Automated data reduction pipelines for the HPF and NEID spectrometers

Chad Bender
University of Arizona

March 21, 2019
EPRV IV: Grindelwald

The HPF/NEID Software Team:
Joe Ninan – Penn State
Kyle Kaplan – UofA
Arpita Roy – Caltech
Gudmunder Steffanson – Penn State
Shubham Kanodia – Penn State

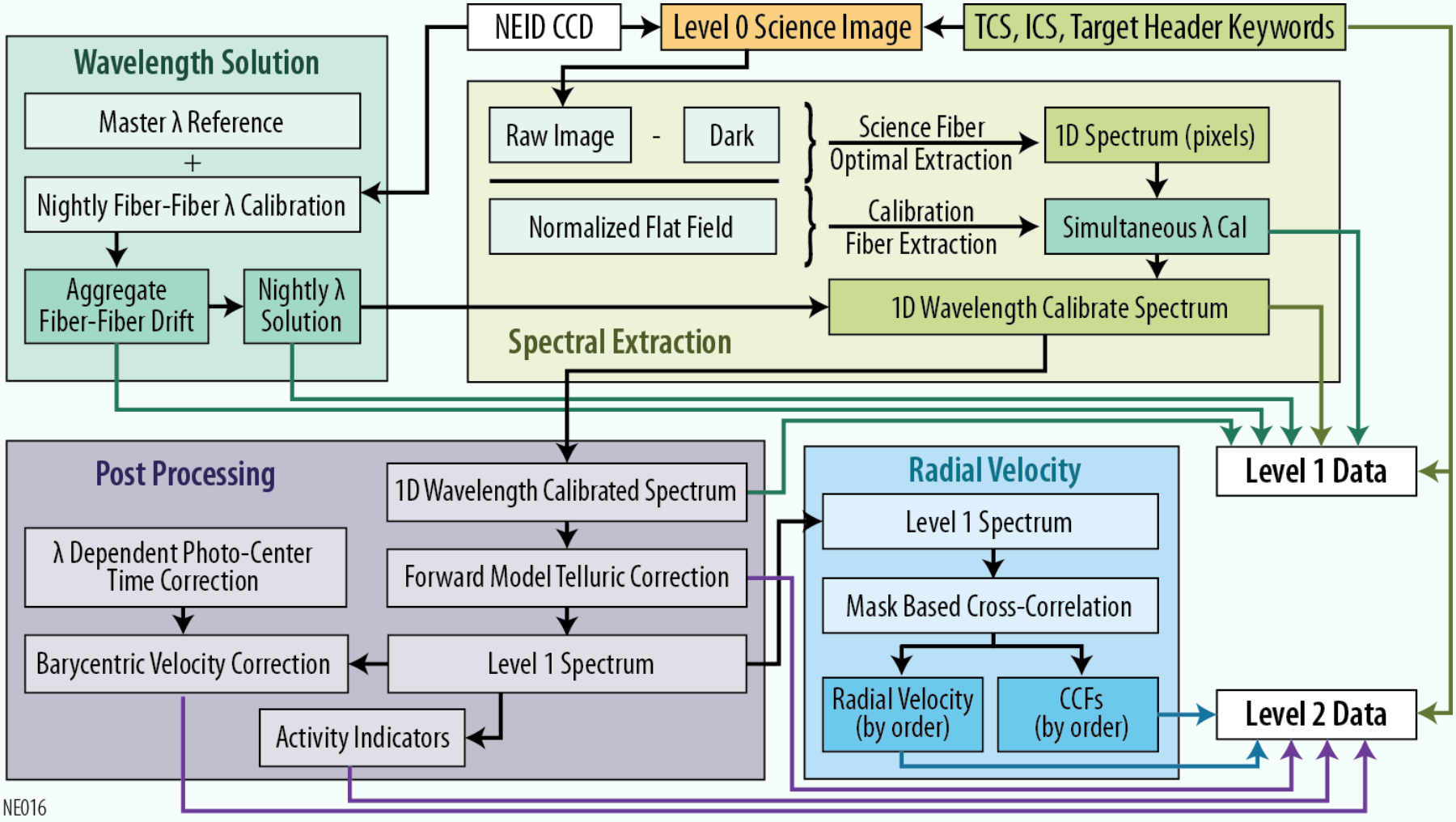
The NExScI Team:

