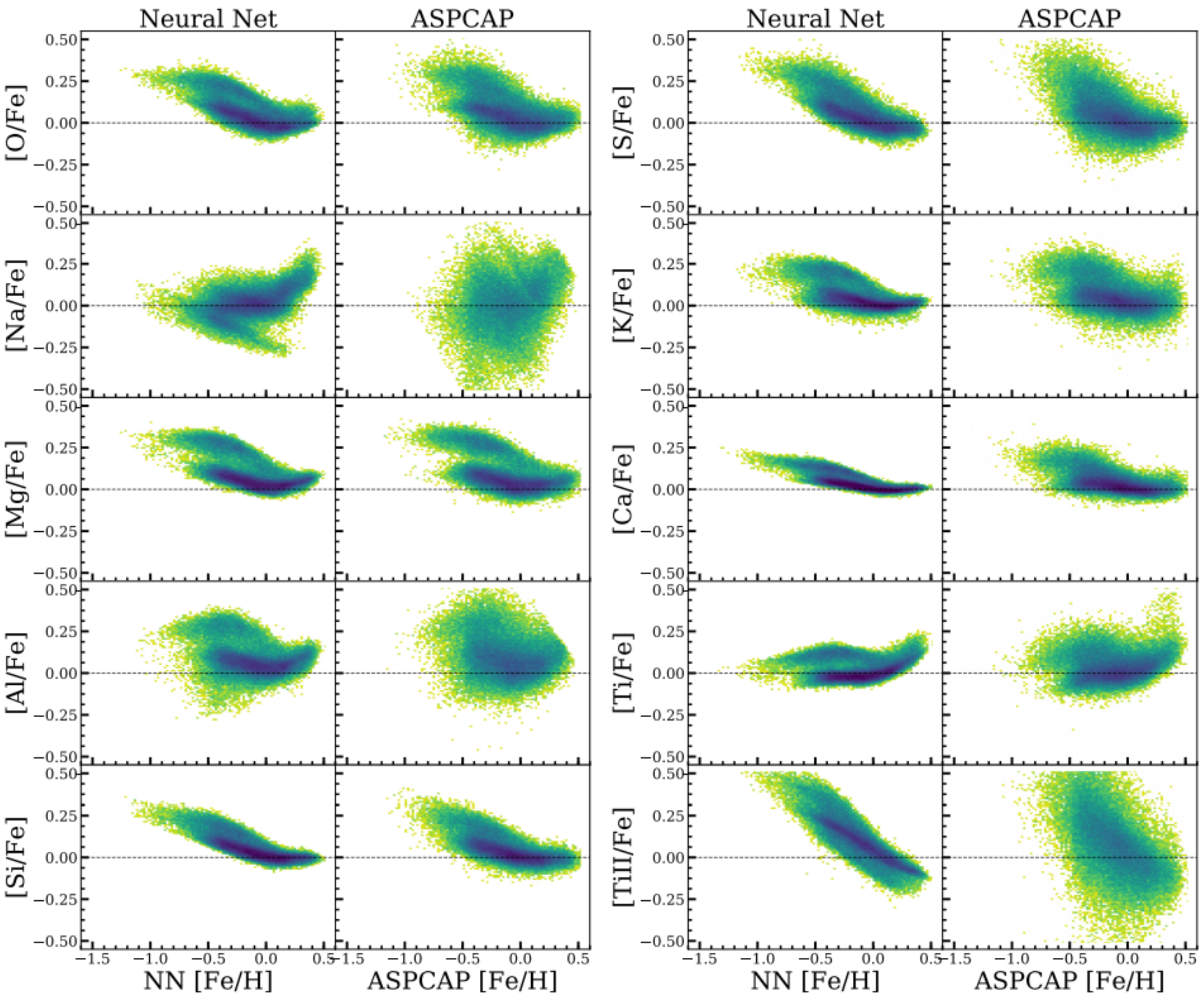




Henry W. Leung, Jo Bovy 2019

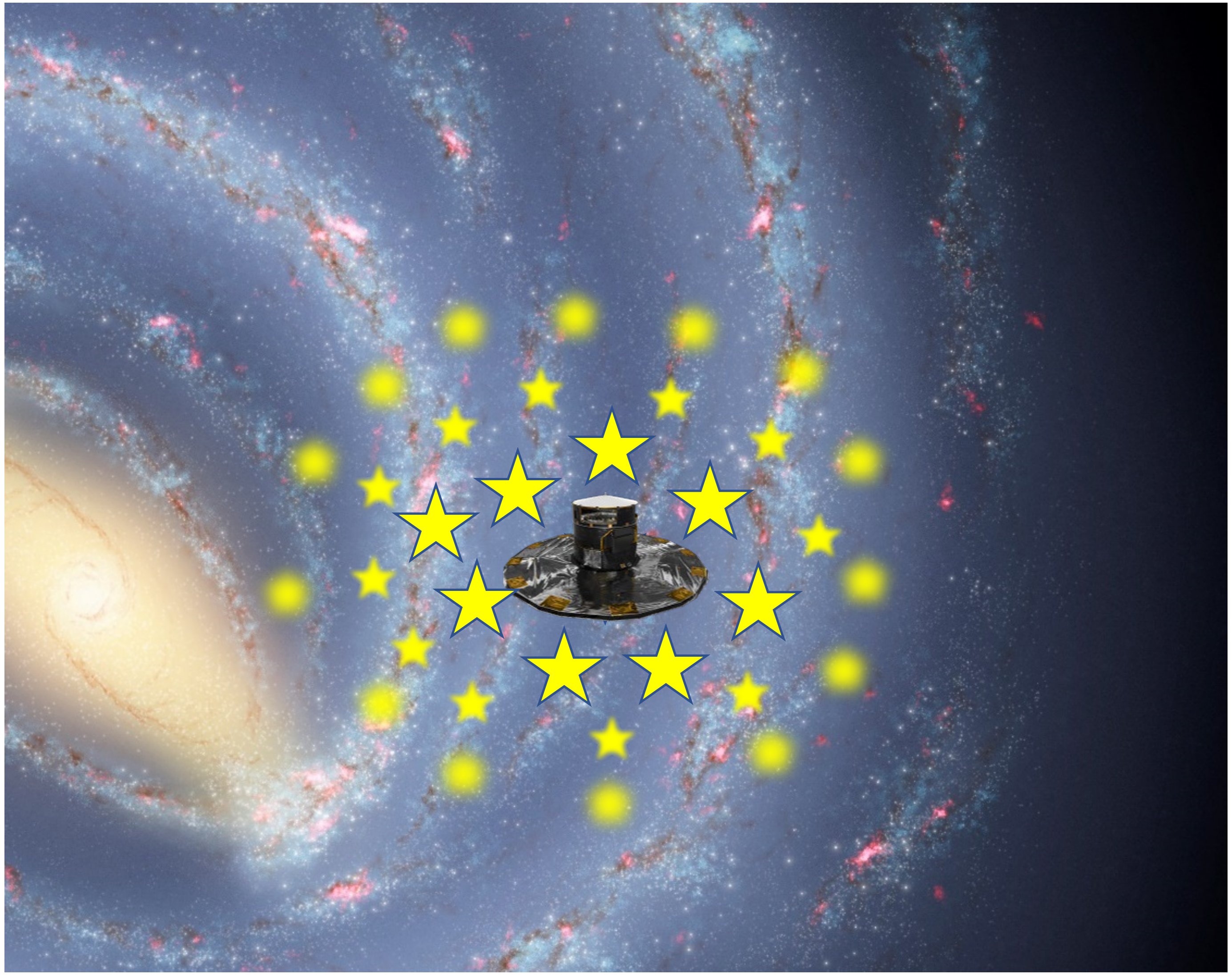
Stellar parameters & Chemical abundances

- Work well for low SNR
- Reasonable uncertainties
- Fast inference
  - 22 parameters for 10 millions 7514 pixels spectra in 300 seconds on GTX1060 6GB
- For small dataset too!
  - Approx. 5000 training spectra with reasonable result

