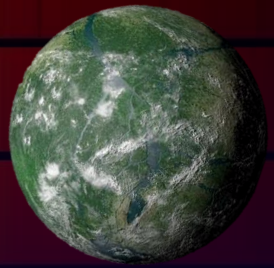
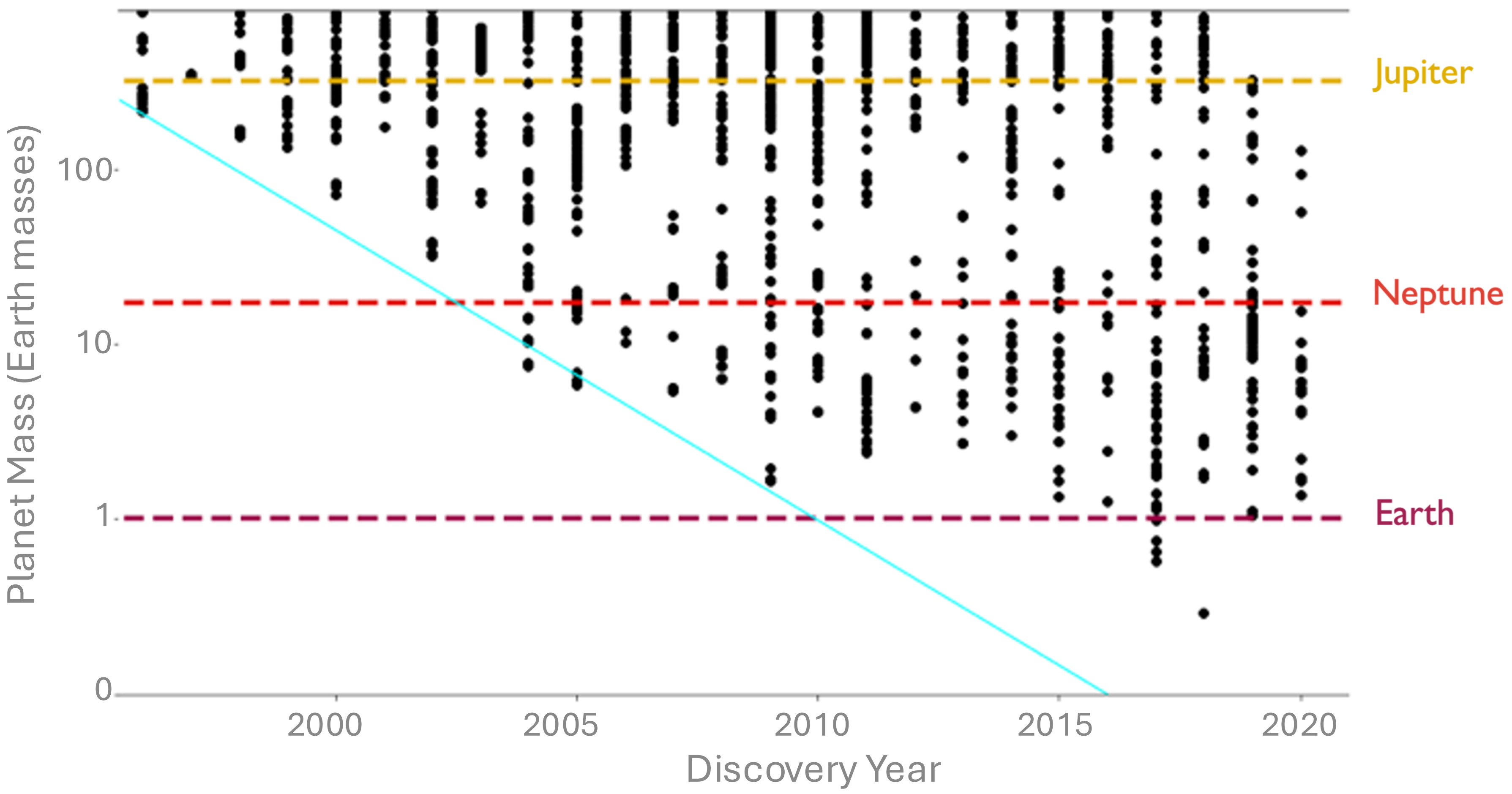


Enhancing radial velocity precision with machine learning

Zoë de Beurs, MIT
Andrew Vanderburg, Julien de Wit,
Naomi McWilliam

Know Thy Star
Know Thy Planet 2,
February 5, 2025

We have reached a plateau in increasing our sensitivity to smaller mass planets due to stellar variability

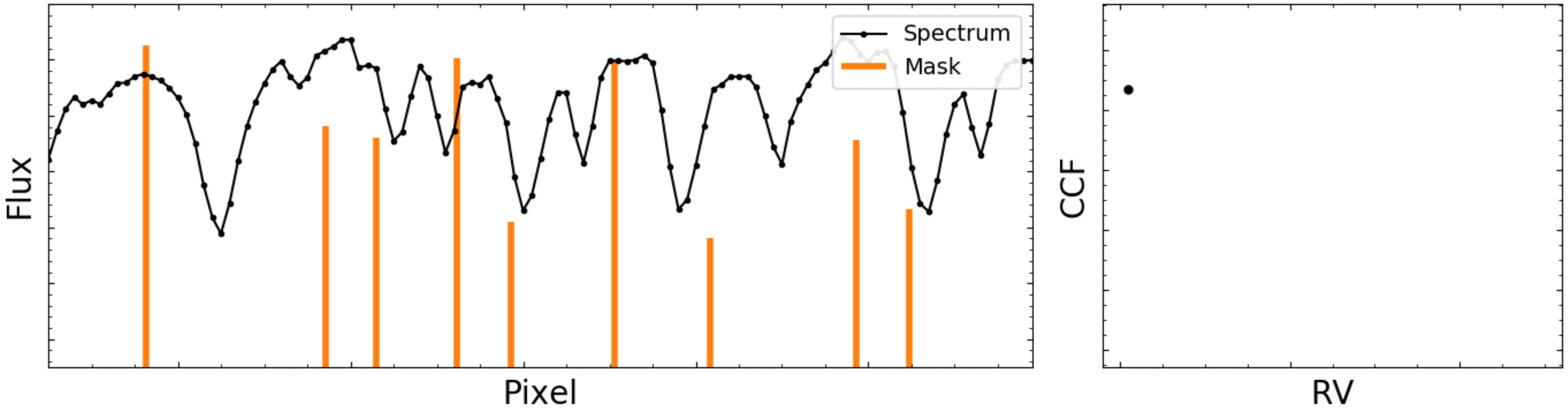


Background | NNs for solar RVs | CALM across the HR diagram | Nonlinear PCA | Beyond CCFs and instrumental noise | Takeaways