Hardware is hard (else ‘easyware’?), but …

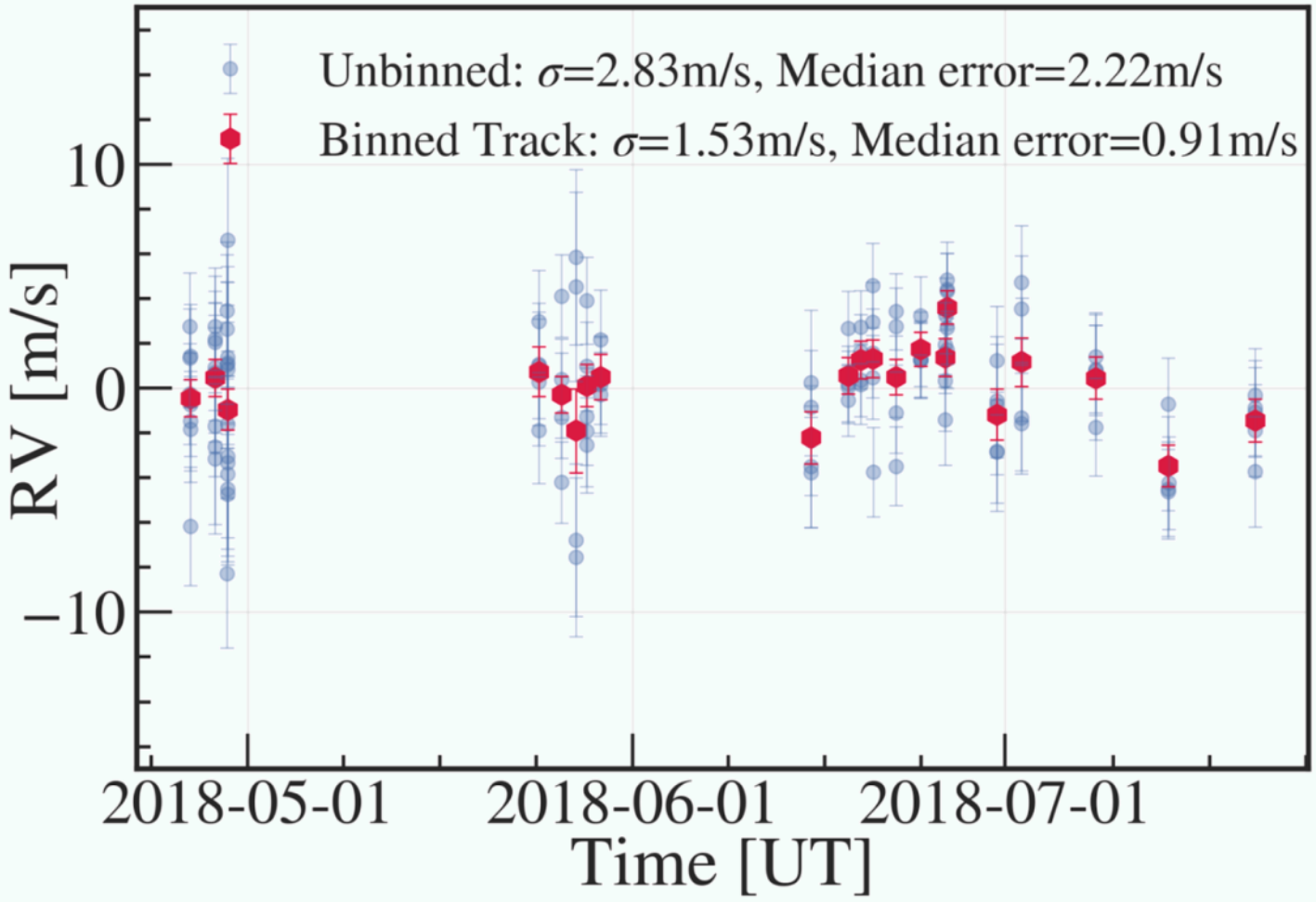
Fix everything you can in hardware…

The software team will thank you

NEID Pipeline Schematic

