

Know Thy Star, Know Thy Planet 2

Pushing down the solar RV precision to 50 cm/s

**with a data-driven
correction at the spectral level**

Sara Tavella

X. Dumusque, M. Cretignier



**UNIVERSITÉ
DE GENÈVE**



**Swiss National
Science Foundation**

03 February 2025

Why EPRVs

Atmospheric characterization
of Earth-like planets

Space missions
(e.g. HWO)



Why EPRVs

Atmospheric characterization
of Earth-like planets

Space missions
(e.g. HWO)

Future
Goals

Discovery of (many)
Earth-like planets

Within ~ 15 pc from us, likely
through RVs only

Why EPRVs

Reach a very high
precision in the radial
velocity technique



~10 cm/s precision
needed to detect an
Earth-like planet

Atmospheric characterization
of Earth-like planets

Space missions
(e.g. HWO)

Future
Goals

Discovery of (many)
Earth-like planets

Within ~15 pc from us, likely
through RVs only

Compute Radial Velocities

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta v}{c}$$

Cross Correlation Functions (CCFs)

- Cross correlation of observed spectra with a tailored mask of stellar lines
- Gaussian fit to an average line profile to determine the radial velocity



Information loss

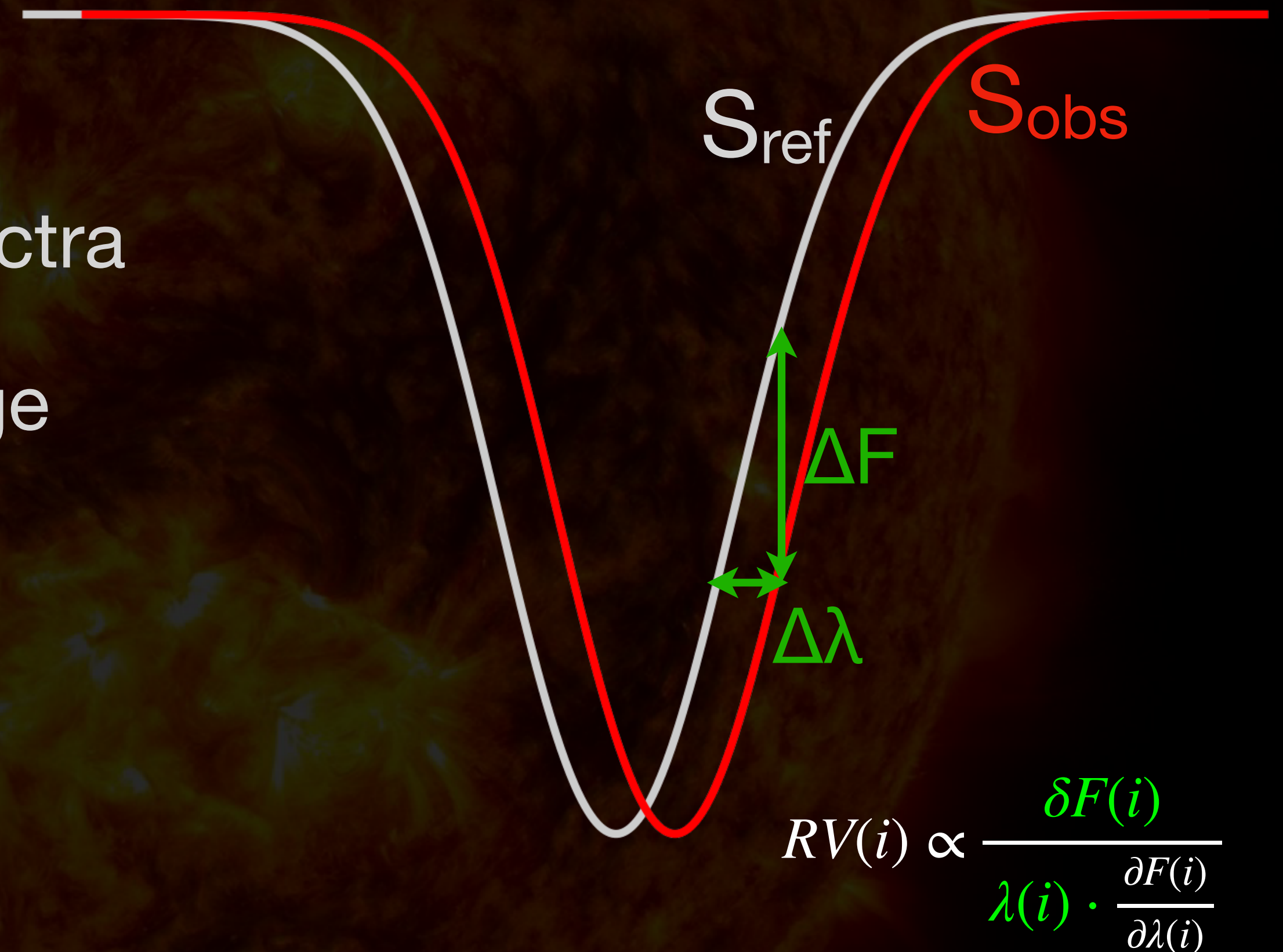
Wavelength-dependent features may be averaged out

Compute Radial Velocities

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta v}{c}$$

Line-by-line (LBL) Radial Velocities

- Doppler shift measured locally on the spectra
- Final Radial Velocity as a weighted average

