Masses (and implied ages) from Stellar Spectra: Using *The Cannon* to exploit the *Kepler*-APOGEE overlap

M. Ness, D. W. Hogg, H.-W. Rix, M. Martig, A. Y. Q. Ho, *in prep*

Anna Ho (Caltech)
Synoptic Surveys: Boutique & Experiments
29 August 2015

*Named after Annie Jump Cannon, who classified stars without using stellar models.*
Summary

• *The Cannon* is a data-driven method for modeling stellar spectra as a function of stellar labels*

*parameters & abundances, collectively*
Summary

- *The Cannon* is a data-driven method for modeling stellar spectra as a function of stellar labels*
- APOKASC is a catalog of 2,000 stars measured in common between APOGEE and *Kepler*.

*parameters & abundances, collectively
Summary

• *The Cannon* is a data-driven method for modeling stellar spectra as a function of stellar labels*

• APOKASC is a catalog of 2,000 stars measured in common between APOGEE and *Kepler*.

• *The Cannon* “learns” from APOKASC to model APOGEE spectra as a function of 5 labels, incl. **mass**

*parameters & abundances, collectively*
Summary

• *The Cannon* is a data-driven method for modeling stellar spectra as a function of stellar labels*
• APOKASC is a catalog of 2,000 stars measured in common between APOGEE and *Kepler.*
• *The Cannon* “learns” from APOKASC to model APOGEE spectra as a function of 5 labels, incl. mass
• **Result:** our (data-driven) model can determine the mass of a giant star directly from its spectrum.

*parameters & abundances, collectively
We want to map stellar ages throughout the MW.
We want to map stellar ages throughout the MW.

- APOGEE provides an extensive sample of giant stars
We want to map stellar ages throughout the MW.

- APOGEE provides an extensive sample of giant stars

Apache Point Observatory Galactic Evolution Experiment
- SDSS near-IR (1.52-1.69 μ) stellar spectroscopic survey of MW disk, bulge, halo
- R ~ 22,500, S/N ~ 100
- 300 fibers
- 150,000 giants observed as of DR12
We want to map stellar ages throughout the MW.

- APOGEE provides an extensive sample of giant stars

Apache Point Observatory Galactic Evolution Experiment

- SDSS near-IR (1.52-1.69 μ) stellar spectroscopic survey of MW disk, bulge, halo
- R ~ 22,500, S/N ~ 100
- 300 fibers
- 150,000 giants observed as of DR12

- For giant stars, mass is a powerful constraint on age (e.g. Martig et al 2014)
We want to map stellar ages throughout the MW.

- APOGEE provides an extensive sample of giant stars

**Apache Point Observatory Galactic Evolution Experiment**
- SDSS near-IR (1.52-1.69 μ) stellar spectroscopic survey of MW disk, bulge, halo
- R ~ 22,500, S/N ~ 100
- 300 fibers
- 150,000 giants observed as of DR12

- For giant stars, mass is a powerful constraint on age (e.g. Martig et al 2014)

- How can we measure masses for APOGEE objects?
We want to map stellar ages throughout the MW.

- APOGEE provides an extensive sample of giant stars

**Apache Point Observatory Galactic Evolution Experiment**
- SDSS near-IR (1.52-1.69 μ) stellar spectroscopic survey of MW disk, bulge, halo
- R ~ 22,500, S/N ~ 100
- 300 fibers
- 150,000 giants observed as of DR12

- For giant stars, mass is a powerful constraint on age (e.g. Martig et al 2014)
- How can we measure masses for APOGEE objects?
APOKASC: APOGEE-Kepler synergy

APOGEE + Kepler Asteroseismology Science Consortium
Pinsonneault et al 2014

2000 APOKASC objects
APOKASC: APOGEE-Kepler synergy

APOGEE + Kepler Asteroseismology Science Consortium
Pinsonneault et al 2014

APOGEE
Kepler

Spectroscopy \( \rightarrow \) \( T_{\text{eff}}, \log g, [\text{Fe/H}], \text{rv}, \) and 15 chemical abundances

2000 APOKASC objects
APOGEE + Kepler Asteroseismology Science Consortium
Pinsonneault et al 2014

**APOKASC:** APOGEE-Kepler synergy

- **APOGEE**
  - Spectroscopy \( \rightarrow \)
  - \( T_{\text{eff}}, \log g, [\text{Fe/H}], \text{rv}, \) and 15 chemical abundances

- **Kepler**
  - Asteroseismology \( \rightarrow \)
  - density, mass, radius

- **2000 APOKASC objects**

- **Age**
APOKASC: APOGEE-Kepler synergy

APOGEE + Kepler Asteroseismology Science Consortium
Pinsonneault et al 2014

This information is complementary but spatially limited.
Can we determine mass directly from APOGEE spectra?

This information is complementary but spatially limited.