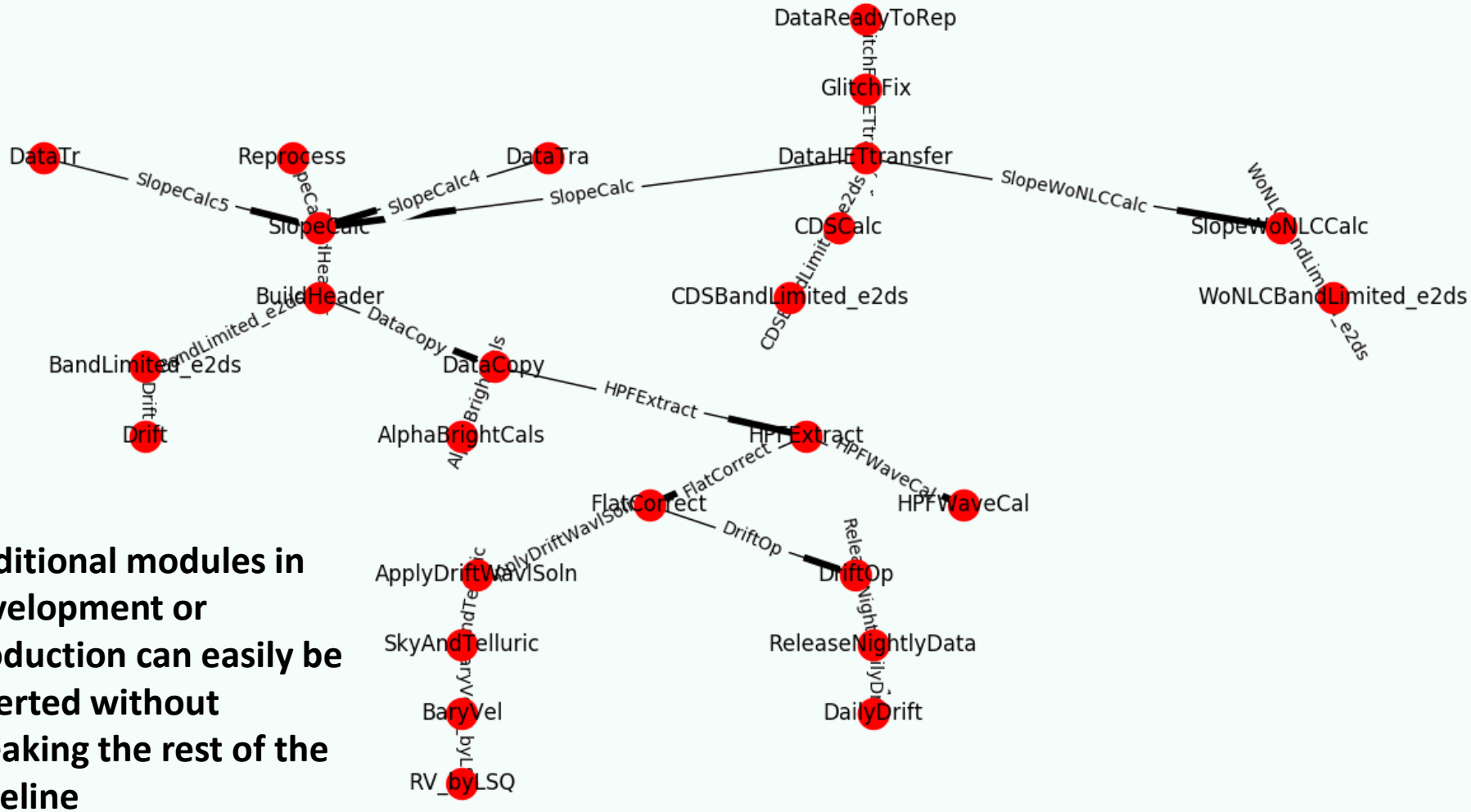


The HPF Pipeline is routinely producing velocities



HPF Directed Acyclic Graph (DAG)

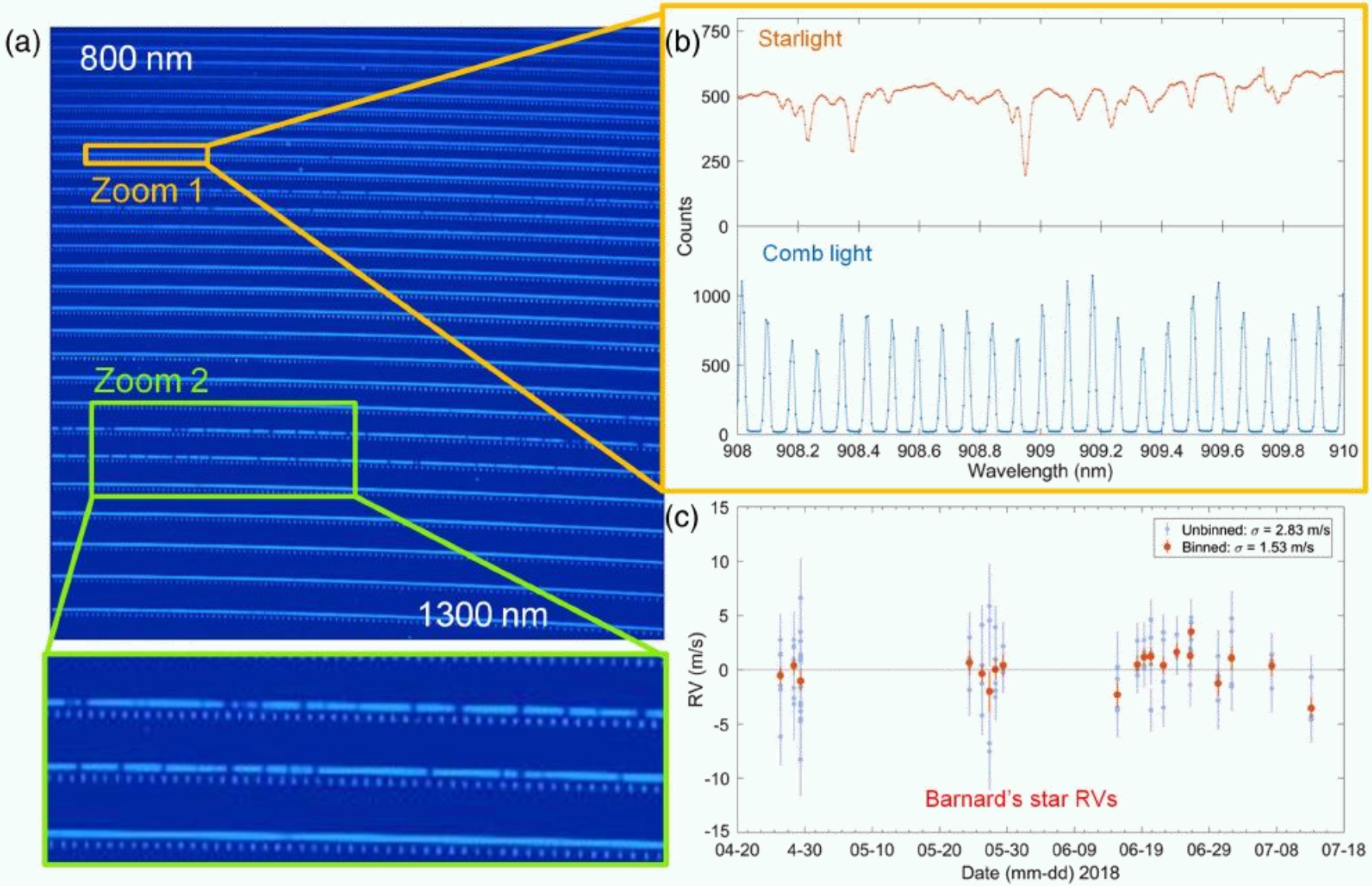
What this looks like in practice for HPF:



Additional modules in development or production can easily be inserted without breaking the rest of the pipeline

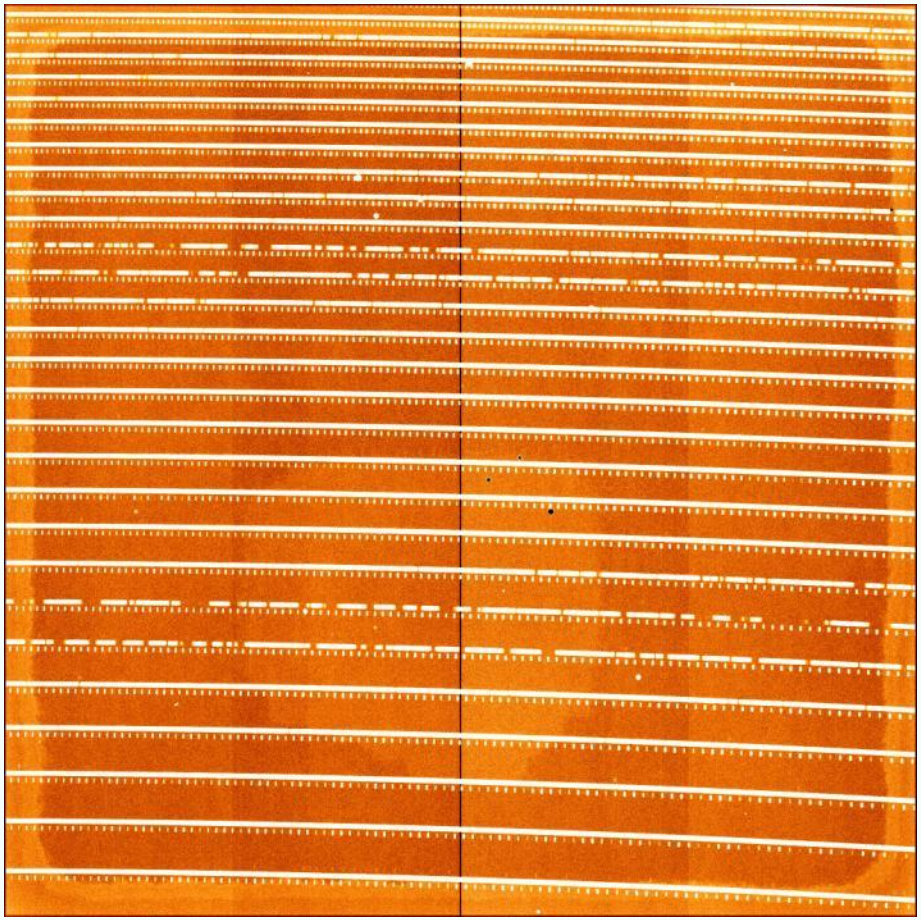
We call this the APCP – Autonomous Pipeline Control Program