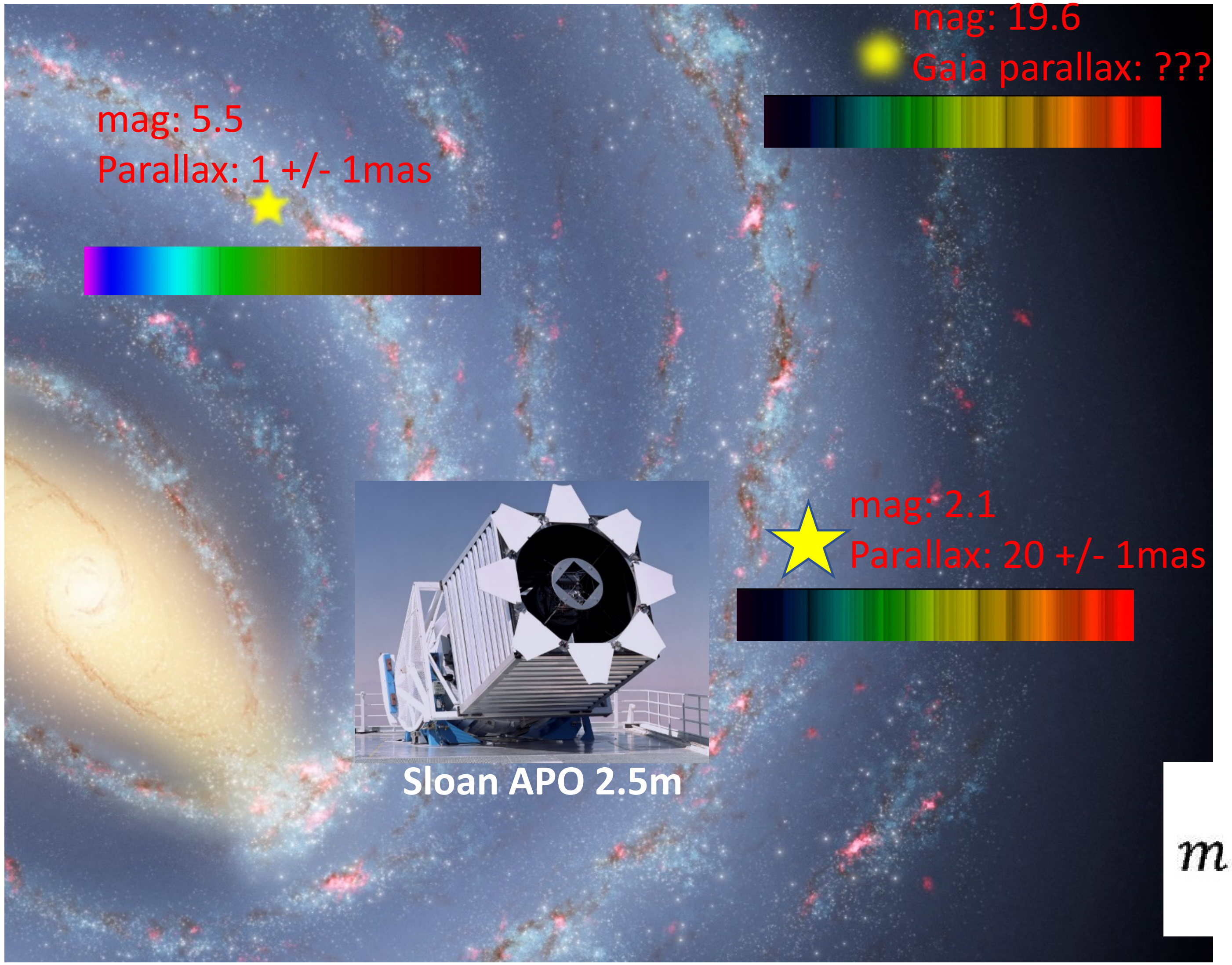


Henry W. Leung, Jo Bovy 2019

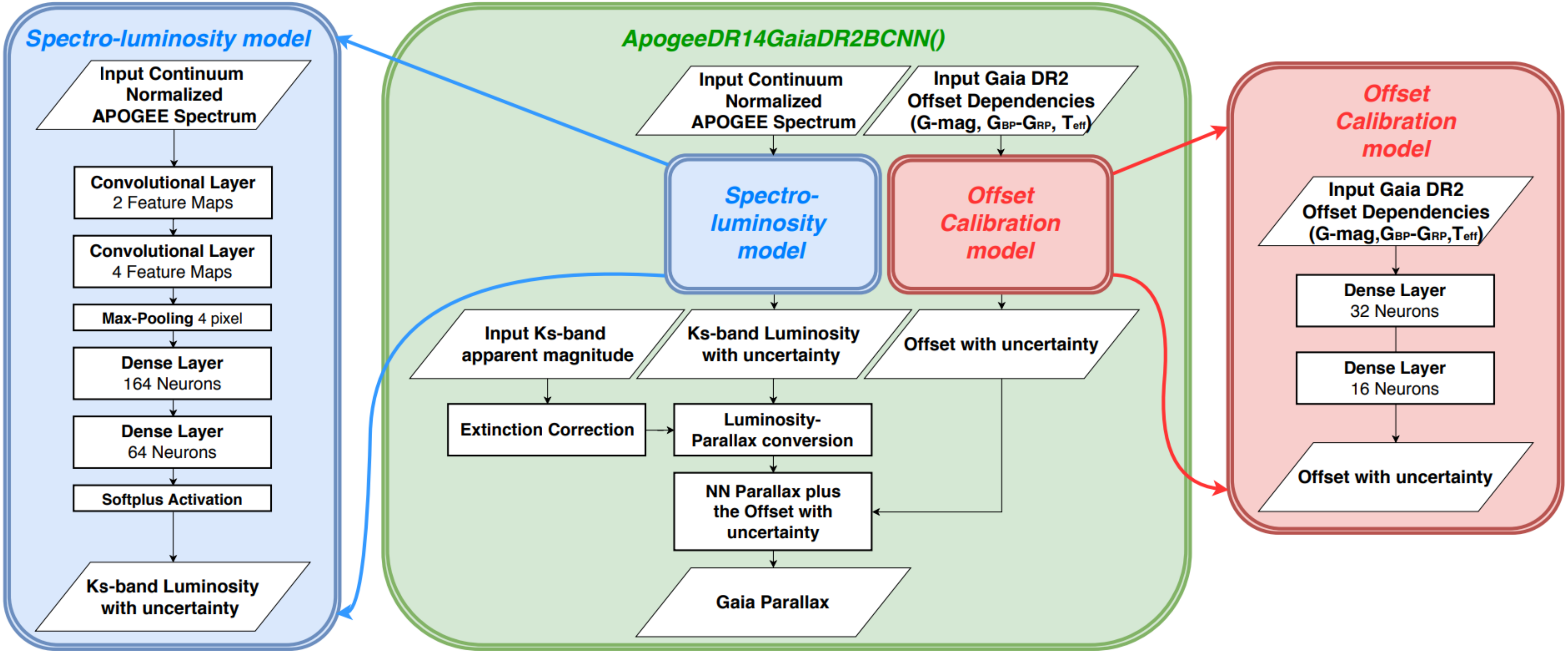
