A Data-Driven Spectral Model of Main Sequence Stars in Gaia DR3

Isabel Angelo (UCLA)

Collaborators: Megan Bedell, Erik Petigura, Smadar Naoz, Melissa Ness
ExSoCal December 11, 2023
Most known planets were identified using photometry

5,550 confirmed planets
2,778 from Kepler
548 from K2
410 from TESS
plus >10,000 candidates from photometry
We need high resolution spectroscopy to learn more about Kepler, K2 and TESS planets.

- How often do planets orbit active stars?
- How often do planets form in binaries?
- What are the demographic trends in the population? (e.g., radius gap, Fulton et al. 2017)

We need spectroscopy to answer these science questions!
We need high resolution spectroscopy to learn more about Kepler, K2 and TESS planets.

- How often do planets orbit active stars?
- How often do planets form in binaries?
- What are the demographic trends in the population? (e.g., radius gap, Fulton et al. 2017)

...we need spectroscopy to answer these science questions!

...but this is challenging :/
What if you didn’t need Keck or Gemini to characterize planet hosts?

Gaia DR3 published spectra for 10,000 Kepler field stars and 37,000 TESS field stars! (~6% of Kepler planet hosts ~20% of TESS planet hosts)
We used the Cannon to train a data-driven spectroscopic model that enables precise characterization of stars with Gaia DR3 spectra. The training set consists of Gaia RVS spectra and corresponding labels (\(T_{\text{eff}}\), \(\log g\), [Fe/H], [\(\alpha/Fe\)], \(v_{\text{broad}}\)) of single stars.

The training step involves supervised learning, where an algorithm fits the flux of real spectra to previously determined labels.

The single star model is given by:

\[ f(\lambda, \text{labels}) \]
Our Gaia RVS model computes stellar properties with comparable accuracy to ground-based surveys.
Our model also computes metrics to identify spectroscopic anomalies

\( \chi^2 \): measures how well the model fits the data

\( \rho_{\text{training}} \): measures how reasonable the predicted star properties are

\( \chi_{\text{Ca}}^2 \): measures how well the model fits the data around activity-correlated Calcium features
A fourth metric, $\Delta \chi^2_{\text{binary}}$, helps to identify spectroscopic binaries by comparing fits of single star and binary (i.e., composite) spectra.
A fourth metric, $\Delta \chi^2_{\text{binary}}$, helps to identify spectroscopic binaries by comparing fits of single star and binary (i.e., composite) spectra.
A fourth metric, $\Delta \chi^2_{\text{binary}}$, helps to identify spectroscopic binaries by comparing fits of single star and binary (i.e., composite) spectra sensitive to binaries missed by AO and RV surveys!
Our spectral model metrics can be used to identify stellar activity, binary stars, and evolved stars.

`main sequence single star`

Gaia spectrum

20 x residuals

`model produces good fit, metrics are reasonable`
Our spectral model metrics can be used to identify stellar activity, binary stars, and evolved stars.

*model produces poor fit everywhere, $\chi^2$ is large and $\rho_{\text{training}}$ is small*
Our spectral model metrics can be used to identify stellar activity, binary stars, and evolved stars.

Active star

Gaia spectrum

Best-fit model

20 x residuals

Model produces poor fit everywhere except activity-correlated lines, $\chi_{Ca}^2$ is large.
Our spectral model metrics can be used to identify stellar activity, binary stars, and evolved stars.

The Gaia spectrum is shown with two models: a best-fit single star model and a best-fit binary model. The residuals for the binary model are significantly lower, indicating that the model is better fit by the binary model. The statistic $\Delta\chi^2_{binary}$ is large, suggesting a significant improvement in fit.
Summary

• most planets are discovered with photometry and require spectroscopic follow-up to characterize

• Gaia DR3 released spectra for thousands of stars in the Kepler and TESS fields (with more to come in DR4!)

• our data-driven spectral model computes stellar properties with comparable precision to ground-based surveys, and establishes metrics to identify binaries, stellar activity, and evolved stars!

stay tuned for published model and planet host properties! (Angelo et al 2024)
Close binaries in Kepler are mistaken for single stars. This can lead to:

- diluted transits
- underestimated planet radii

Errors in planet demographic trends like the radius gap (Fulton et al. 2017)
A complete sample of binaries among planet hosts will allow us to test theories of how binaries sculpt the exoplanet population:

- **Do binaries suppress planet formation?**
  - e.g., Kraus et al. 2016, Hirsch et al. 2021

- **Do binaries sculpt dynamically rich orbits?**
  - e.g., Naoz et al. 2012, Becker & Adams 2017, Li et al. 2014

- **Do binaries deplete planet reservoirs?**
  - e.g., Quintana et al. 2007, Jang-Condell 2015
Previous searches for binaries among Kepler planet hosts

- Moe & Di Stefano 2017, Raghavan et al. 2010
- Kraus et al. 2016, Horch et al. 2014

offset spectra
Kolbl et al. 2015

ground-based imaging + AO
Kraus et al. 2016, Horch et al. 2014
Spectroscopic binaries are traditionally identified by their characteristic double-lined spectra (e.g., Kolbl et al. 2015)

traditional method: identify 2 sets of spectral features

caveat: requires >10 km/s relative velocity between stars (i.e. $a \leq 1\text{au}$)

figure credit: Angelo et al., in prep
Advances in spectrum modeling allow us to identify binaries with *overlapping* spectral features, thereby increasing sensitivity (e.g., Burgasser et al. 2010, El-Badry et al. 2018)

*figure credit: Angelo et al., in prep*
We’ve also uncovered a number of stars with superimposed Calcium emission + absorption

possible explanations:
interacting binaries
young accreting T Tauri stars
chromospheric activity in GKM main-sequence stars
emission linked to the accretion of matter in the stellar magnetosphere