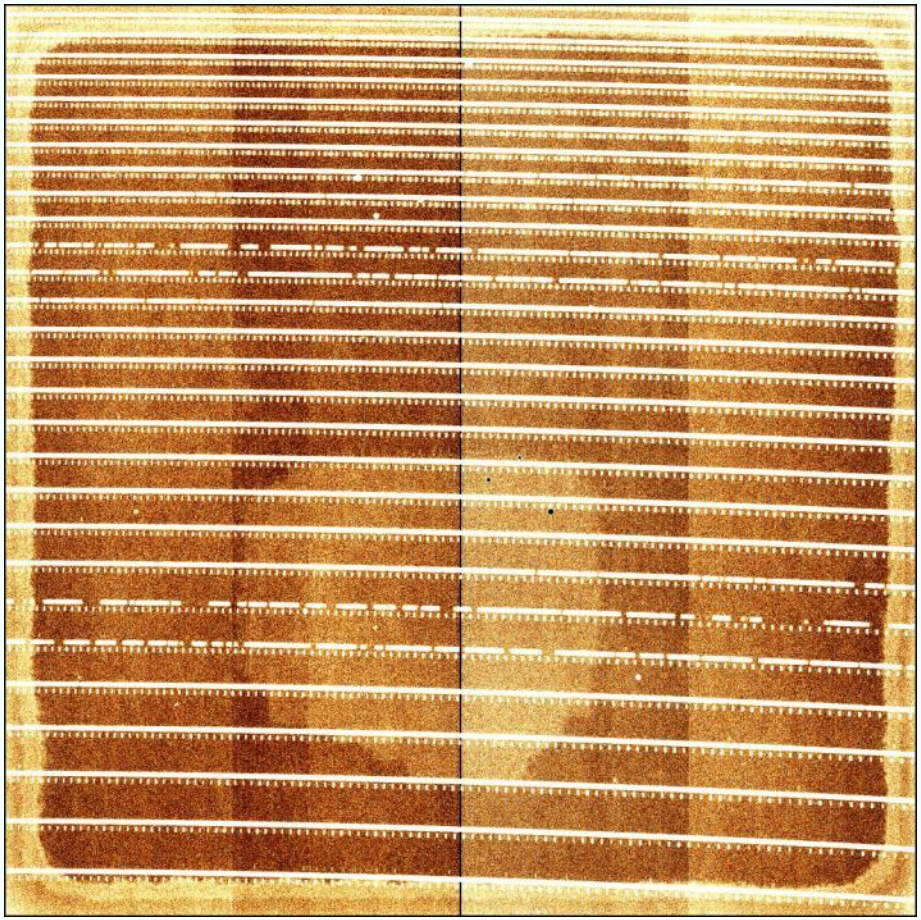
A processed HPF image and spectrum



A sample HPF image of GJ699

Two different stretches of a single HPF up-the-ramp frame



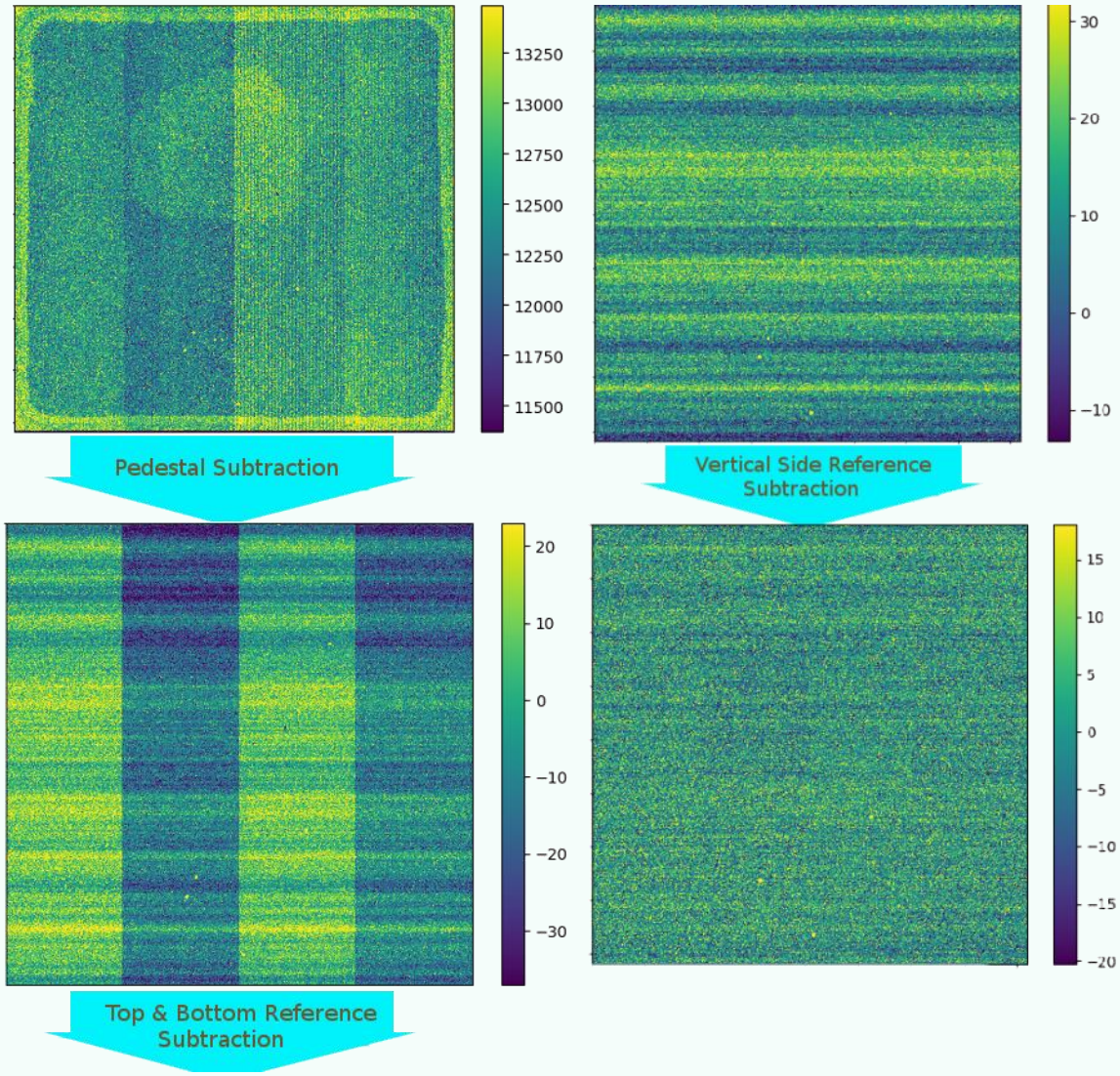


Deeper Dives into a few modules

- H2RG Image Processing
- Spectral Extraction
- Barycentric Correction
- RV Measurements of Late-Type Stars

H2RG Basic Bias Correction

A standard reference pixel based bias correction leaves structured noise in the image