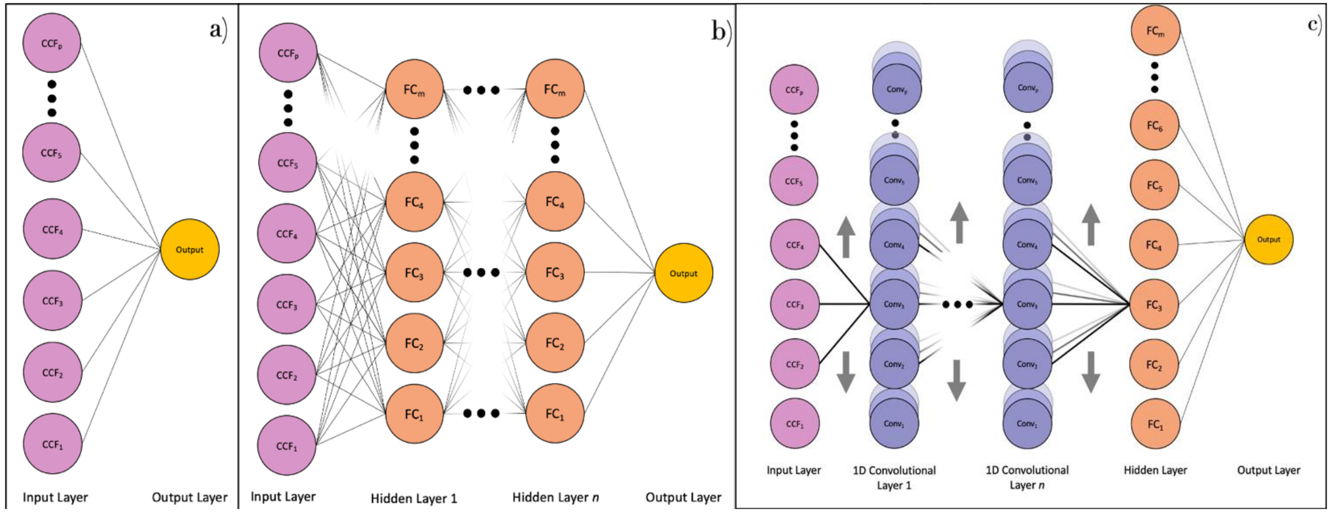
Cross-correlation functions (CCF) capture shape changes in spectral lines

- CCFs are represent the average line profile of a star
- Stellar variability encodes its signatures in the individual spectral lines
- CCFs are sensitive to line shape changes that persist in most spectral lines

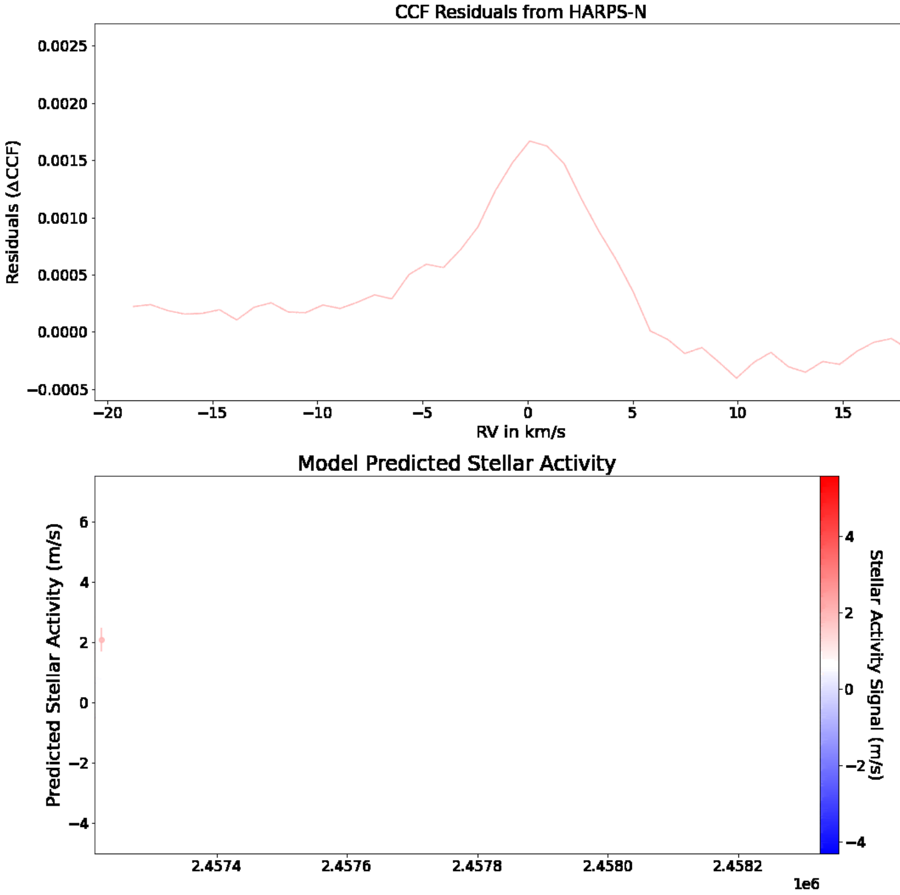
Animation by Marina Lafarga



Neural networks reduce RV scatter for solar observations by 40% using CCF shape changes



Linear | Fully Connected | Convolutional



de Beurs et al. 2022

For other stars, we developed the CALM linear regression method to model stellar variability

Designed to address two Challenges:

1. Modeling stellar variability patterns with relatively small datasets (50-100 datapoints) requires a smaller number of free parameters
2. Simultaneously modeling planetary signals

CCF Activity Linear Model (CALM)

$$a_1 \cdot CCF_1 + a_2 \cdot CCF_2 + \ldots\ a_n \cdot CCF_n + b_1 \cdot Kepler_1 + b_2 \cdot Kepler_2 + \ldots + b_m \cdot Kepler_m = \mathbf{RV}$$