Spectro-photometric Distances to stars



Spectro-photometric Distances to stars

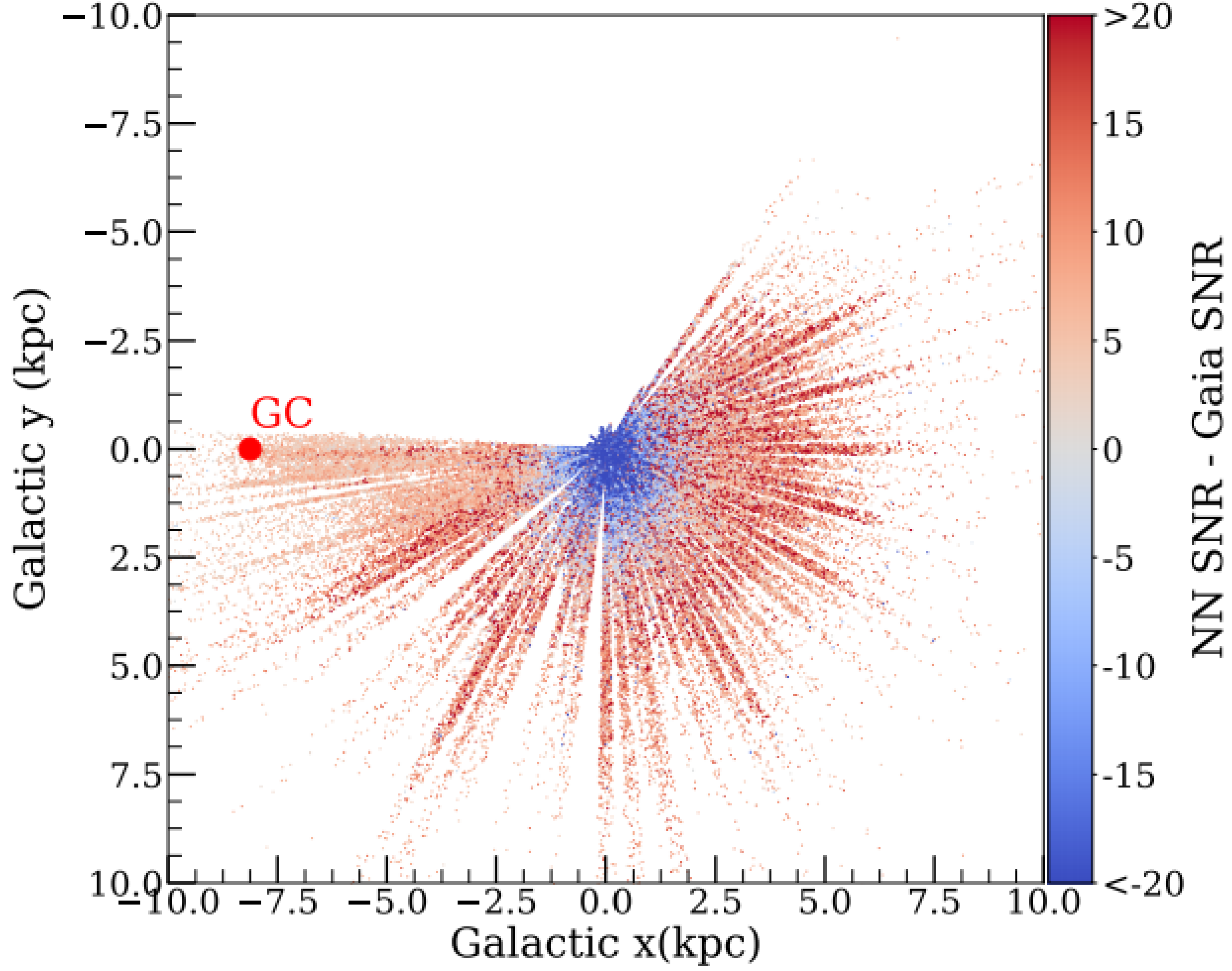