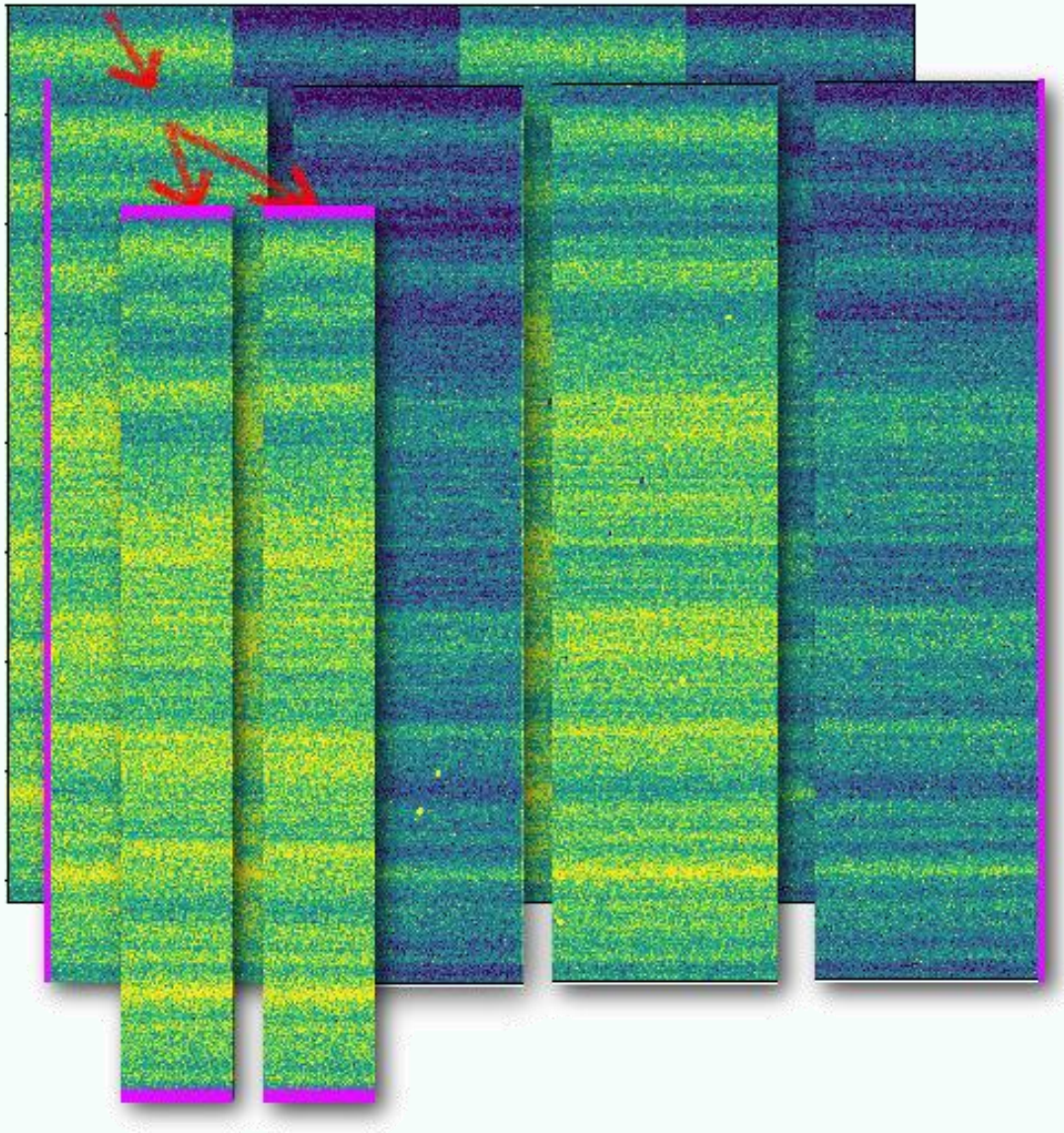


Modified clocking schemes can interleave additional reference pixels:

IRS^2 in H2RG (Roucher et. al)

Interleaved ref pixel (H4RG)

H2RG Better Bias Correction