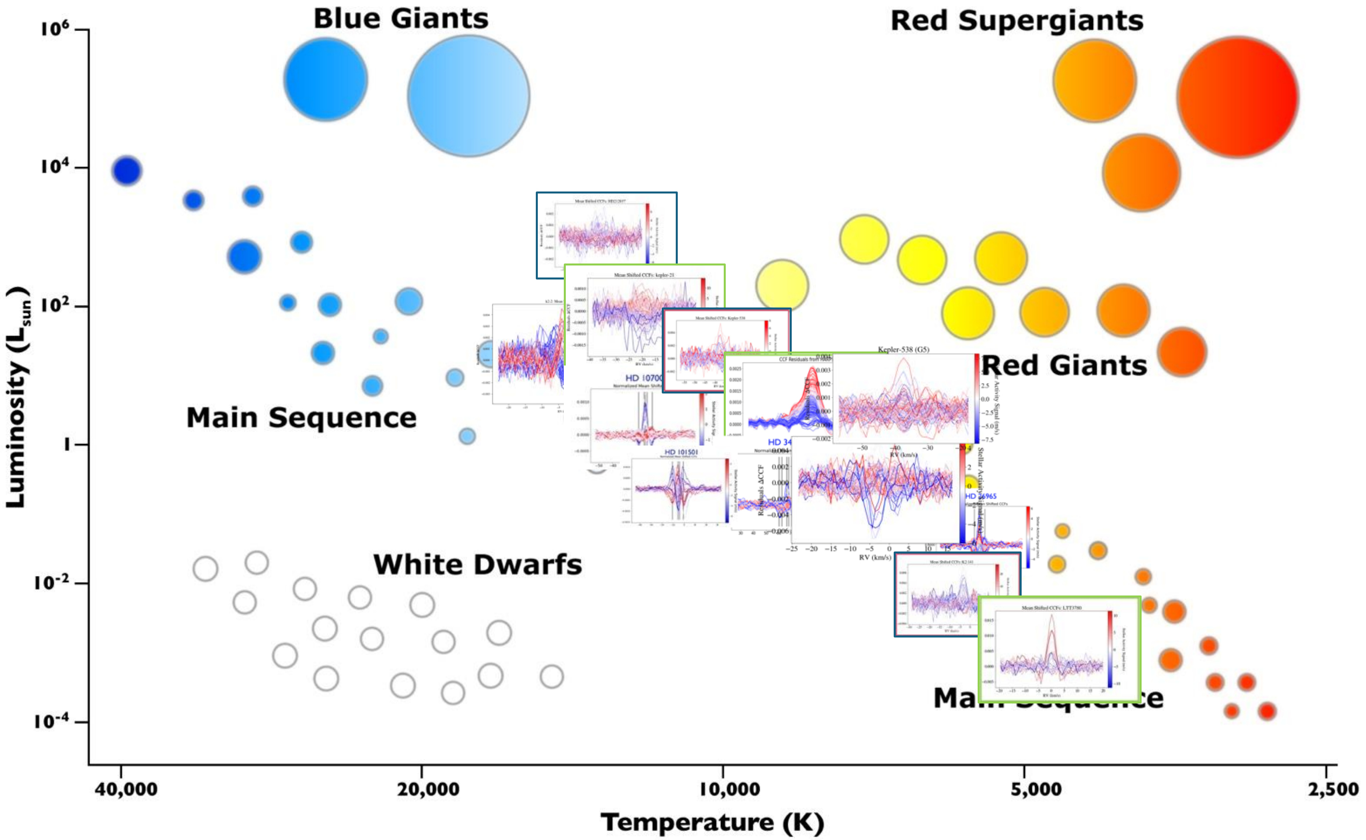
Stellar variability

Planetary Doppler Shifts

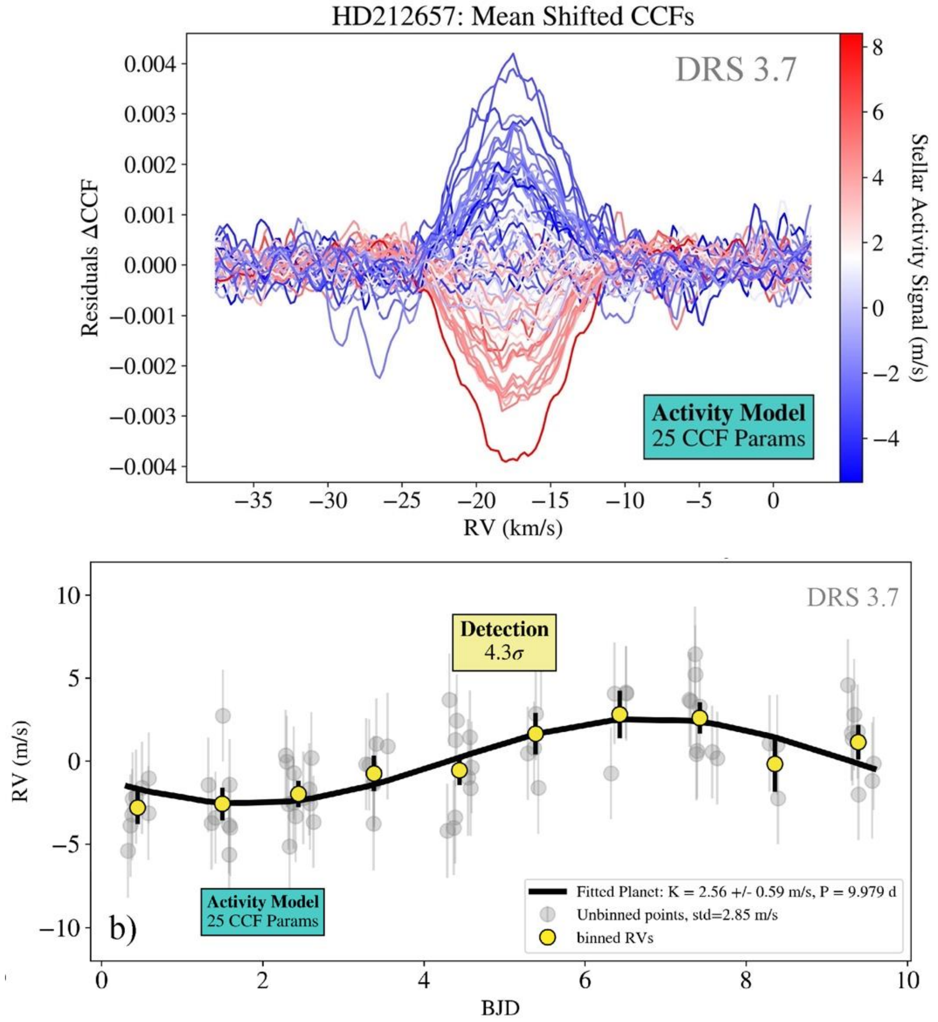
Background | NNs for solar RVs | CALM across the HR diagram | Nonlinear PCA | Beyond CCFs and instrumental noise | Takeaways

Our stellar sample

- F type
  - Kepler-21 – F6
  - K2-167 – F7
  - HIP41378 – F8
- G type
  - HD 34411 – G0
  - The Sun – G2
  - Kepler-538 – G5
  - HD 101501 – G8
  - Tau Ceti – G8
- K-type
  - K2-2 – K0
  - HD 26965 – K1
  - Kepler-78 – K2
  - K2-141 – K4
- M-type
  - HD 79211 – M0
  - TOI-700 – M2
  - LTT3780 - M4

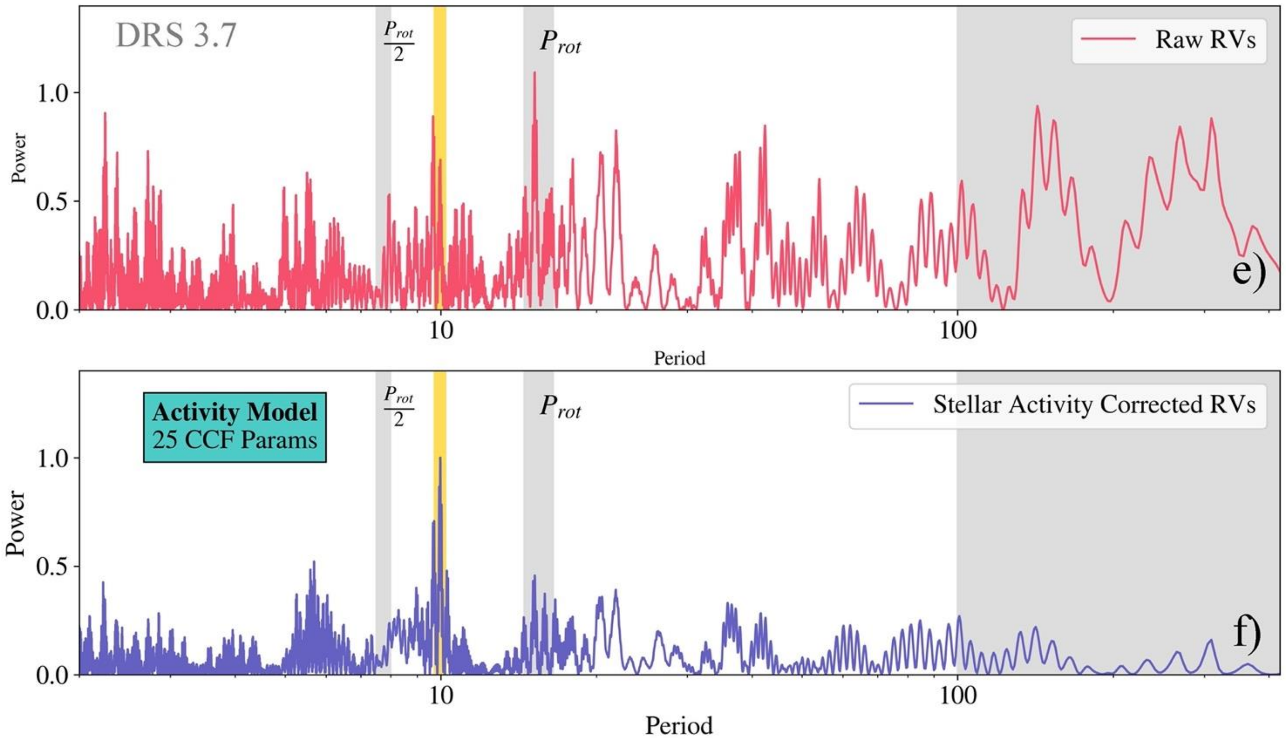


CALM Results - K2-167 (F7)