$$m - M = 5\log\left(\frac{d}{10}\right)$$

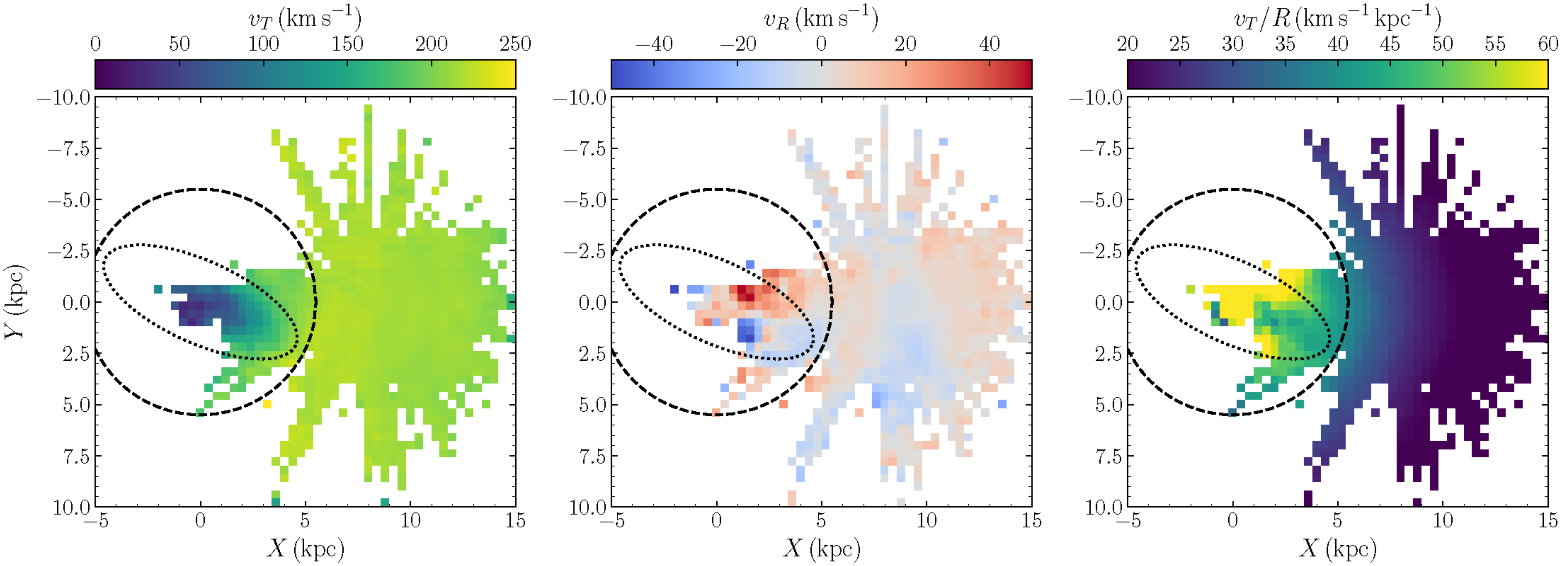


Henry W. Leung, Jo Bovy 2019