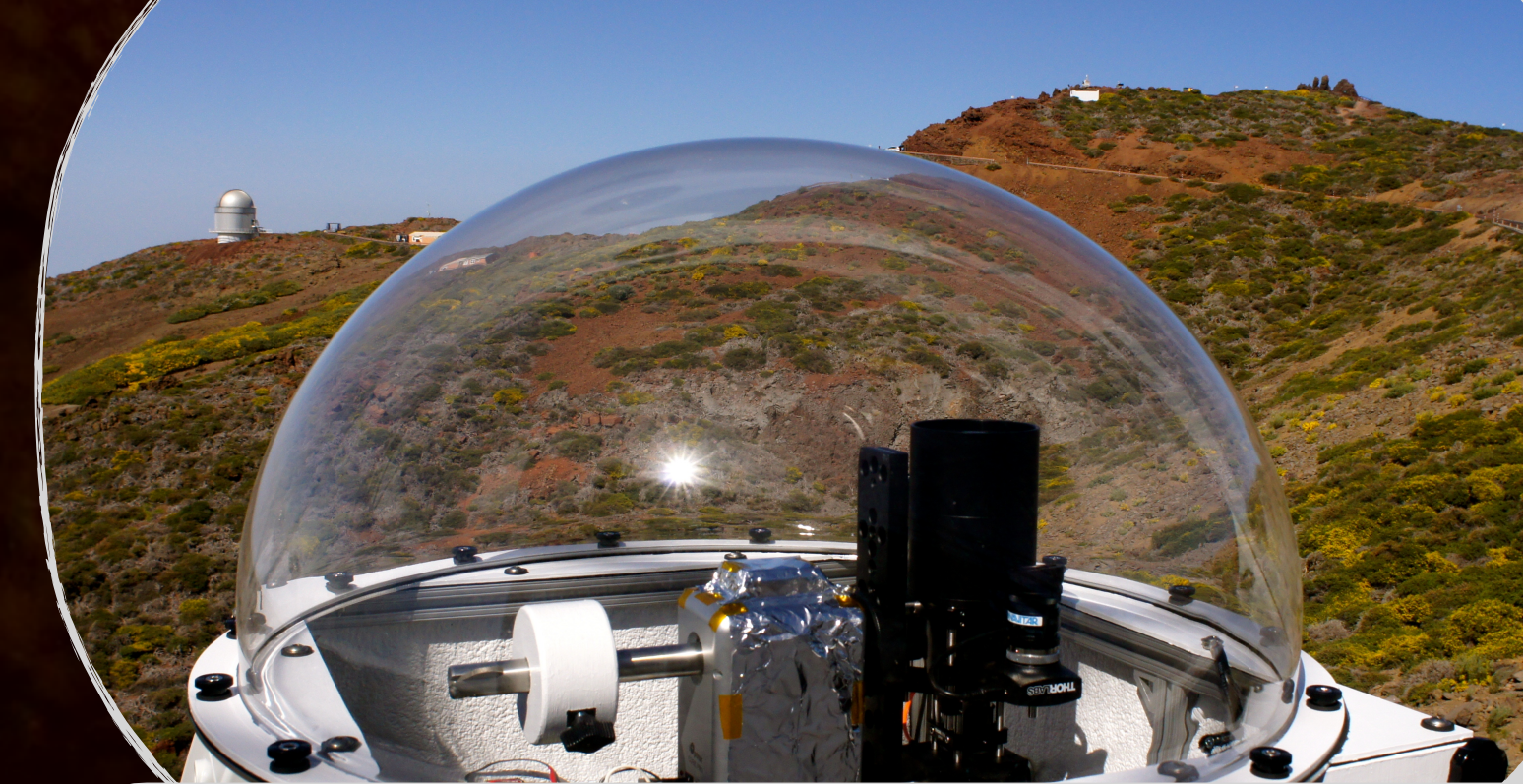


Need for high-fidelity spectra

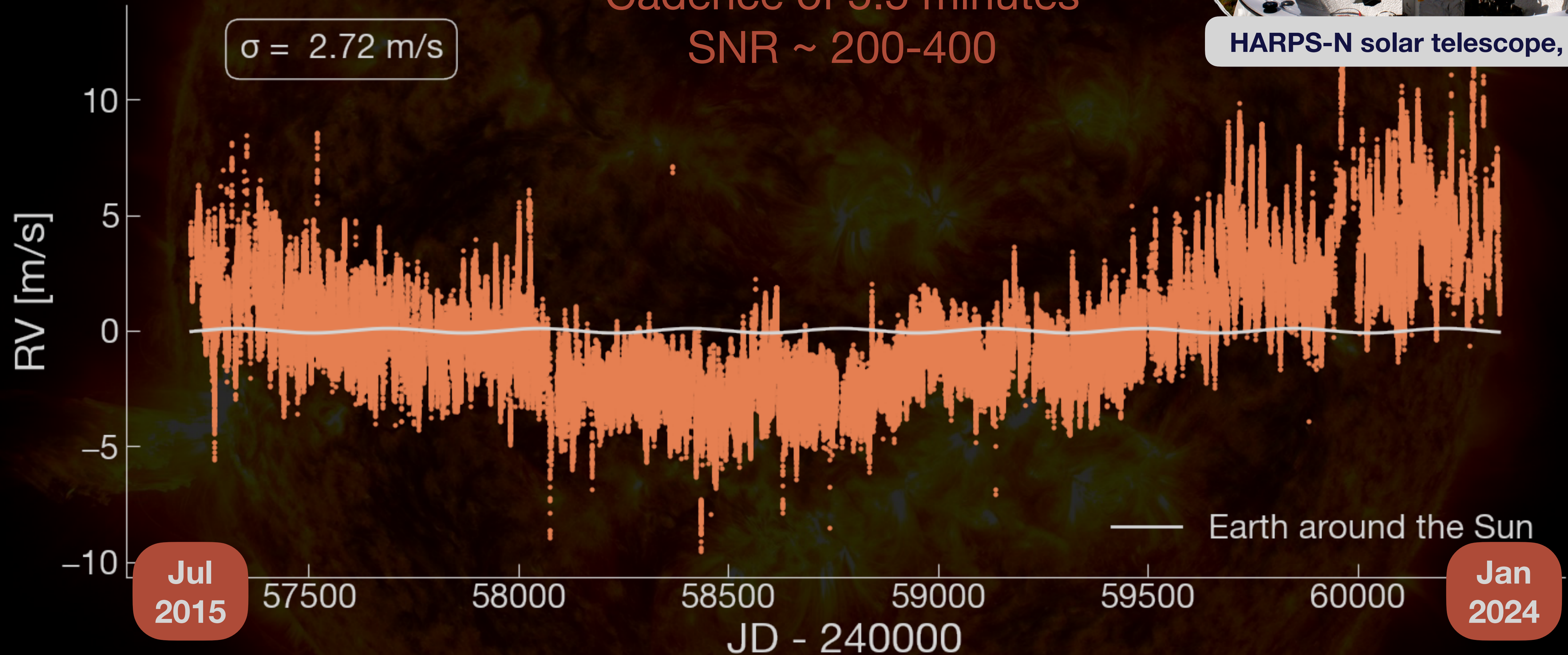
Correction for spurious signals must occur at the spectral level

8 Years of HARPS-N Solar Data

Observations occur daily
Cadence of 5.5 minutes
SNR ~ 200-400



HARPS-N solar telescope, La Palma



8 Years of HARPS-N Solar Data



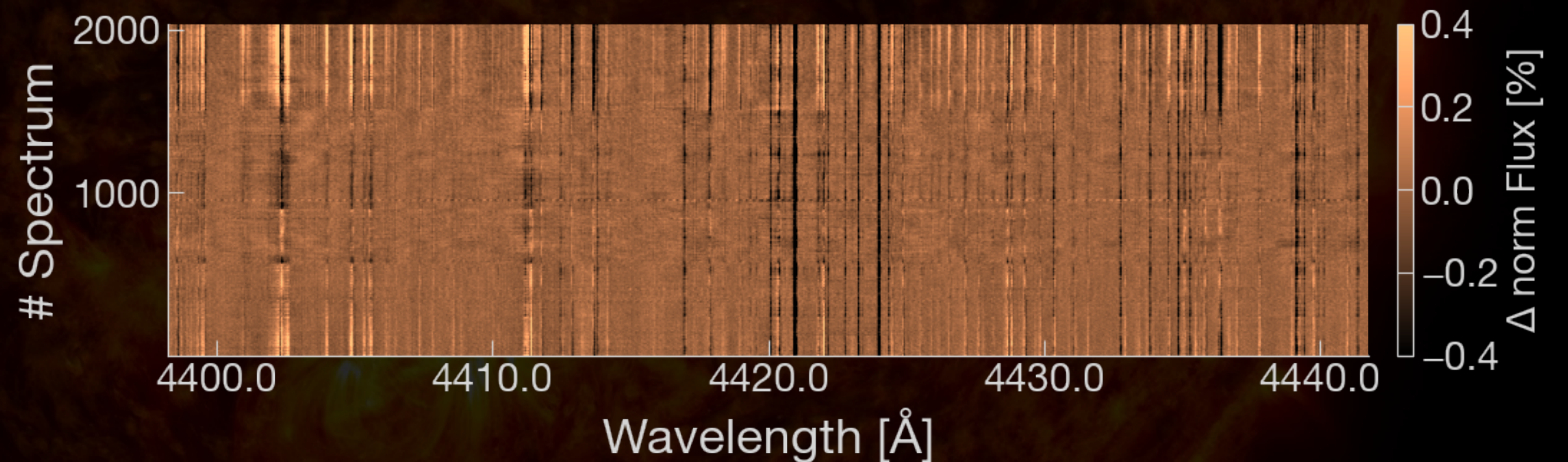
Daily binning + quality check → 2000+ spectra over 8 years

8 Years of HARPS-N Solar Data



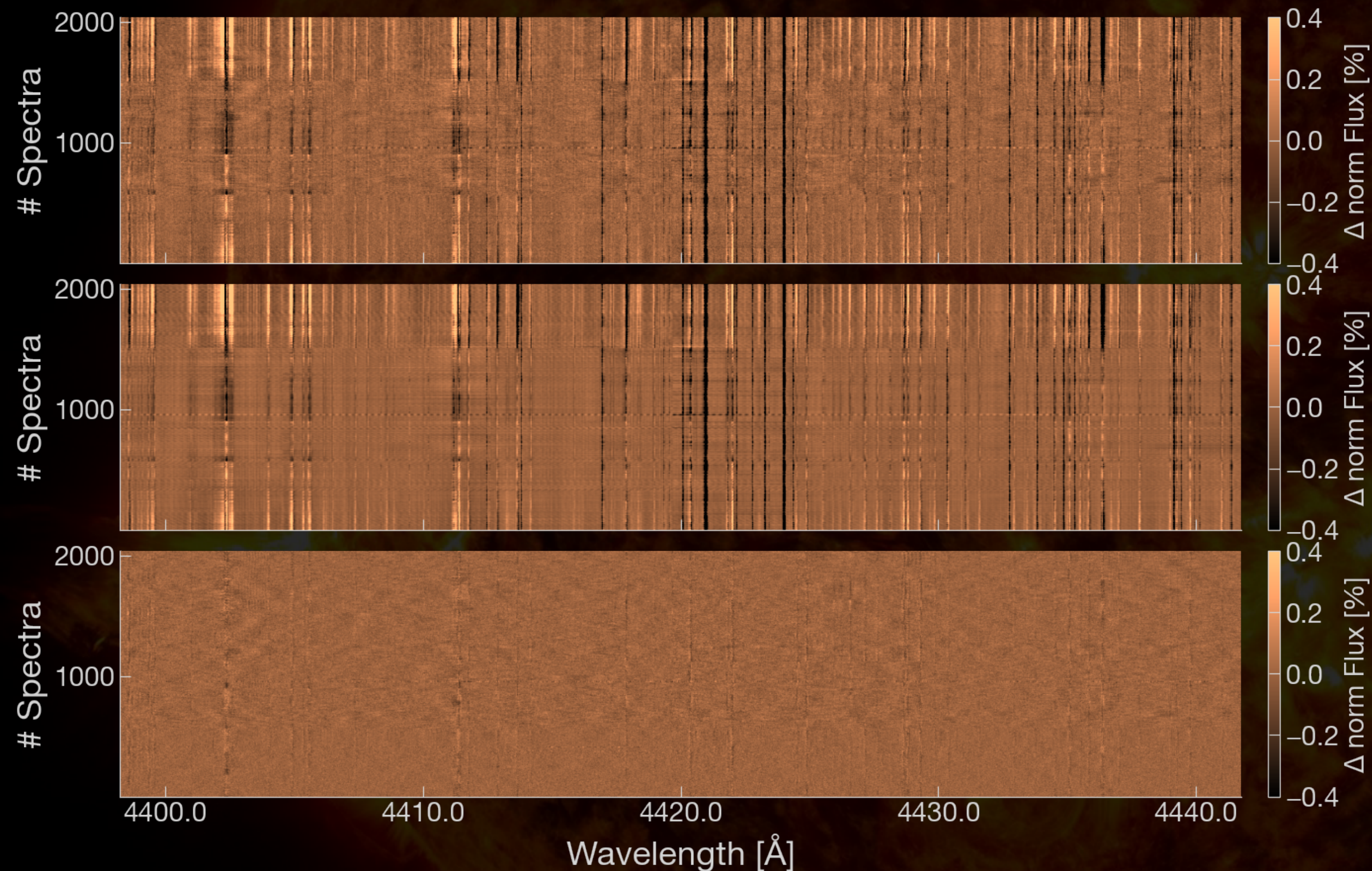
Daily binning + quality check $\rightarrow$ 2000+ spectra over 8 years

$$\frac{F_{obs} - F_{ref}}{F_{ref}}$$



Correction at The Spectral Level

YARARA: Tellurics, Cosmics, Instrument, Activity [...]