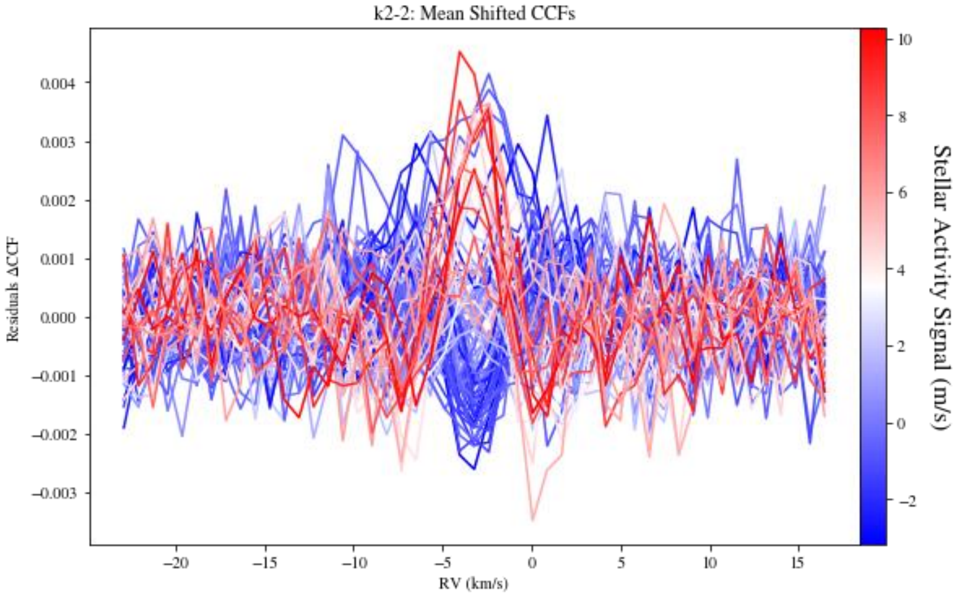
σ_{raw} = 4.02 m/s
σ_{us} = 2.04 m/s
Planet detection significance: 4.3σ

de Beurs et al. 2024



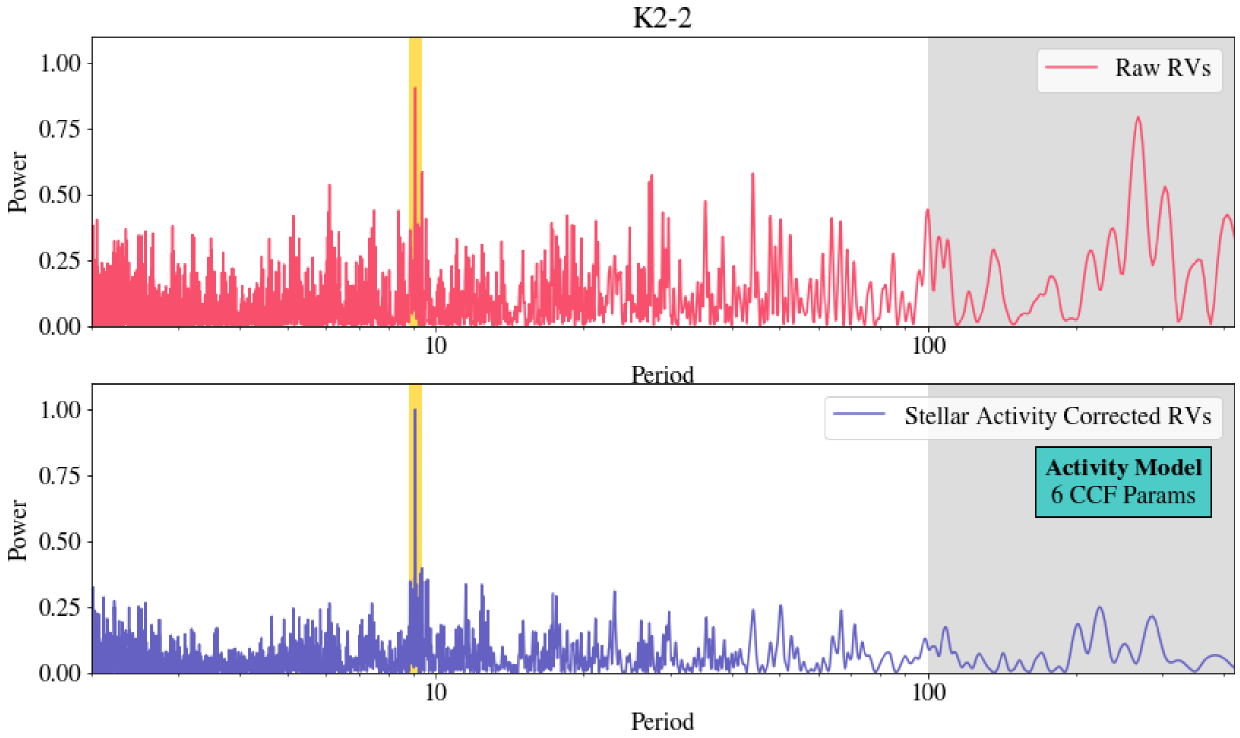
Background | NNs for solar RVs | Ch.2: CALM across the HR diagram | Nonlinear PCA | Beyond CCFs and instrumental noise | Takeaways

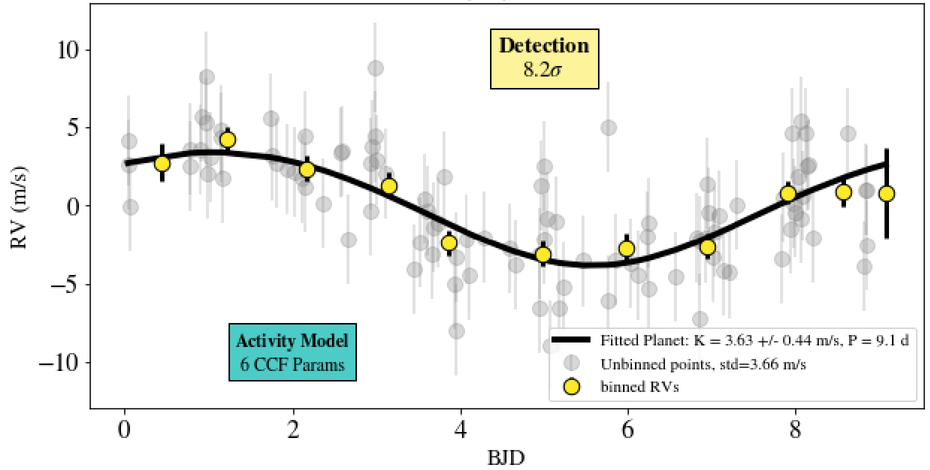
CALM Results - K2-2 (K0)