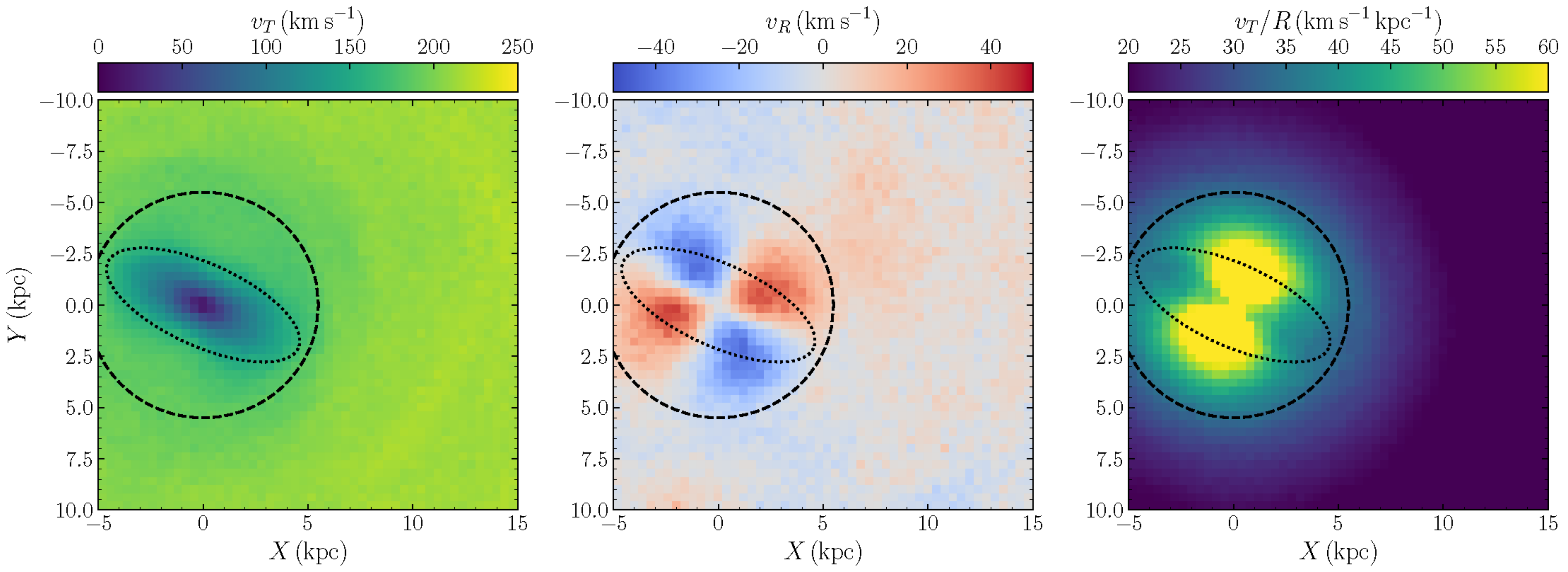
Signal-to-Noise comparison to Gaia DR2 using APOGEE DR14 spectra