$$\frac{F_{obs} - F_{ref}}{F_{ref}}$$

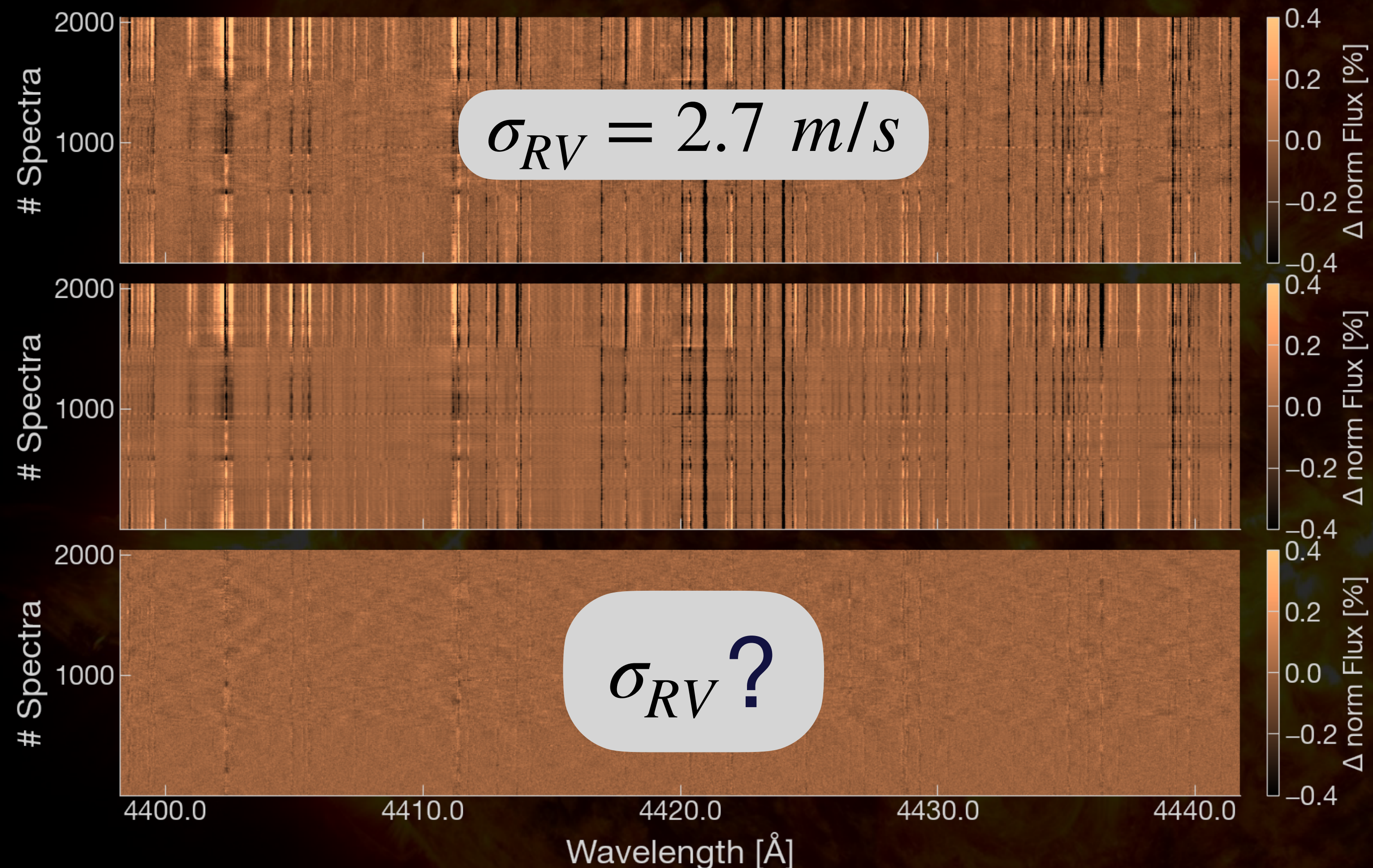
Correction mask

$$\frac{F_{corr} - F_{ref}}{F_{ref}}$$

Cretignier+2021

Correction at The Spectral Level

YARARA: Tellurics, Cosmics, Instrument, Activity [...]



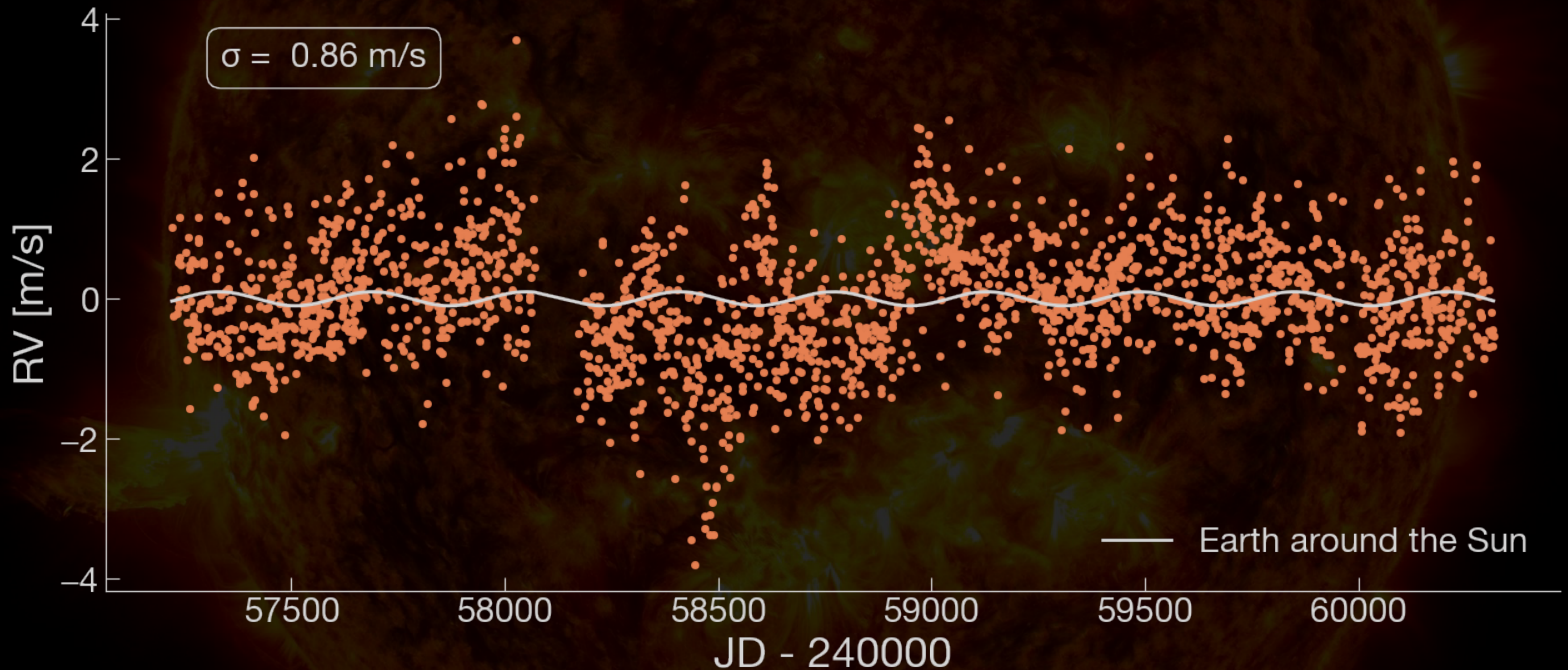
$$\frac{F_{obs} - F_{ref}}{F_{ref}}$$

Correction mask

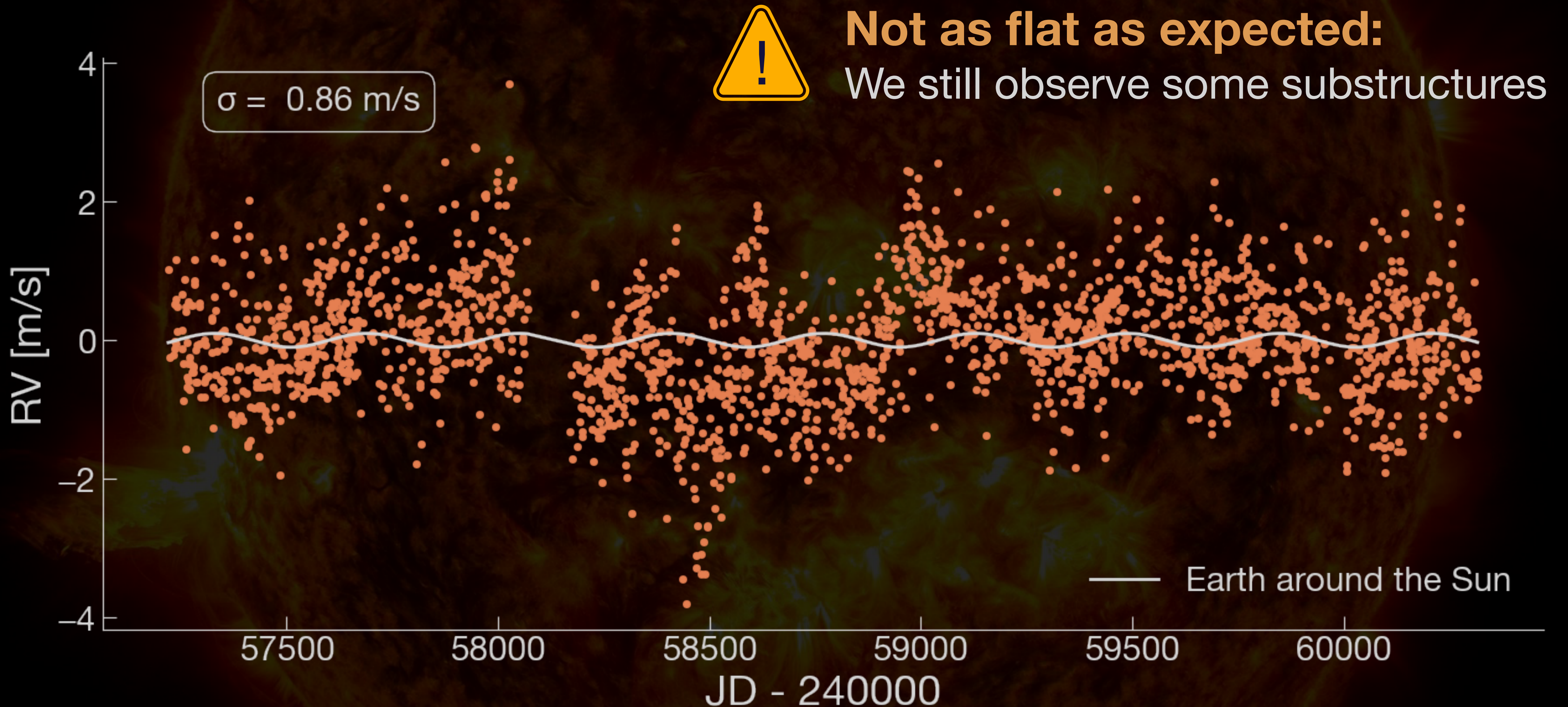
$$\frac{F_{corr} - F_{ref}}{F_{ref}}$$

Cretignier+2021

Radial Velocities After YARARA v1



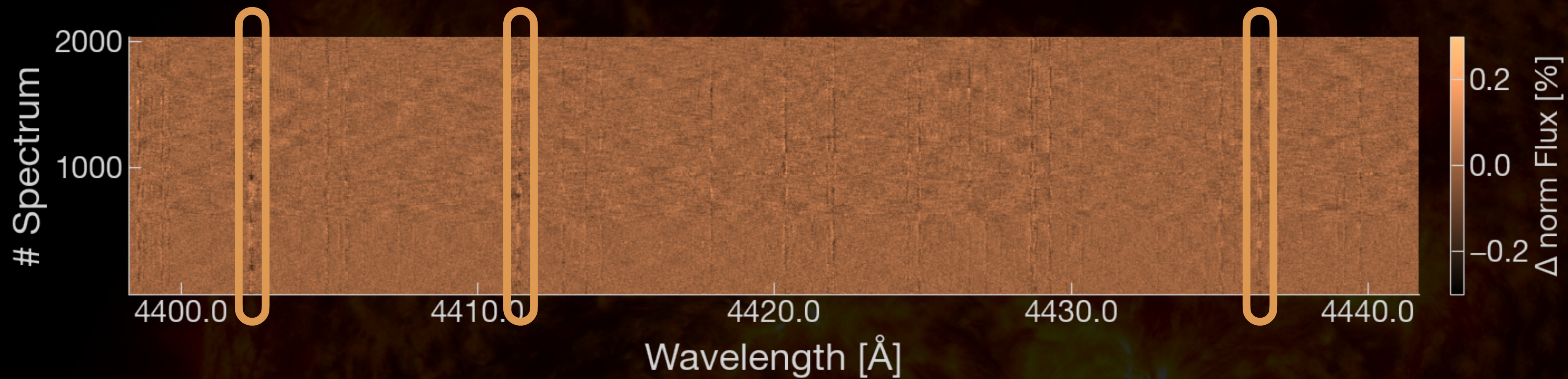
Radial Velocities After YARARA v1



Let's Look at the Spectral Time Series



Let's Look at the Spectral Time Series



Periodograms of the spectral time series:

Sun rotation

Clearly stellar origin

Yearly signal

Likely instrumental origin + tellurics

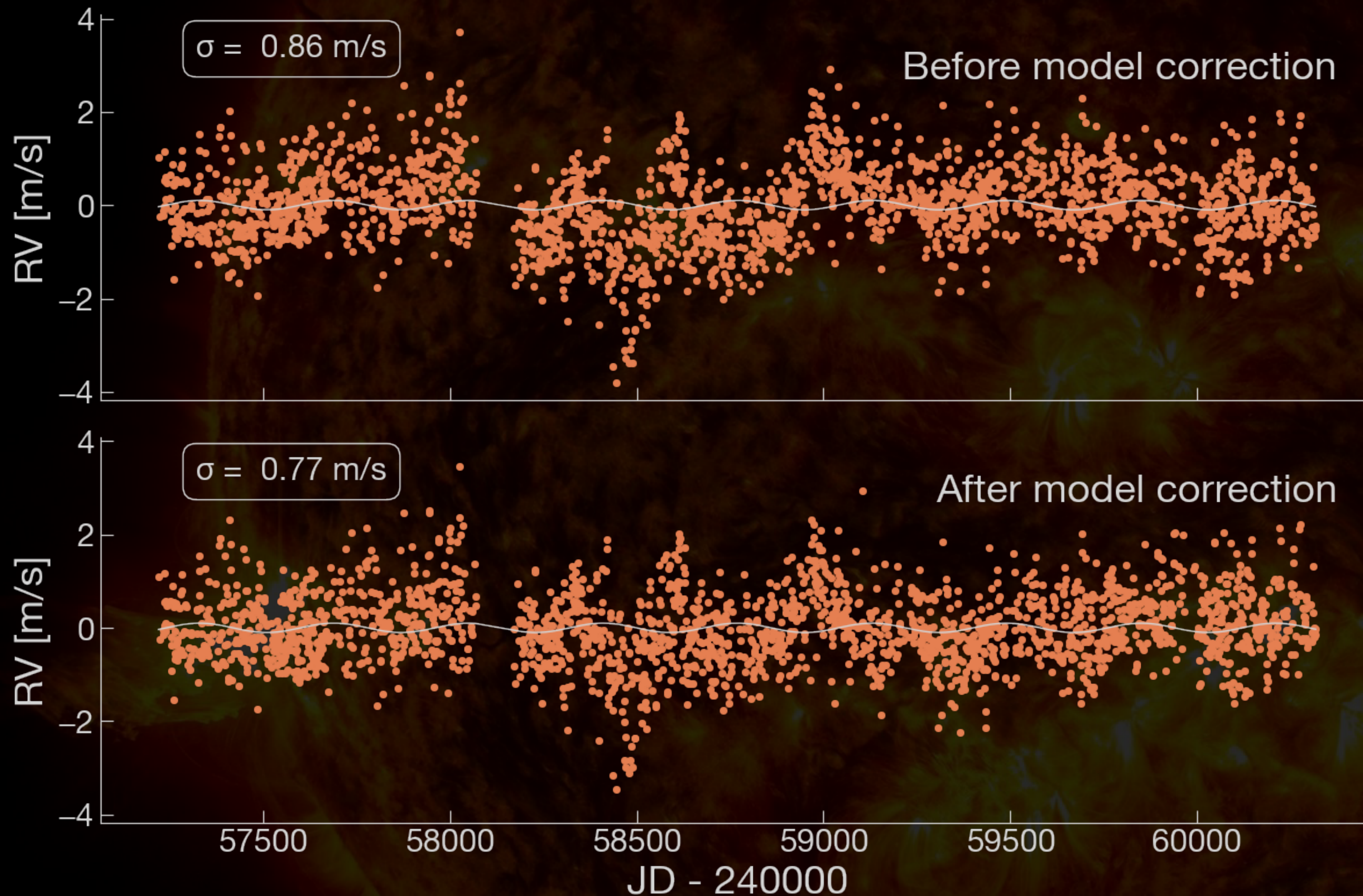
Long term trend

Instrumental and stellar origin



**Some significant signals
persist after YARARA**

Preliminary Linear Detrending



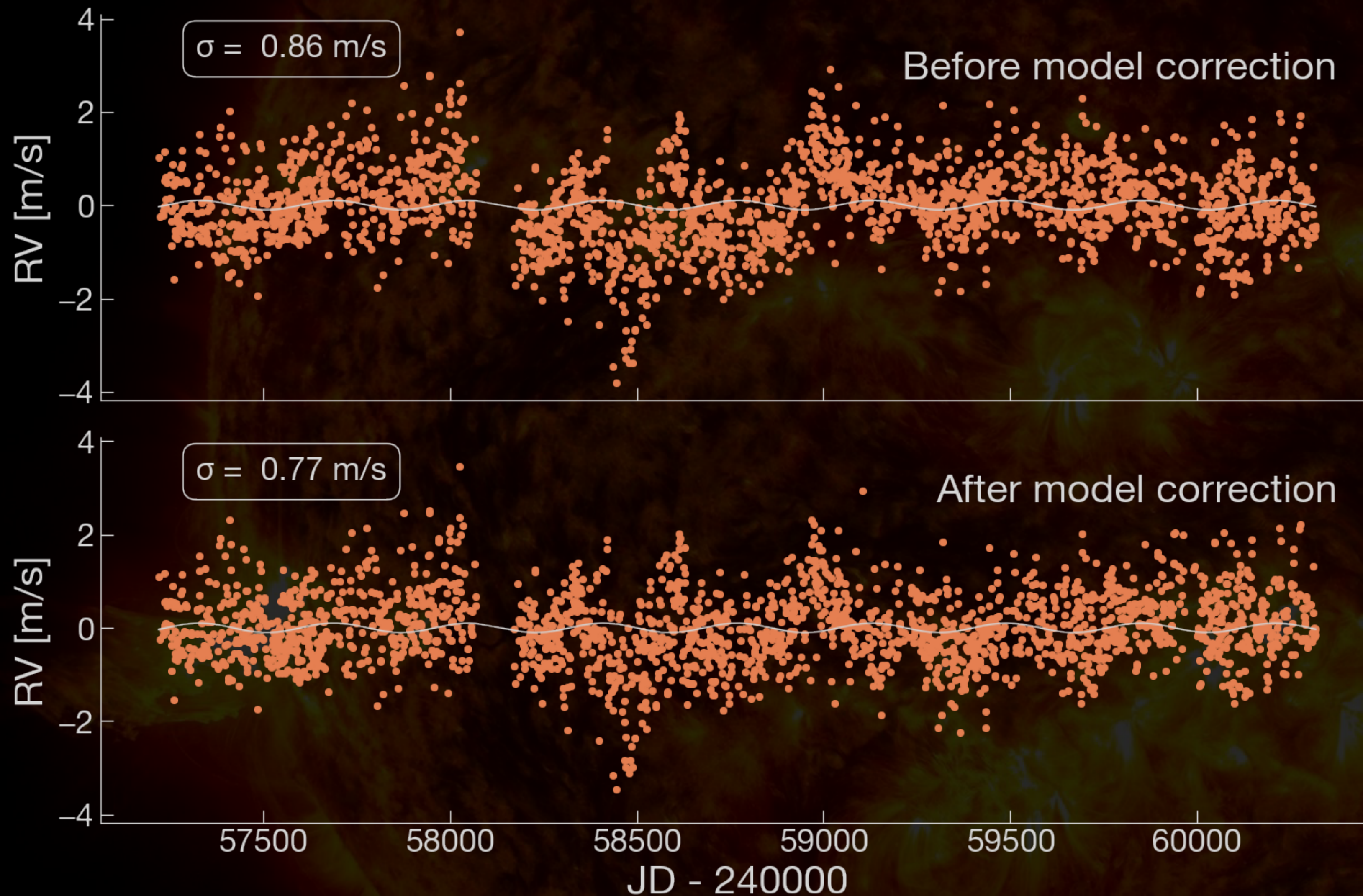
Linear model

Long term trend

Sun rotation

Yearly signal

Preliminary Linear Detrending