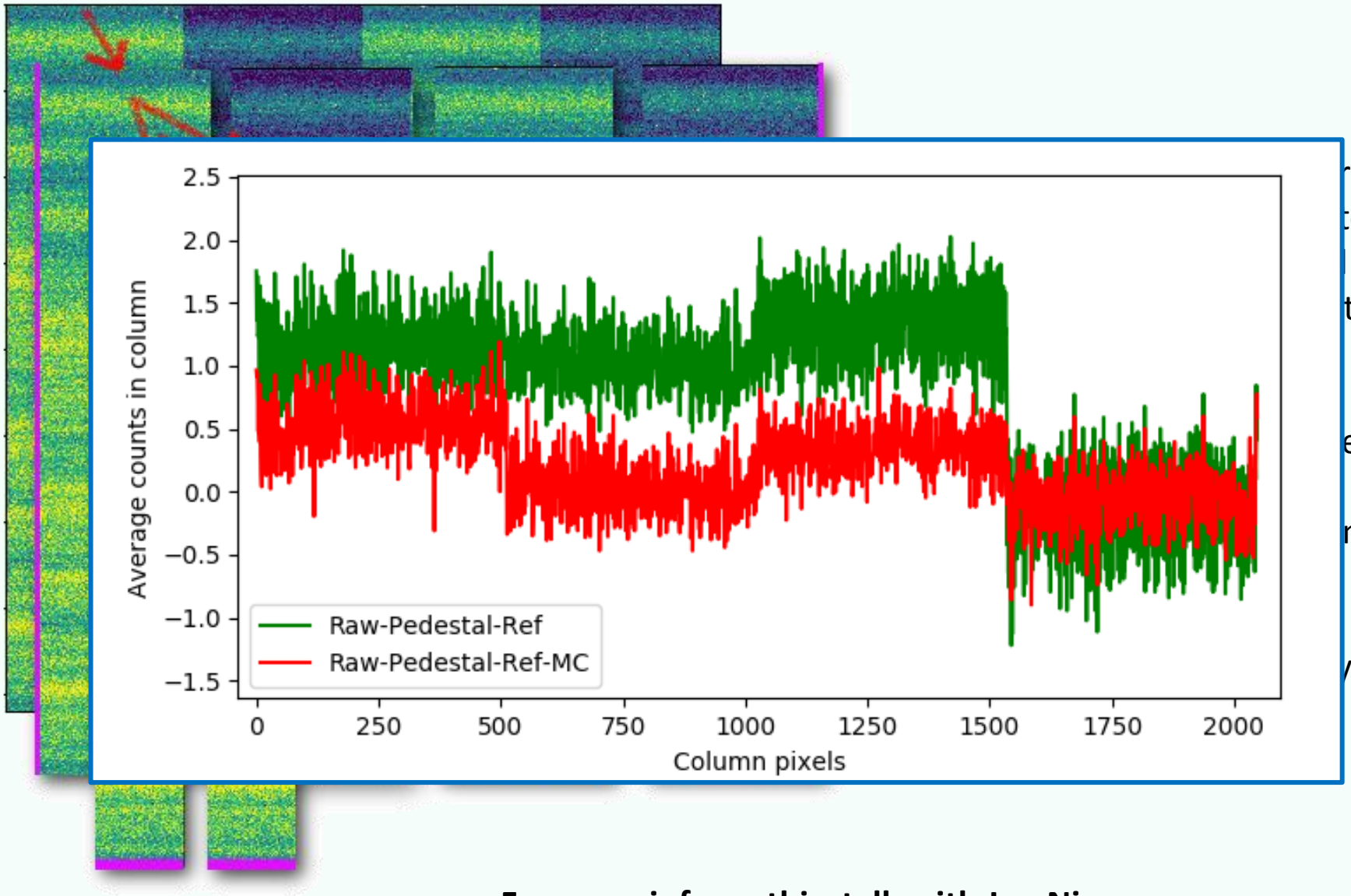


We have implemented a data driven scheme that partitions pixels into even & odd column groups, and uses median values between beams to correct the pattern noise.