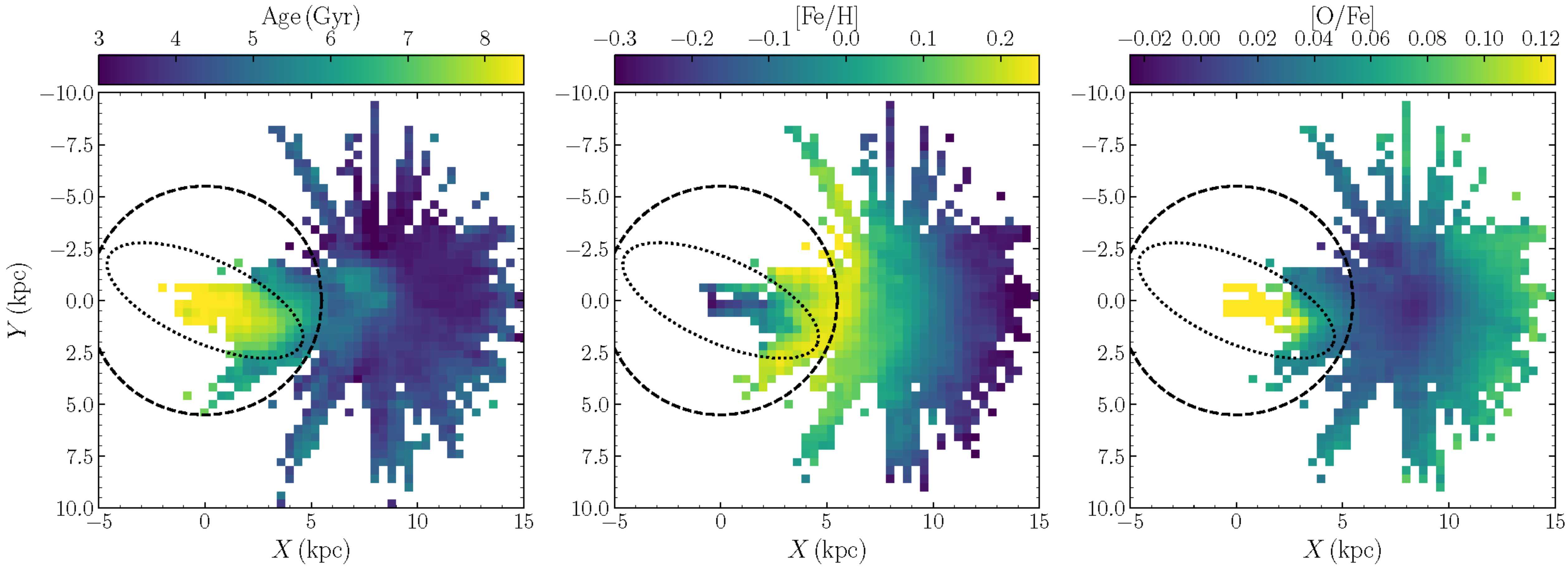
Henry W. Leung, Jo Bovy 2019



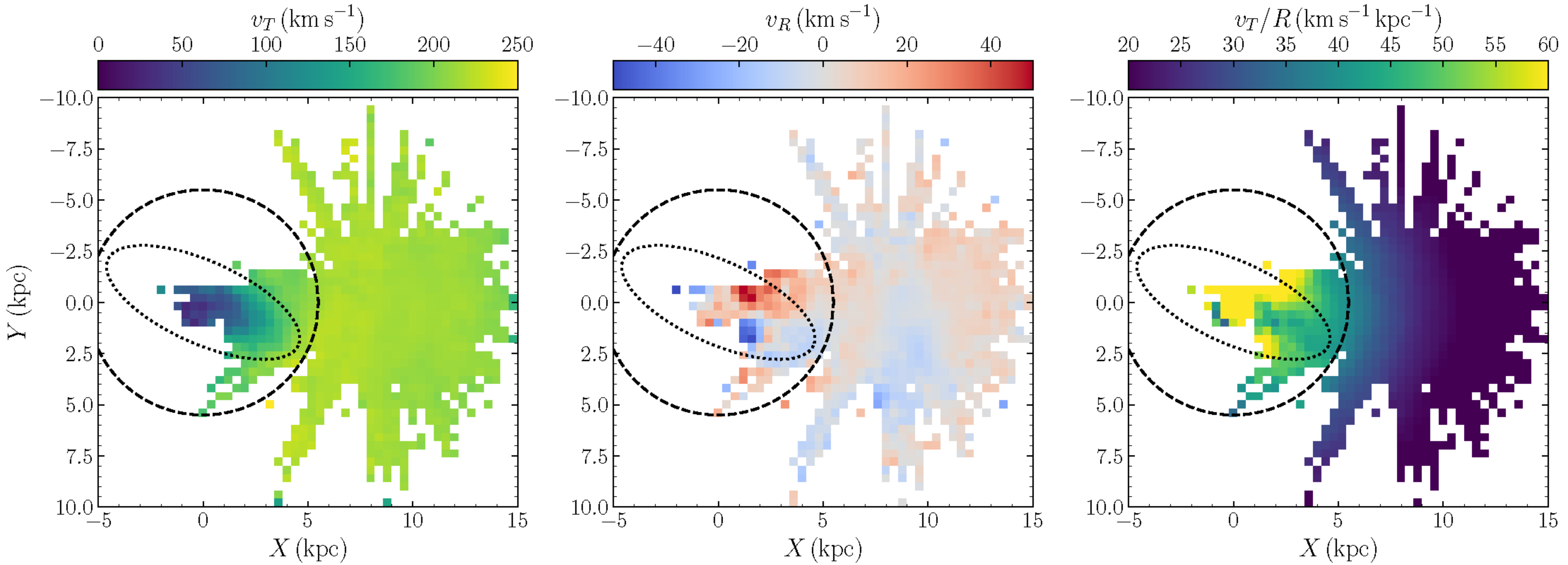
Jo Bovy, **Henry W. Leung**, et al 2019



Jo Bovy, **Henry W. Leung**, et al 2019



Jo Bovy, **Henry W. Leung**, et al 2019



Bar length: ~5kpc

Bar pattern speed: ~41 km/s/kpc

Bar age: ~8 Gyr

Jo Bovy, **Henry W. Leung**, et al 2019

Summary

- Bayesian NNs with Dropout VI w/ modified loss function
  - Account for Uncertainty/Incompleteness in training data, also produce uncertainty in prediction
- Structures NNs to reflect physical knowledge
- Develops a python package – astroNN (https://github.com/henrysky/astroNN/)

- Data from NNs to map milky way (chemical/kinematics)
  - Real world scientific progress!
  - Formation, evolution and structure of milky-way
  - Data product available as SDSS Value-Added Catalog (DR16; Dec 2019)

Email: henrysky.leung@utoronto.ca henrysky Website: henrysky.github.io