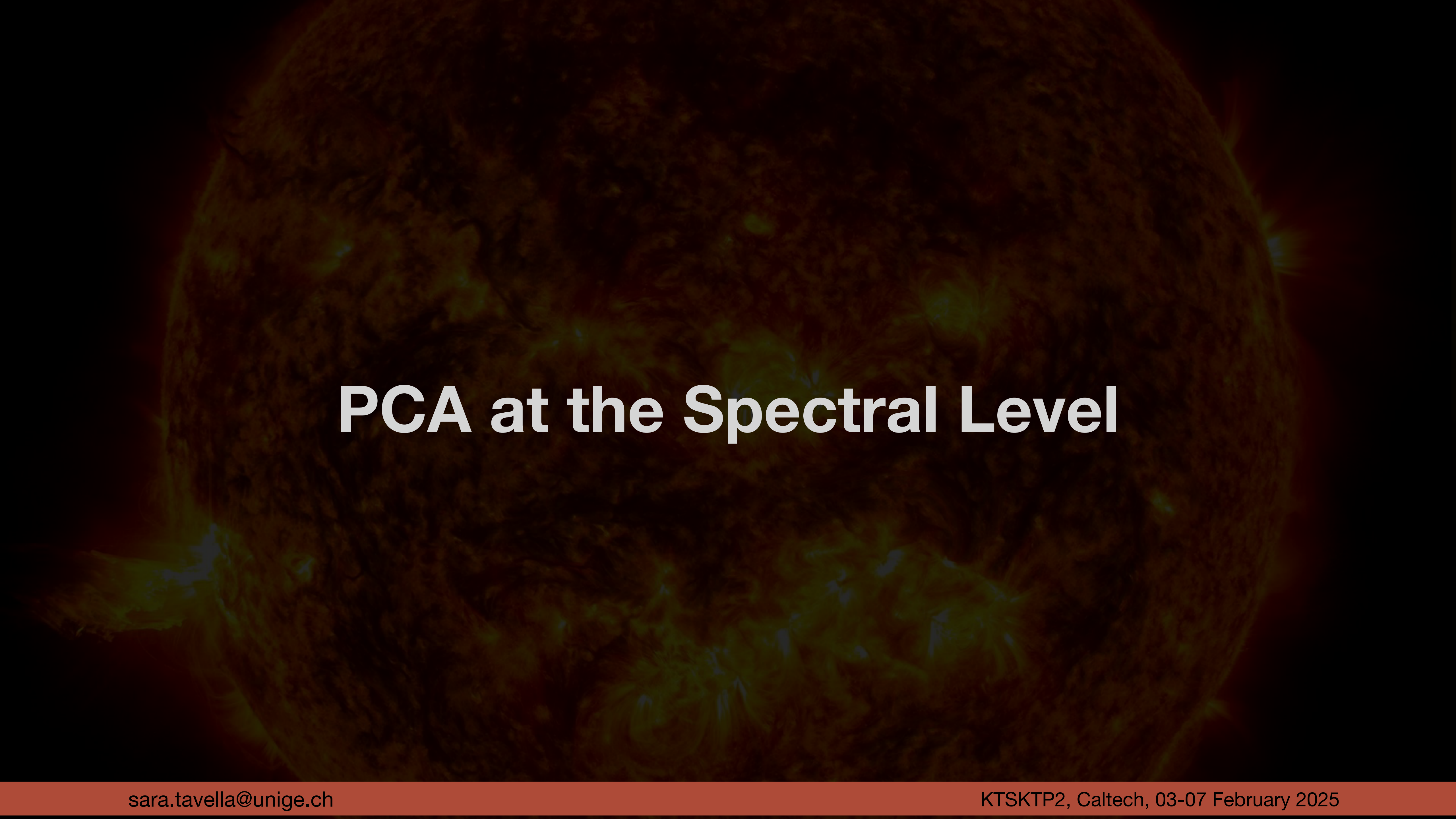
Linear model

Long term trend

Sun rotation

Yearly signal

**10% improvement
in the RV rms**



PCA at the Spectral Level

1. Reduce Dimensionality



1. Reduce Dimensionality



“Forget” about the Physics

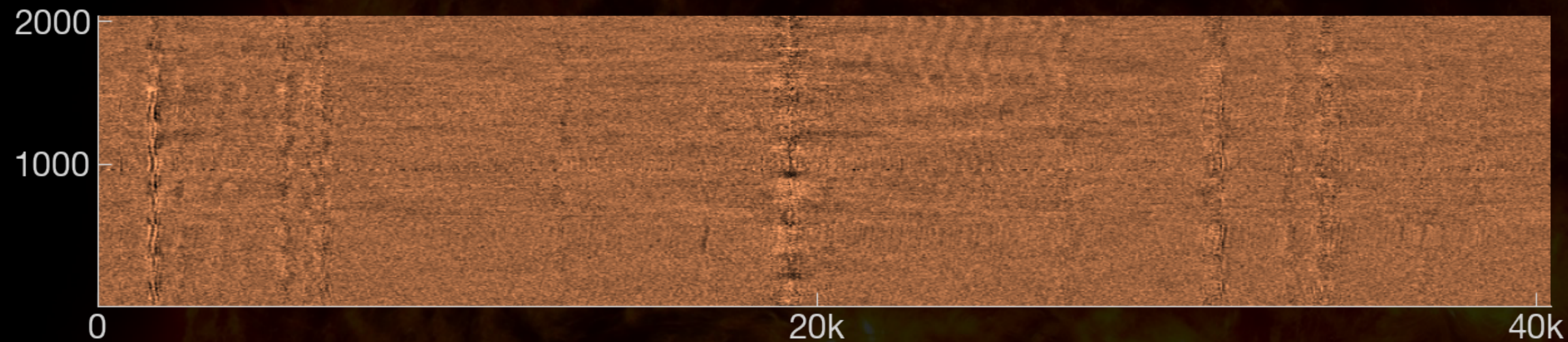
1. Reduce Dimensionality



“Forget” about the Physics

Keep only columns showing significant signals (from periodogram analysis)

