HiCannon: Data-driven spectroscopy of cool stars at high spectral resolution

Aida Behmard, Erik Petigura, Andrew Howard
Why is it important to determine stellar parameters / abundances ($R_*, M_*, T_{\text{eff}}, [\text{Fe/H}], \text{ etc.}$)?

Transit detections of exoplanets

Need $R_*$ to determine planet radius $R_p$
Why is it important to determine stellar parameters / abundances ($R_*, M_*, T_{\text{eff}}, [\text{Fe/H}],$ etc.)?

Correlations between stellar abundances and planets properties

Stars with high $[\text{Fe/H}]$ more likely to host giant planets (Fischer & Valenti + 2005)
How do we determine stellar parameters / abundances before?

Synthetic spectral libraries

**SpecMatch-Syn (Petigura+2015)**

![Graph showing spectral lines for different effective temperatures](image)
Synthetic spectral libraries make good abundance/parameter predictions for these stars.
The Cannon

Data-driven (ML) approach for predicting stellar “labels” (parameters + elemental abundances) from spectra

Stellar “labels”:
- Temperature $T_{\text{eff}}$
- Bulk metallicity $[\text{Fe}/\text{H}]$
- Stellar radius $R_*$

The Cannon was developed by Melissa Ness (MPIA), Andy Casey (Monash U.), and Anna Ho (Caltech)
The Cannon

“Training Step”

Construct flux model

\[ f_{jn} = V(\lambda_n) \cdot \Theta_j + \text{noise} \]

“Test Step”

Apply \( f_{jn} \) to test set spectra
**Cannon flux model fitting:**

\[ l_n = [1, T_{\text{eff}}, R_*, [\text{Fe/H}] \ldots ] \]
High Resolution Echelle Spectrometer (HIRES) sample

141 stars, K and M dwarfs:

- $3000 \text{ K} < T_{\text{eff}} < 5200 \text{ K}$
- No giants ($R_* < 1 R_\odot$)

Labels ($T_{\text{eff}}$, $R_*$, [Fe/H]):

- Interferometry
- SED Modeling
- Gaia Parallaxes
Evaluating **Cannon** performance: Cross-validation bootstrap scheme

iteratively:

Whole Sample

"Training Set": $n-1$

"Validation Set": 1
The Cannon cannot make predictions for spectra not well-represented in training set!

**GL896A**: a fast rotator
...diversify the training set!

Copy each spectrum in training set and artificially broaden

+0-20 km/s
$T_{\text{eff}} \approx 70$ K; $R_\ast \approx 5\%$; $[\text{Fe/H}] \approx 0.08$ dex
Conclusion
(with some modifications to the spectral sample), *The Cannon* is able to make label predictions for cool stars comparable to the best alt. methods (but is easier to use – data driven!)
Future work

Consider including prior information (line lists, etc.)

Elemental abundance studies
Prediction residuals

- 

\[\Delta T_{\text{eff}} \text{ [K]}\]

- 

\[\Delta R_*/R_*\]

- 

\[\Delta [\text{Fe/H}] \text{ [dex]}\]
Flux model

“complex vectorizer” function \( \mathbf{V} \) model coefficients

\[
f_{jn} = \mathbf{V}(l_n) \cdot \mathbf{v}_j + \text{noise}
\]

\[
l_n = [1, T_{\text{eff}}, R^*, [\text{Fe/H}] \ldots]
\]

“Training Step” : fit for model coefficients \( \mathbf{v}_j \) for each flux model

“Test Step” : fit for labels \( l_n \) for each star in the validation set that best reproduces empirical flux
Synthetic spectral libraries struggle with small, cool stars…

Too bad! Good for finding small, cool planets

Transit method
Radial velocity method

...And most common stars in the galaxy

M-dwarfs = ~75% of stars in solar neighborhood!