APOGEE + Kepler Asteroseismology Science Consortium
Pinsonneault et al 2014

**APOGEE**

Spectroscopy → $T_{\text{eff}}$, logg, [Fe/H], rv, and 15 chemical abundances

**Kepler**

Asteroseismology → density, mass, radius

2000 APOKASC objects

age
The Cannon: *Data-driven labels* from spectra

M. Ness, D.W. Hogg, H.-W. Rix, A. Y. Q. Ho, G. Zasowski
Ness et al. 2015

*(Stellar parameters and abundances)*
The Cannon: *Data-driven labels* from spectra

1. Training Step
2. Test Step

*(Stellar parameters and abundances)*
**Training Step:** use the training set (spectra + labels) to fit a spectral model at each wavelength.

\[ f_{n\lambda} = g(\ell_n|\theta_{\lambda}) + \text{noise} \]
**Training Step:** use the training set (spectra + labels) to fit a spectral model at each wavelength.

\[
f_{n\lambda} = g(\ell_n | \theta_\lambda) + \text{noise}
\]

\[
\{T_{eff}, \log g, [Fe/H], [\alpha/Fe], M\}
\]
Training Step: use the training set (spectra + labels) to fit a spectral model at each wavelength.

\[ f_{n\lambda} = g(\vec{l}_n | \theta_\lambda) + \text{noise} \]

\[ f_{n\lambda} = \theta_\lambda^T \cdot \vec{l}_n + \text{noise} \]

\[ f_{n\lambda} = a_\lambda + b_\lambda (T_{\text{eff}})_n + c_\lambda (\log g)_n 
+ d_\lambda ([Fe/H])_n + e_\lambda ([\alpha/Fe])_n 
+ f_\lambda (M)_n + \text{(quadratic terms)} 
+ \text{scatter}_\lambda \]
Test Step: use the spectral model to infer new labels.

\[ f_{n\lambda} = a_\lambda + b_\lambda (T_{\text{eff}})_n + c_\lambda (\log g)_n + d_\lambda ([Fe/H])_n + e_\lambda ([\alpha/Fe])_n + f_\lambda (M)_n + (\text{quadratic terms}) + \text{scatter}_\lambda \]

\[ f_{n\lambda} = a_\lambda + b_\lambda (T_{\text{eff}})_n + c_\lambda (\log g)_n + d_\lambda ([Fe/H])_n + e_\lambda ([\alpha/Fe])_n + g_\lambda (M)_n + (\text{quadratic terms}) + \text{scatter}_\lambda \]
Cross-validation: 90% training, 10% test

\[ f_{n\lambda} = a_{\lambda} + b_{\lambda} (T_{\text{eff}})_n + c_{\lambda} (\log g)_n + d_{\lambda} ([Fe/H])_n + e_{\lambda} ([\alpha/Fe])_n + g_{\lambda} (M)_n + (\text{quadratic terms}) + \text{scatter}_\lambda \]
Cross-validation results (Ness et al. in prep)

RMS shown in box
Cross-validation results (Ness et al. *in prep*)

Mass determined directly from APOGEE spectra!
Cross-validation results (Ness et al. *in prep*)

Mass determined directly from APOGEE spectra!

None of these objects were used in training the model
Strengths & Limitations

Limitations:
• Training set

Strengths:
• Data-driven (no physical models required)
• Model is not a black box
• Fast
• Works at low-SNR
• Can handle noise and missing data in spectra
What’s next for *The Cannon*?

- Age distribution for 85,000 APOGEE giants (Ness et al. *in prep*)
- Improve training set (8000 new APOKASC objects in January)
- Cross-calibration & label transfer for large spectroscopic surveys (e.g. APOGEE & LAMOST; Ho et al. *in prep*)
- Part of label determination pipelines (e.g. 4MOST, GALAH)
- Classification?
- Gaussian processes?
- Galaxy spectra? Low-resolution spectra?

*The Cannon* is a powerful (and fast!) technique for squeezing out the information present in a spectrum

[https://github.com/annayqho/TheCannon](https://github.com/annayqho/TheCannon)
[https://annayqho.github.io/TheCannon/](https://annayqho.github.io/TheCannon/)
Where does The Cannon get information on mass and age? (Ness et al. *in prep*)
Sample RC star spectral model (red=model, black=APOGEE spectrum)
Results:

APOGEE objects NOT observed by *Kepler*

\[ f_{n\lambda} = a_\lambda + b_\lambda (T_{\text{eff}})_n + c_\lambda (\log g)_n + d_\lambda ([Fe/H])_n + e_\lambda ([\alpha/Fe])_n + g_\lambda (M)_n + (\text{quadratic terms}) + \text{scatter}_\lambda \]
$[\text{Fe/H}]-[\alpha/\text{Fe}]$ Plane, 20,000 APOGEE RC Stars
(Ness et al in prep)
(NOT Kepler objects)