2. Run PCA

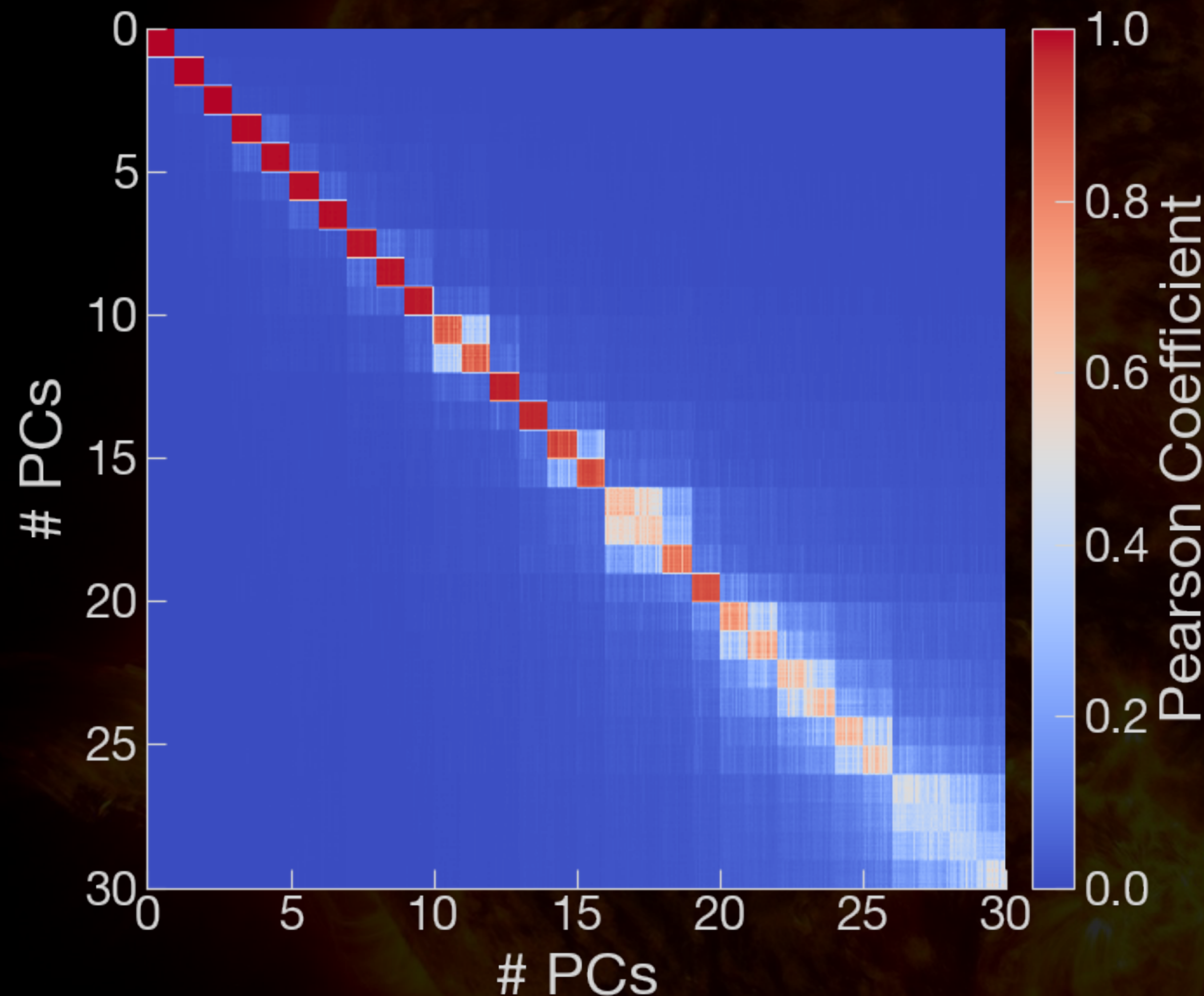


PCA matrix: 2k x 40k pixels

Principal components (PCs): 1 x 40k

N.B. We are mainly interested in the time coefficients of the PCs

3. Validate the First Principal Components



Leave-percentage-out cross-validation

20% of the data removed

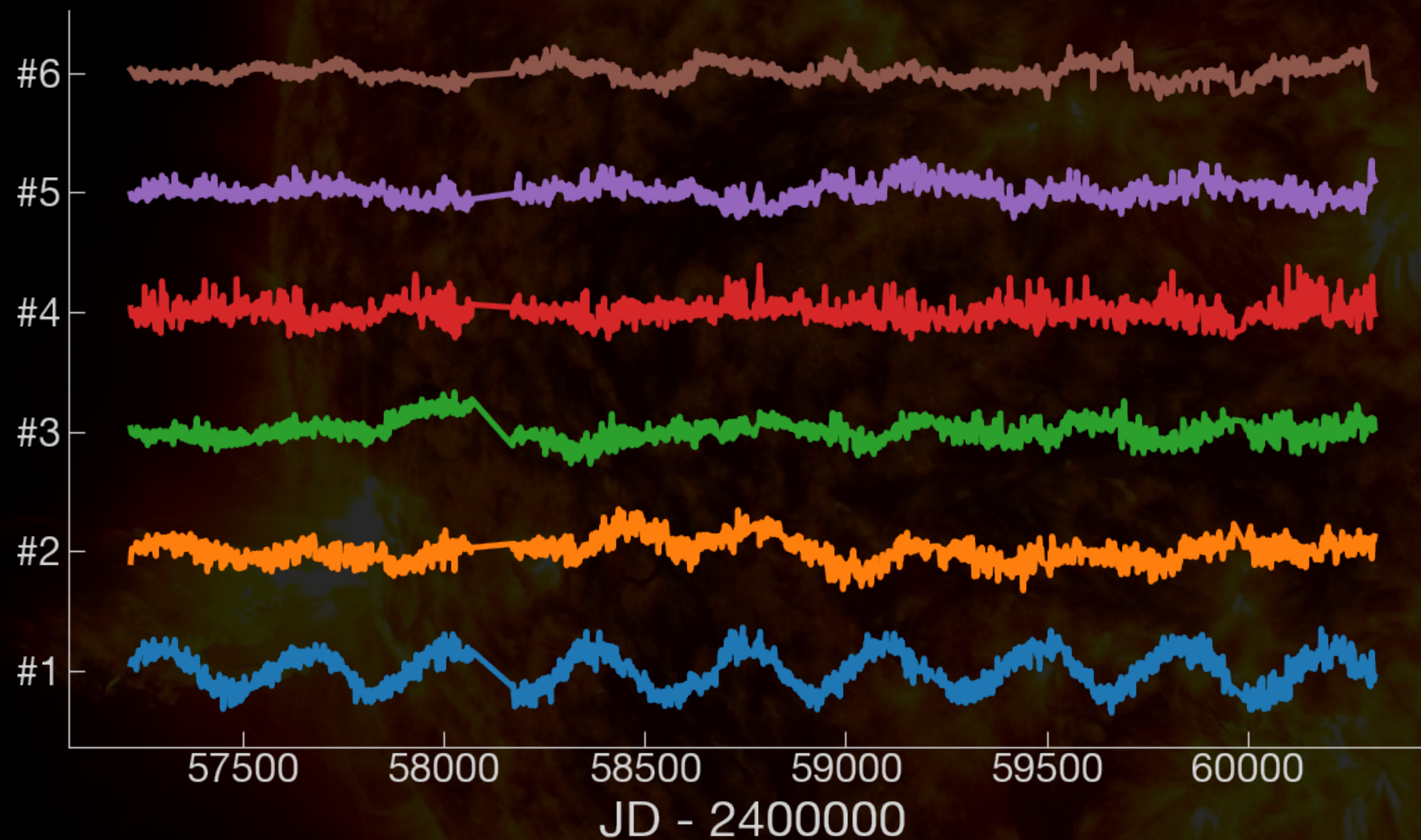
10000 iterations

Significance threshold: $R = 0.95$

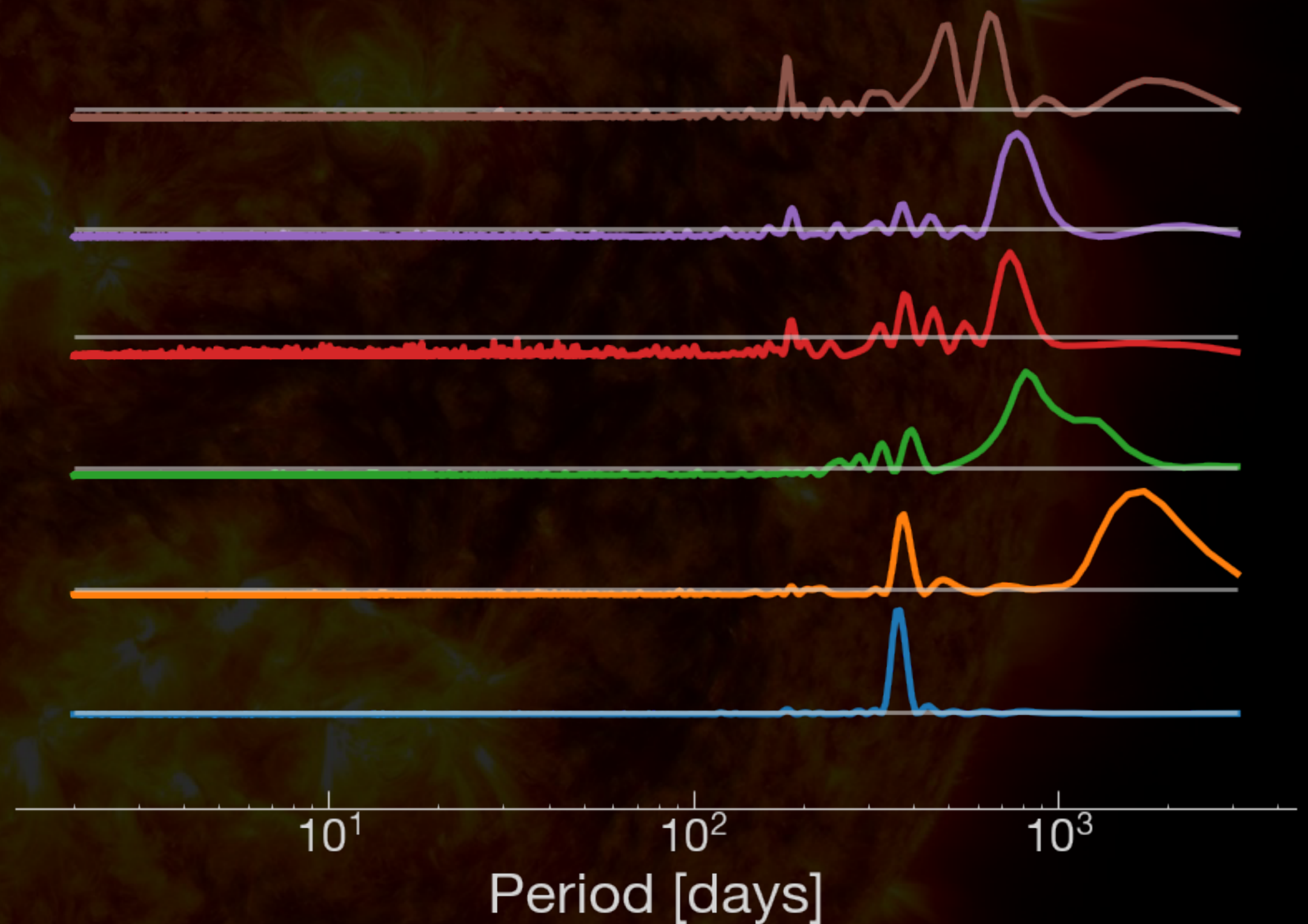
**16 significant
principal components**