σ_{raw} = 4.72 m/s
σ_{us} = 2.70 m/s
Planet detection significance: 8.2σ

Thygesen, Rodriguez, de Beurs et al. 2024



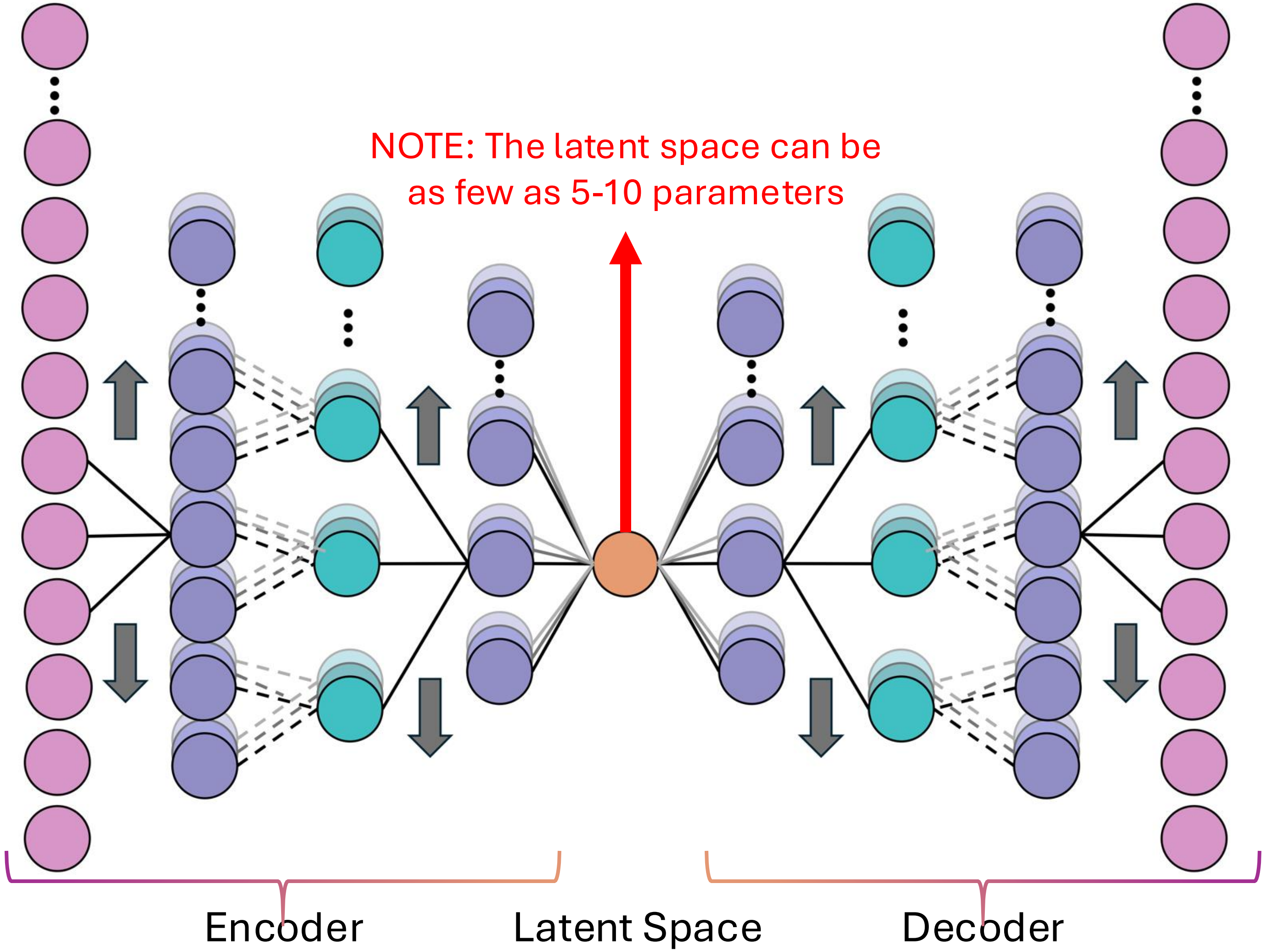


Background | NNs for solar RVs | CALM across the HR diagram | Nonlinear PCA | Beyond CCFs and instrumental noise | Takeaways

CALM uses only a subset of the CCF to model activity

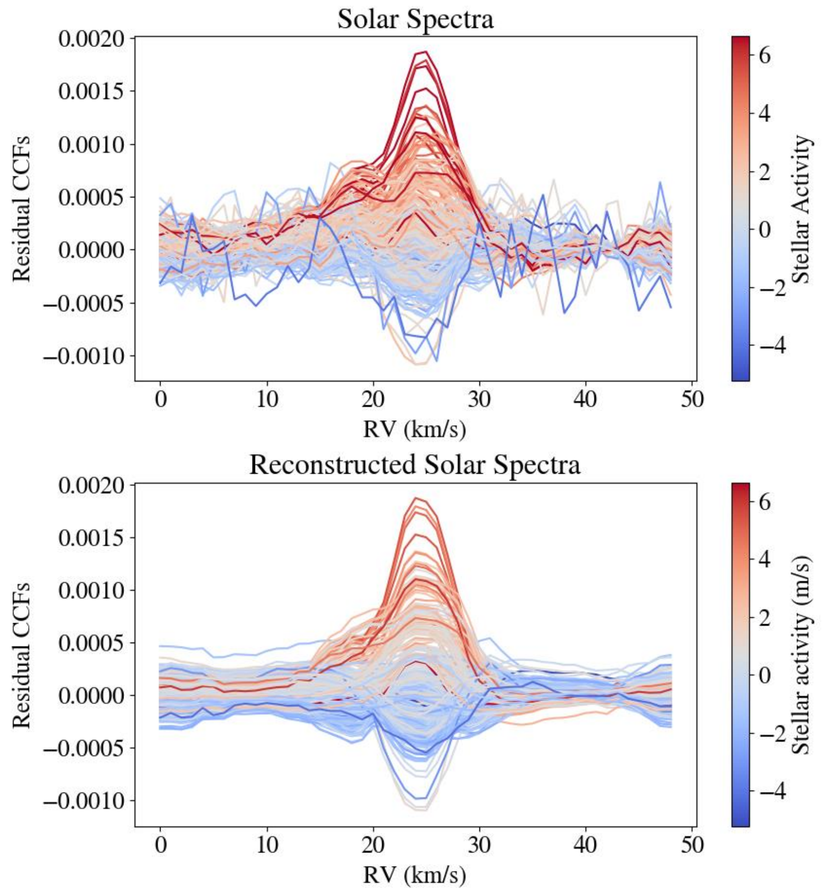
Using a subset of the CCF decreases the number of free parameters, but results in a loss of information.

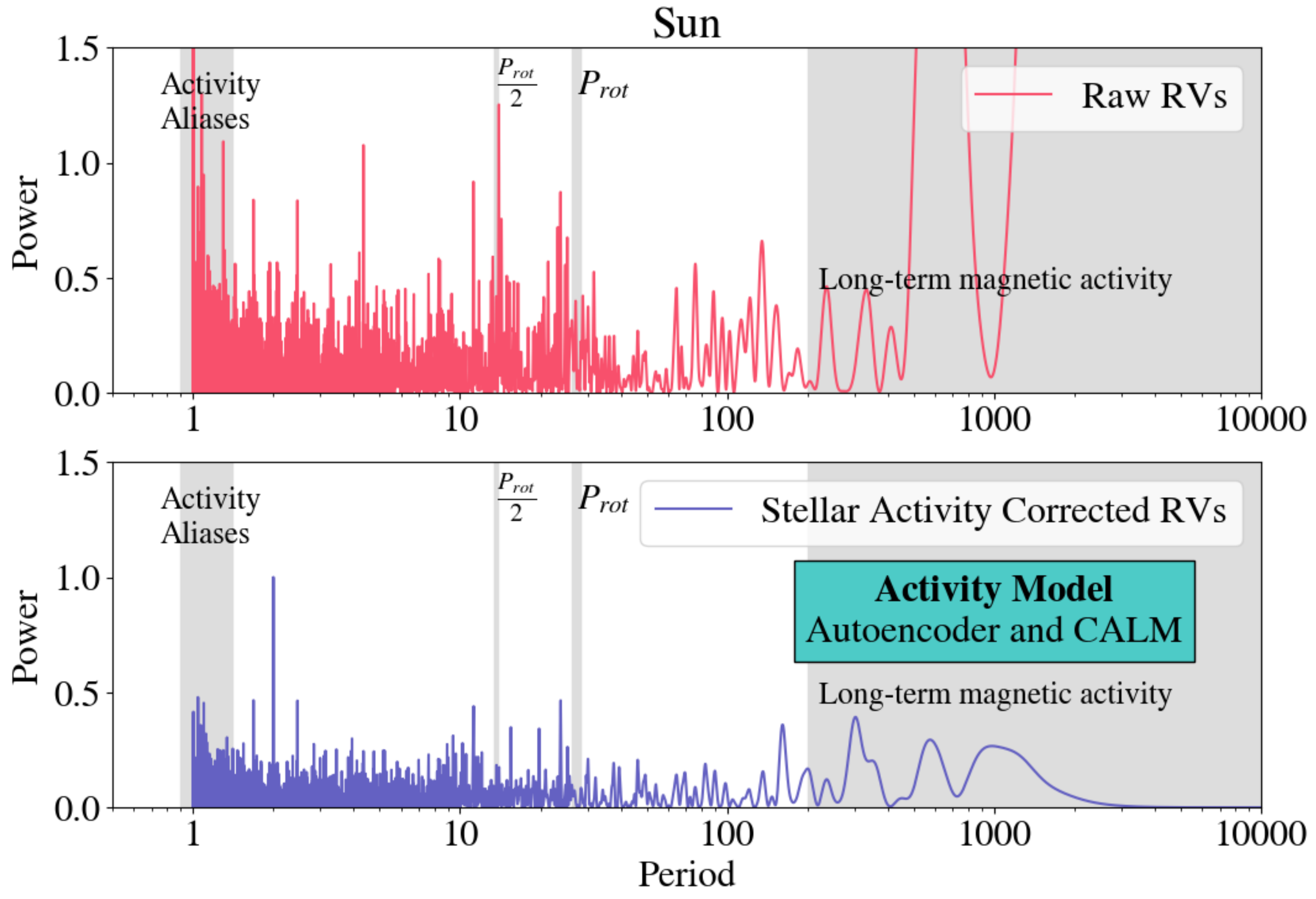
- Could autoencoders (nonlinear PCA) extract the most important information from CCFs?
- Can autoencoders do this without timing information and without stellar activity labels?
- Could this work for stars beyond the Sun?



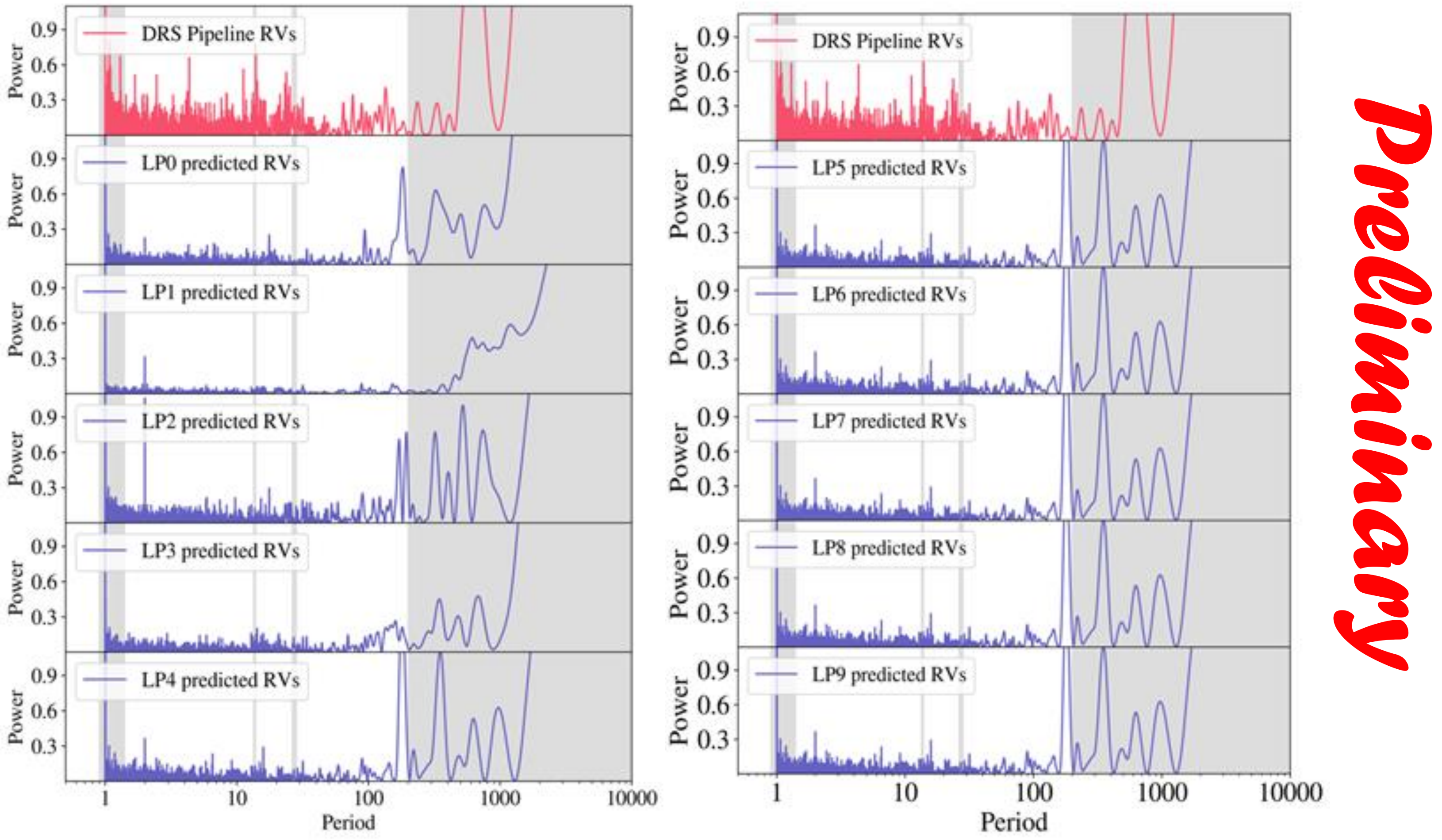
Background | NNs for solar RVs | CALM across the HR diagram | **Nonlinear PCA** | Beyond CCFs and instrumental noise | Takeaways

Autoencoders can reproduce key features from solar CCFs using 10 latent params and reduce the RV scatter