3. Validate the First Principal Components

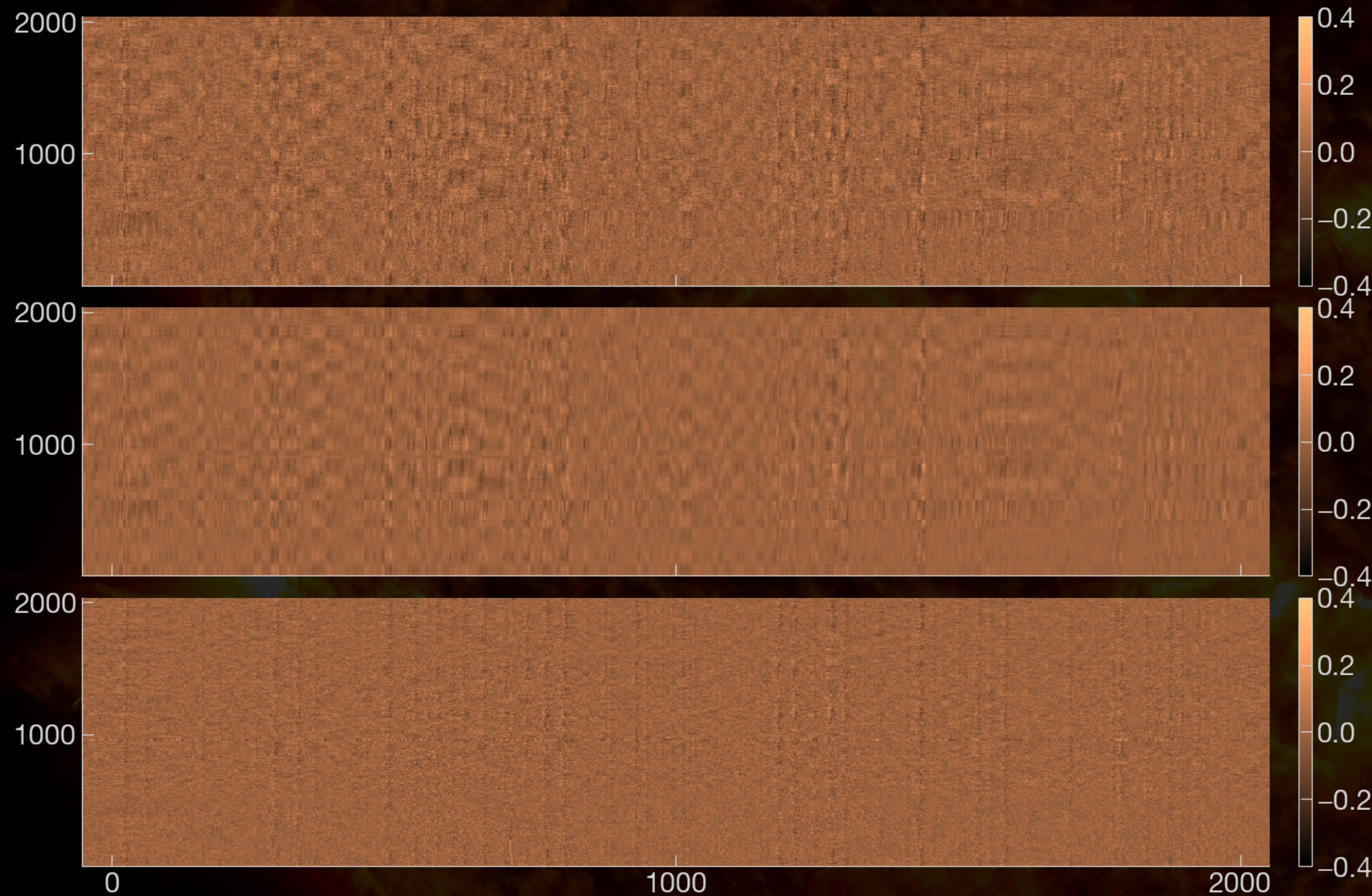
Principal components



Periodograms



4. Correct for the PCA Model

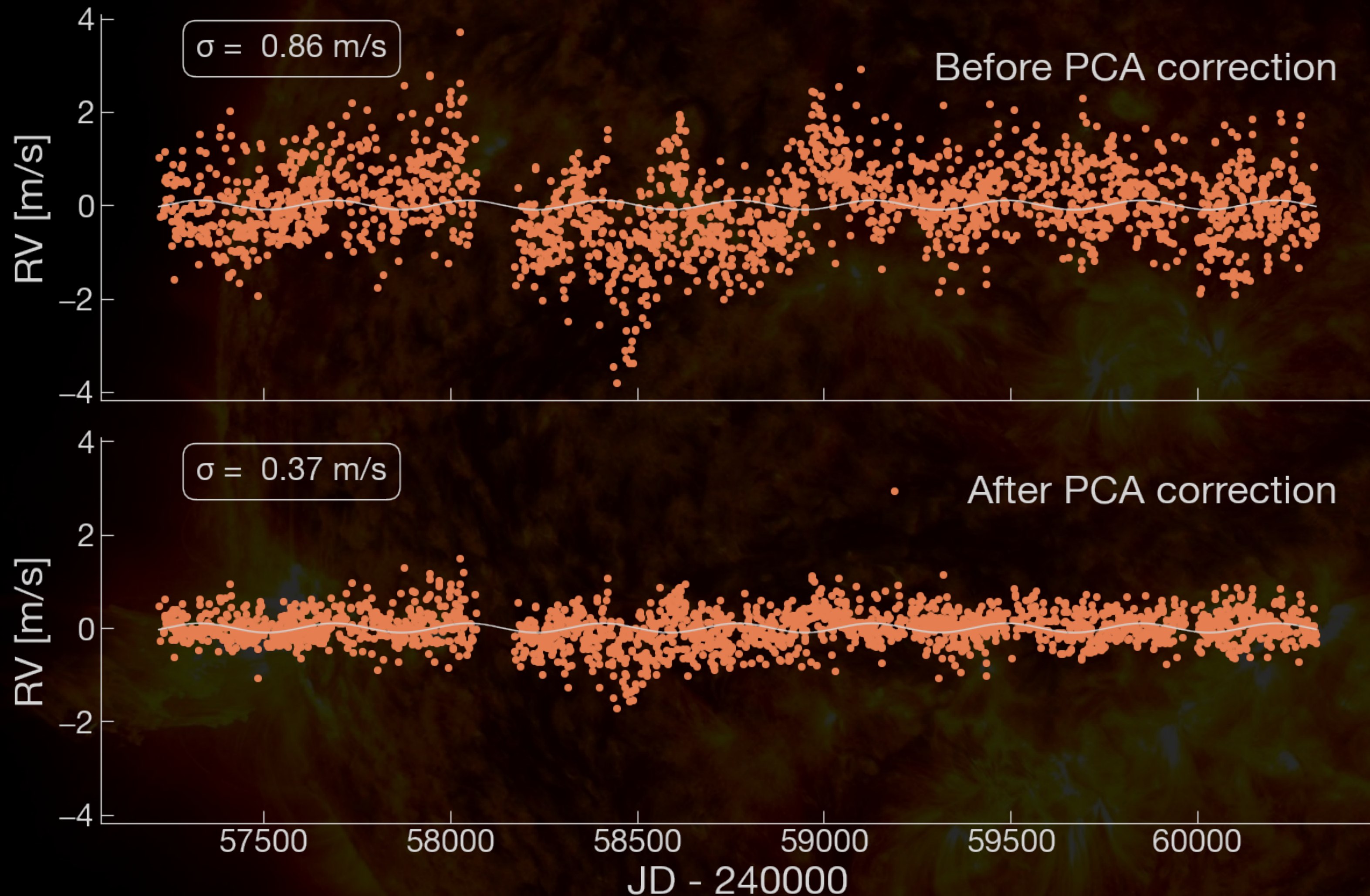


PCA matrix
before correction

PCA model

PCA matrix
after correction

Result #1. Improvement in RV Precision



**~60% improvement
in the RV rms**

Result #2. Improvement in Detection Limits

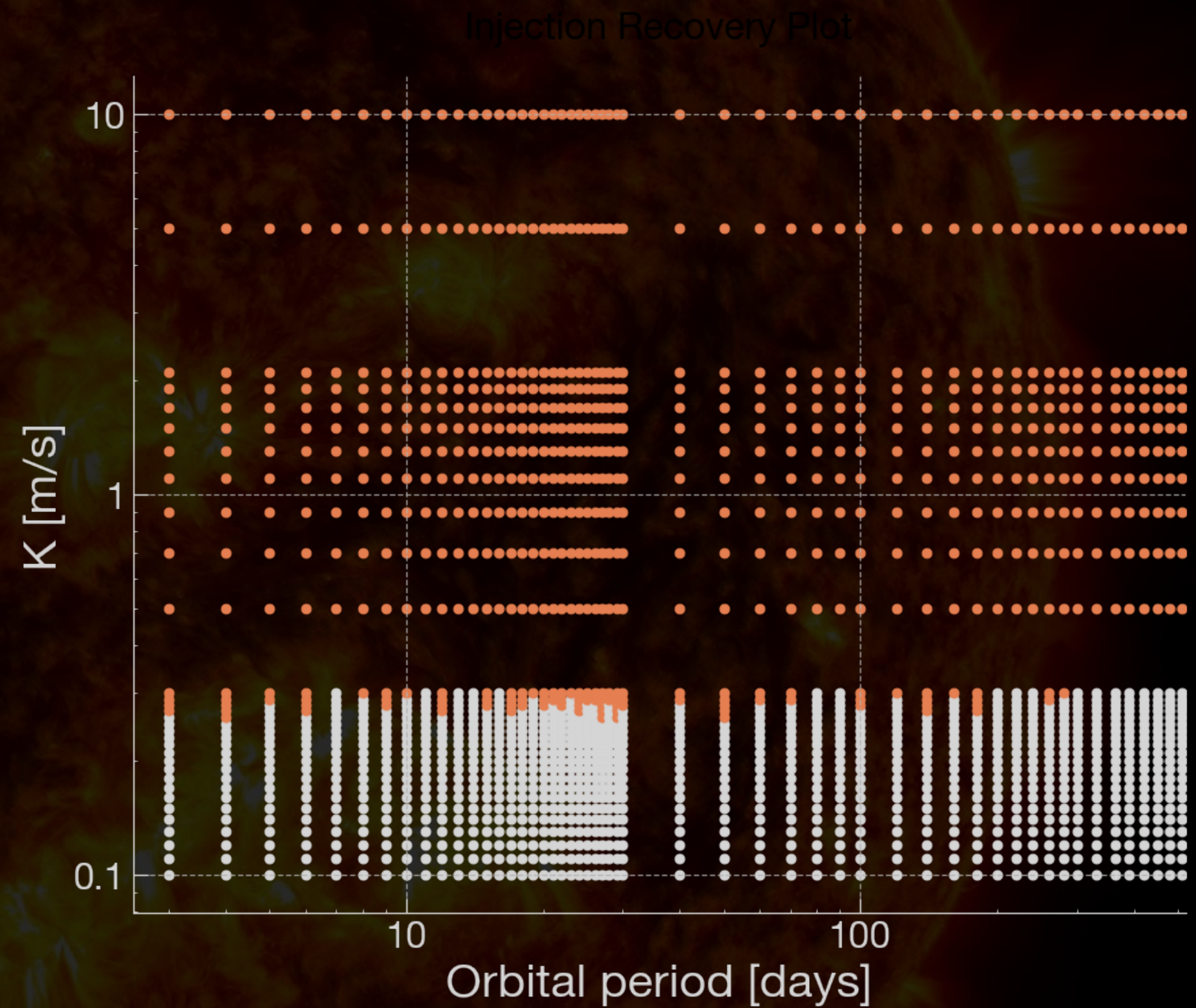
Injection-recovery test

Grid of planetary signals:

Period from 1 day to 510 days

Amplitude from 10 cm/s to 10 m/s

- Planets injected at the spectral level



Result #2. Improvement in Detection Limits

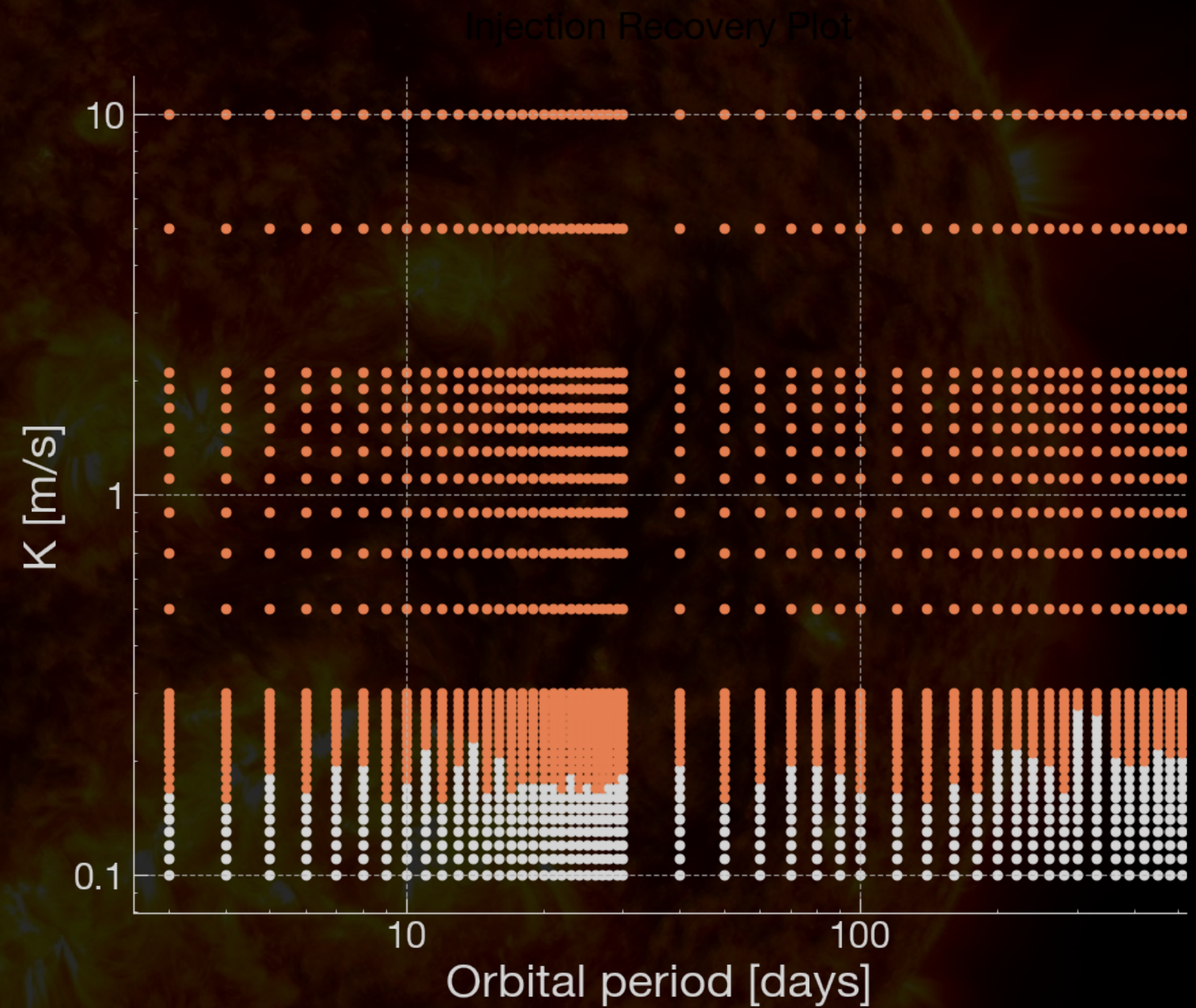
Injection-recovery test

Grid of planetary signals:

Period from 1 day to 510 days

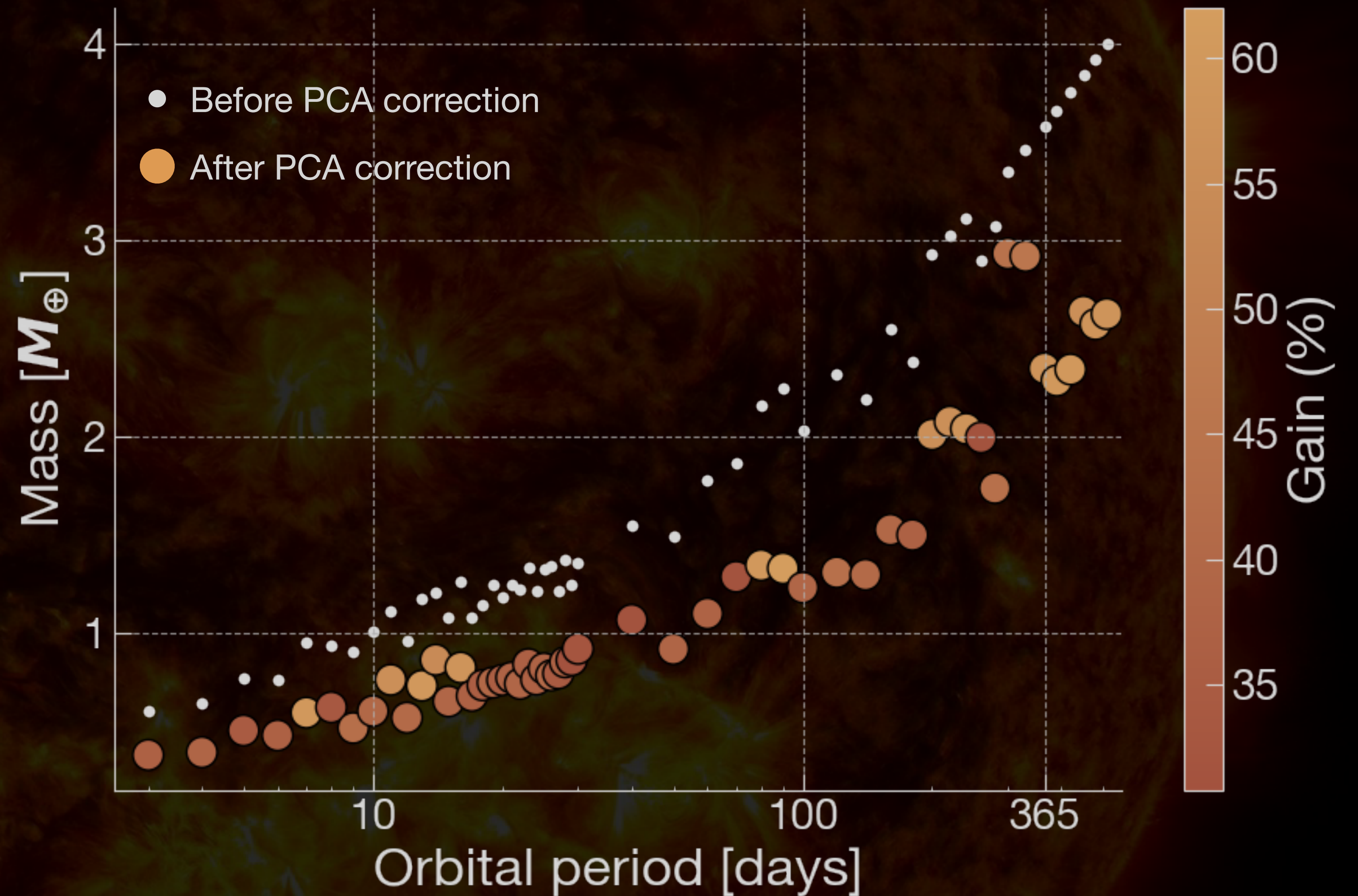
Amplitude from 10 cm/s to 10 m/s

- Planets injected at the spectral level
- PCA run again for every injection



Result #2. Improvement in Detection Limits

**30% to 60%
improvement
in detection limit**



Take-Home Messages

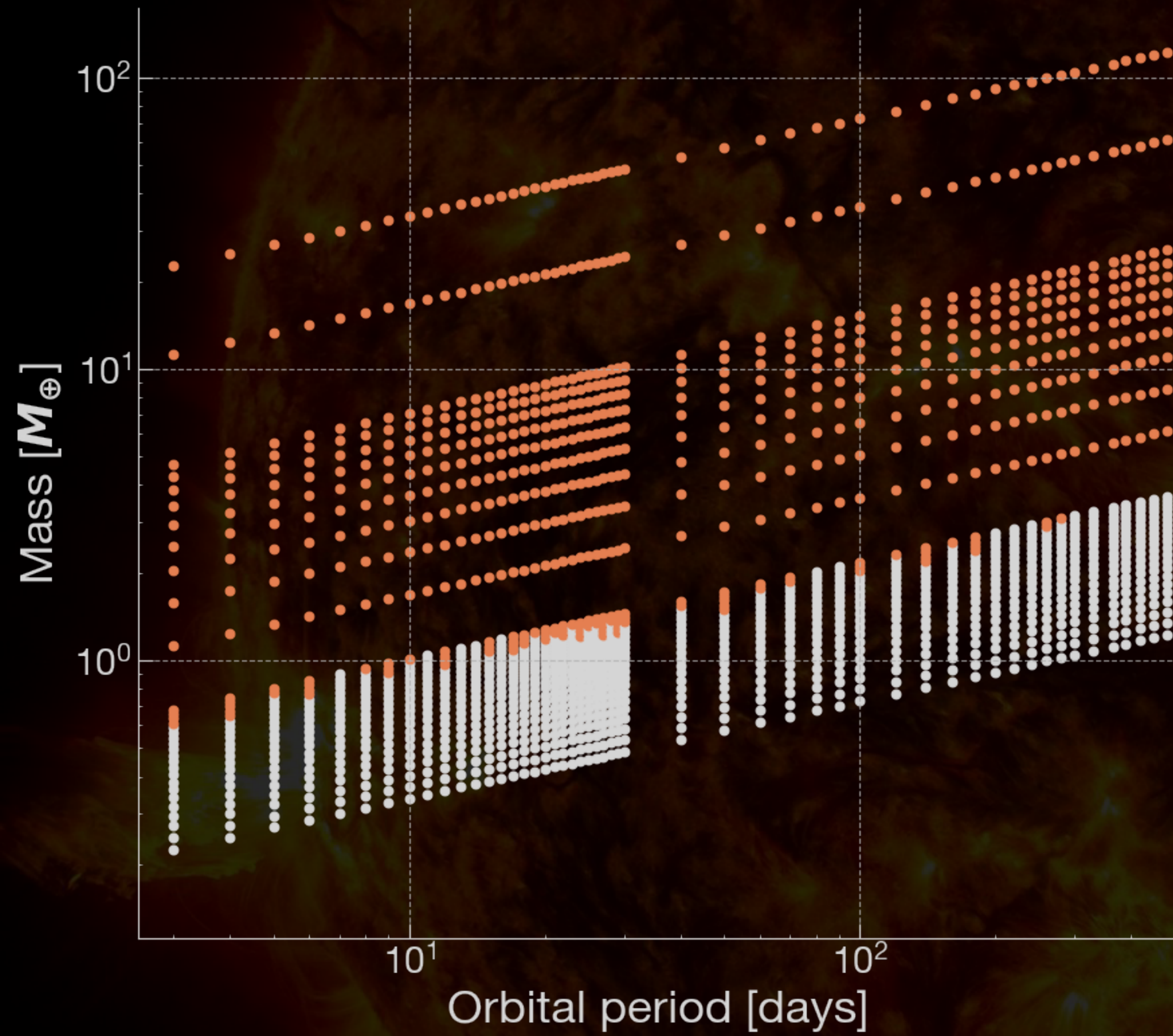
We work at the spectral level to catch wavelength-dependent features

We want to correct for residual signals of both instrumental and stellar origin

- Data drive correction of the spectral time series:
The PCA model mainly captures long term and yearly signals
!! RV precision of 8 years of HARPS-N solar data down to 40 cm/s
- Injection-recovery test to validate the PCA correction:
Reproduces by construction a real case of exoplanet detection
Shows that the PCA does not absorb the planetary signal
!! ~40% improvement of detection limits (~2.5 Earth Masses @365d)

BACKUPS

Injection Recovery Plot



Injection Recovery Plot