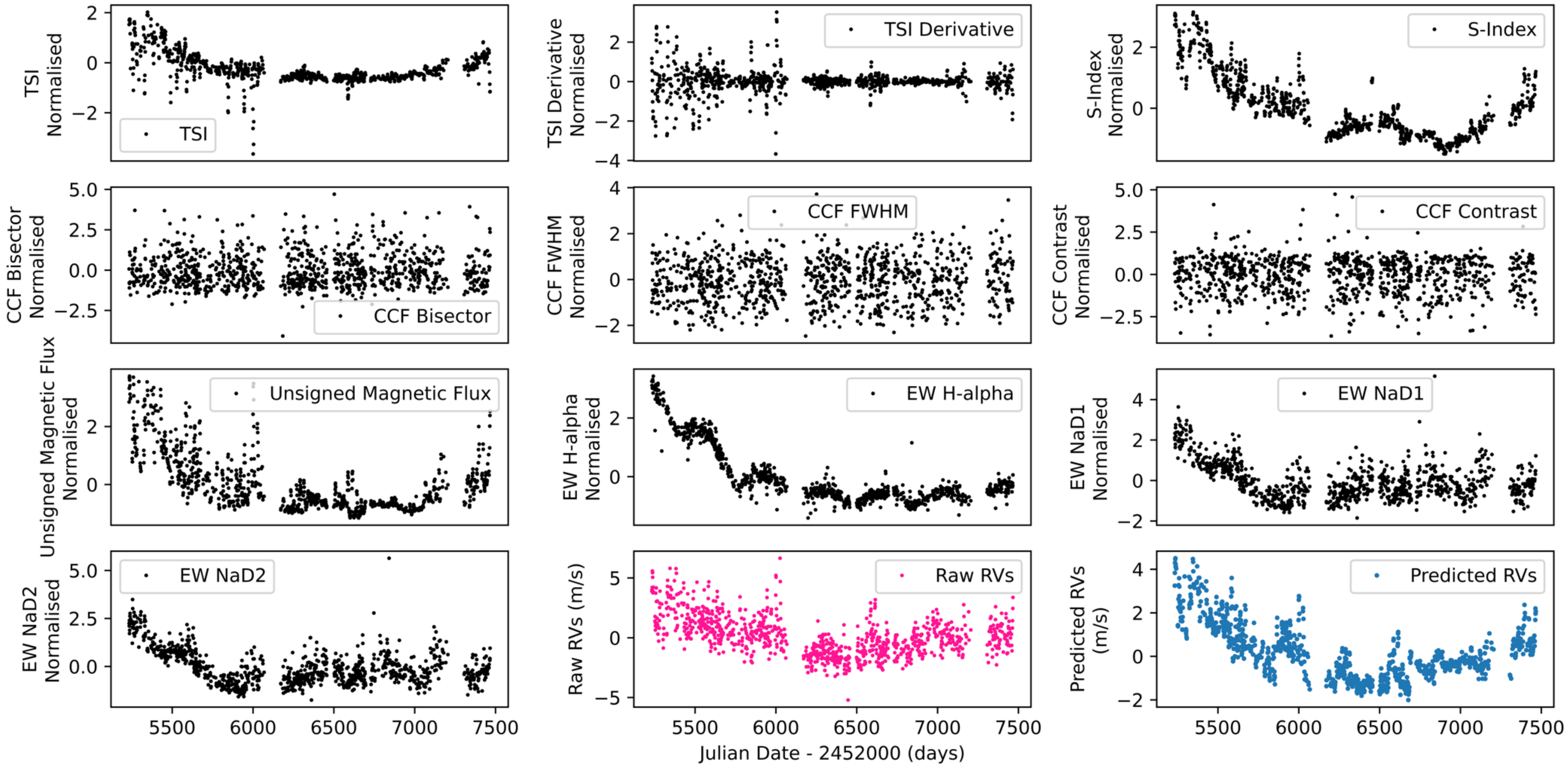




The latent parameters appear to trace CCF phenomena at different timescales



Beyond CCFs, we trained a neural network a multitude of common activity tracers to find orthogonal information

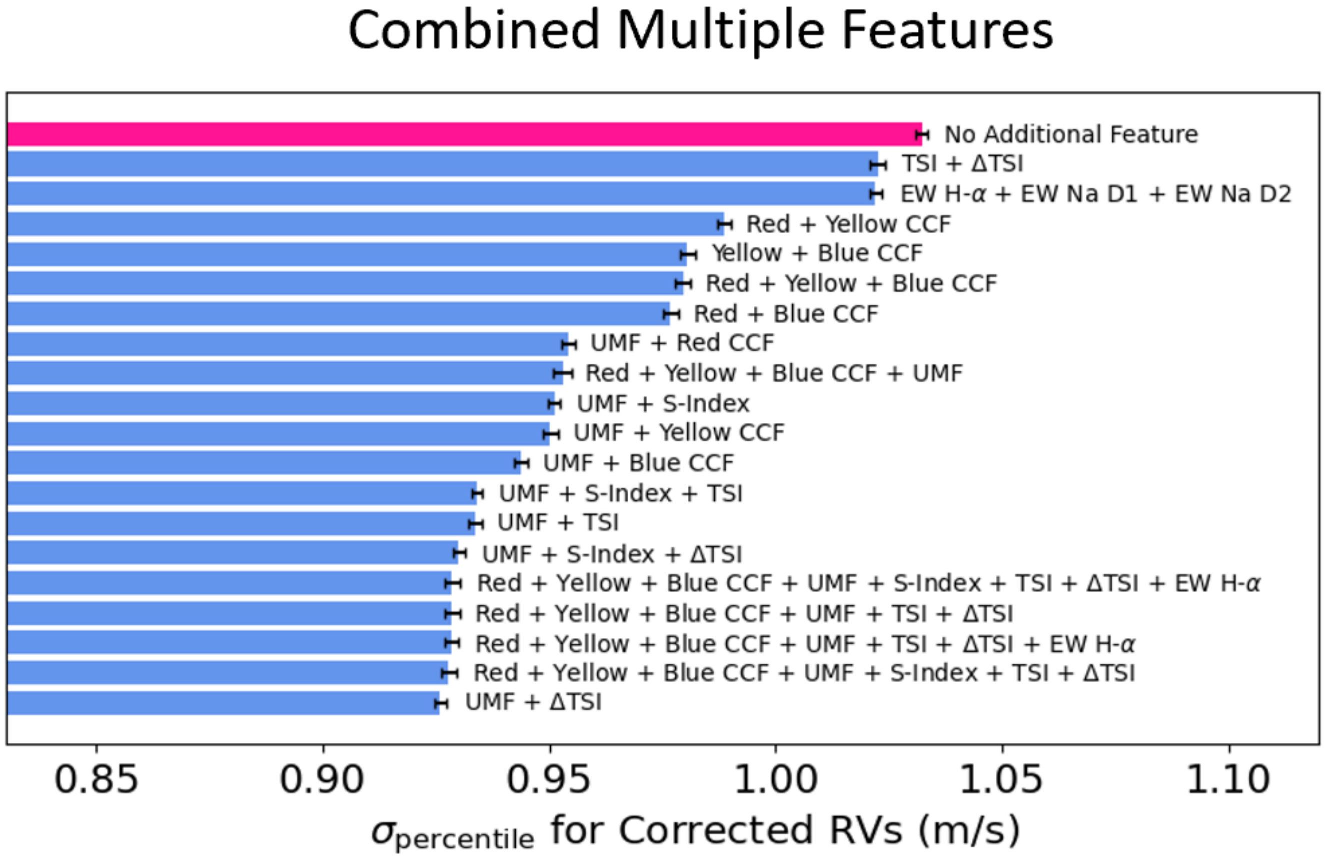


Work by Naomi McWilliam, PhD student at ICL

McWilliam, de Beurs, Vanderburg et al. Under Review

Background | NNs for solar RVs | CALM across the HR diagram | Nonlinear PCA | **Beyond CCFs and instrumental noise** | Takeaways

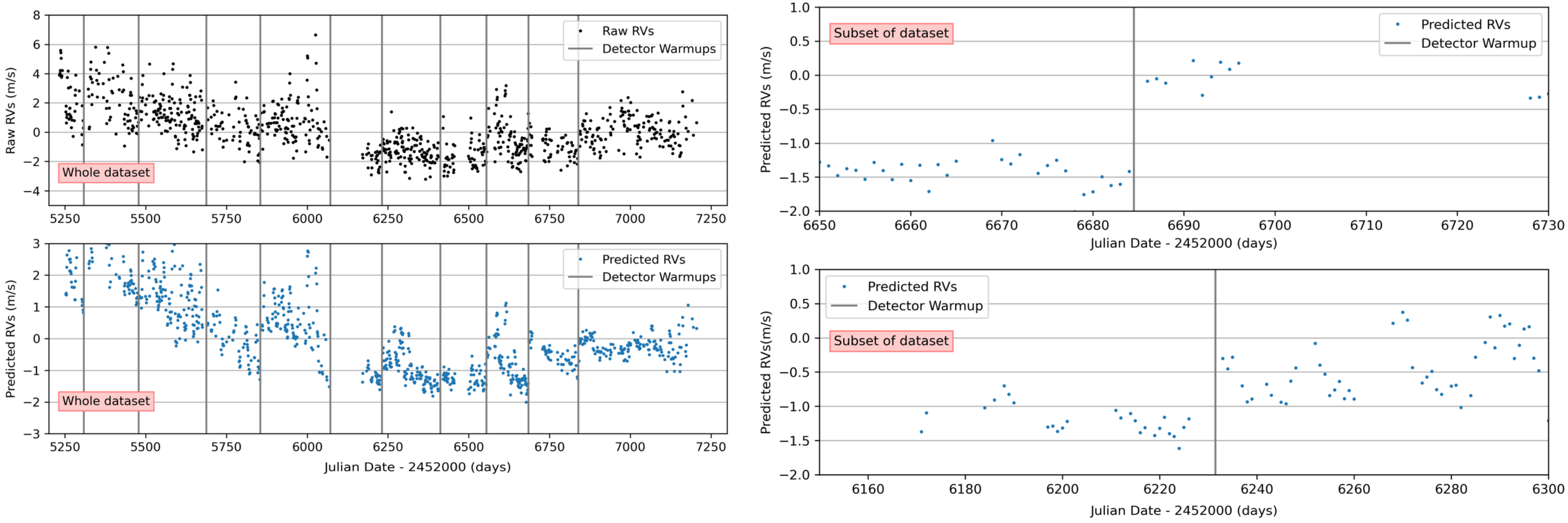
Adding photometry and unsigned magnetic flux improves the neural network performance



McWilliam, de Beurs, Vanderburg et al.
Under Review

Background | NNs for solar RVs | CALM across the HR diagram | Nonlinear PCA | Beyond CCFs and instrumental noise | Takeaways

Know Thy Star and Know Thy Instrument?



Background | NNs for solar RVs | CALM across the HR diagram | Nonlinear PCA | Beyond CCFs and instrumental noise | Takeaways

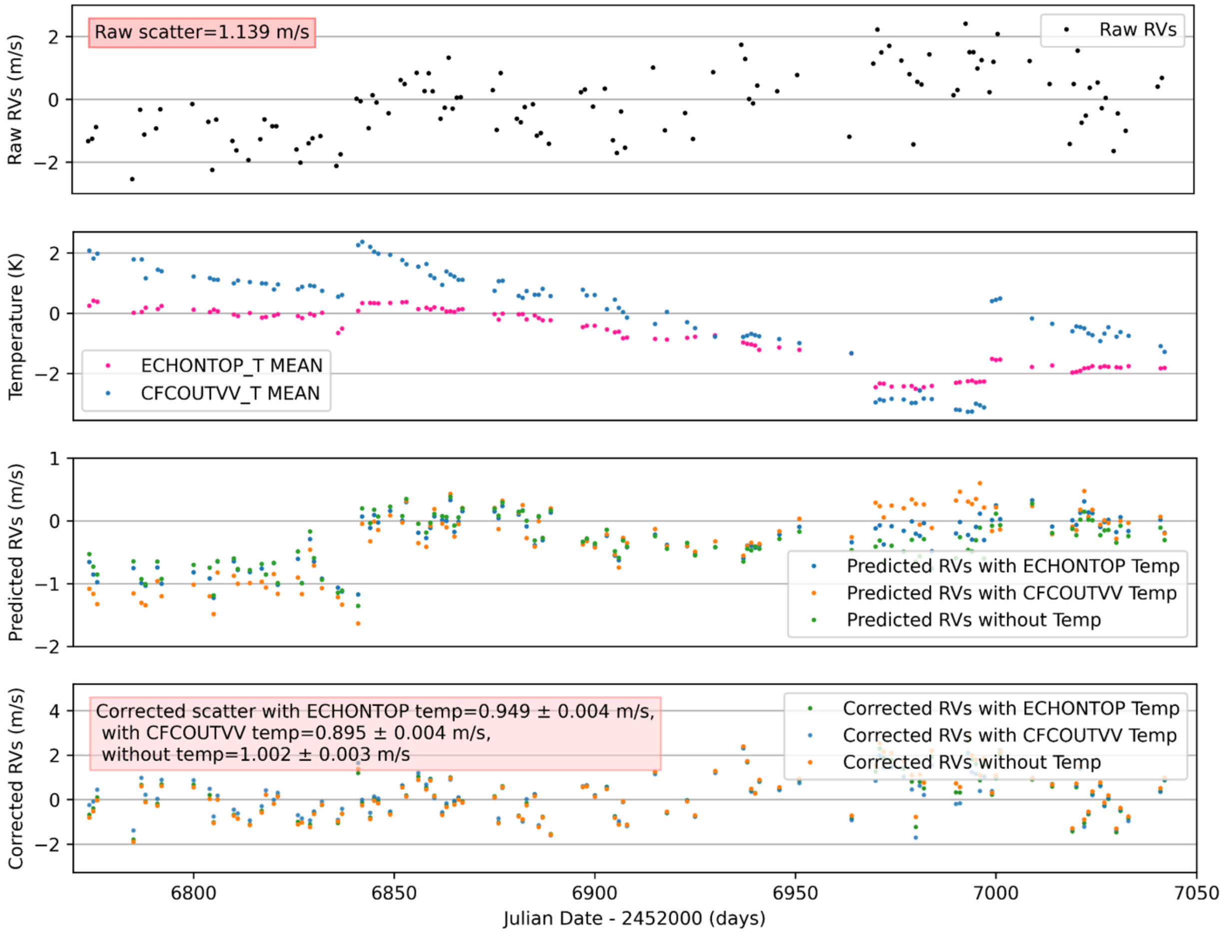
Including temperature readings in our NN model reduces the RV scatter by 10% during low magnetic activity

Low Magnetic Activity [Oct 2019 to July 2020]

Raw RV Scatter: 113.9 cm/s

Corrected RV scatter

- Without temperature readings: 100.2 ± 0.3 cm/s
- With ECHONTOP* temperature: 94.9 ± 0.4 cm/s
- With CFCOUTVV** temperature: 89.5 ± 0.4 cm/s



* ECHONTOP = Echelle grating on top

** CFCOUTVV = Temperature of cryostat (output vacuum vessel)

Background | NNs for solar RVs | CALM across the HR diagram | Nonlinear PCA | Beyond CCFs and instrumental noise | Takeaways

Takeaways

- We trained neural networks on solar CCFs to predict the stellar variability contributions and reduced the RV scatter by 40% (de Beurs et al. 2022)
- Our CALM method reduces the RV scatter for HARPS-N data, constraining the mass of K2-167 b (de Beurs et al. 2024) and increase the detection significance of K2-2 (Thygeson, Rodriguez, de Beurs et al. 2024)
- Autoencoders seems a promising path towards constructing stellar activity tracers that maximizing information content while minimizing the number of free parameters
- Including photometry and unsigned magnetic flux in addition to CCFs improves NN performance and reduces solar RV scatter by additional 10% (McWilliam, de Beurs, et al. Under review)
- Neural networks can correct for instrumental systematics → Know Thy Star and Know Thy Instrument