This removes the noise structure while still utilizing the standard Teledyne SIDECAR microcode, and standard readout electronics.

It can be applied to data already taken and archived.

H2RG Better Bias Correction



riven
to
uses
to

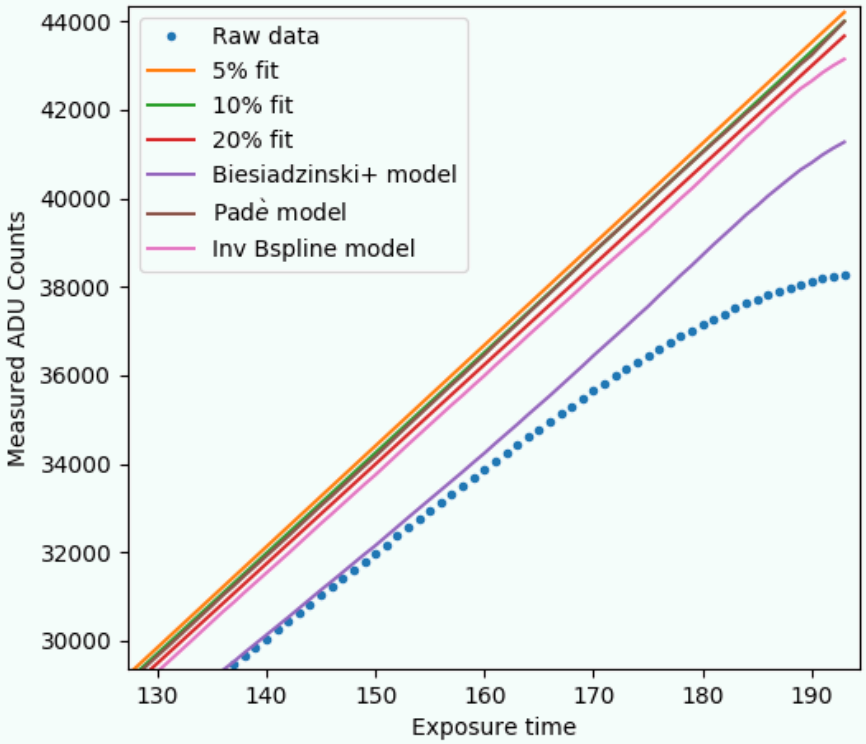
e

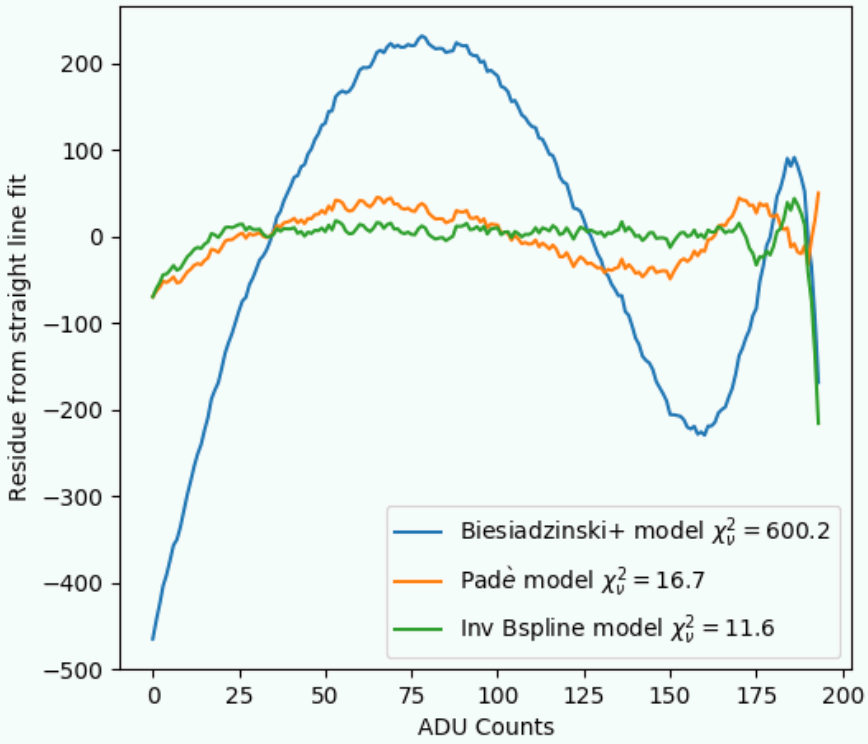
nd

y

For more info on this, talk with Joe Ninan

H2RG Linearity correction



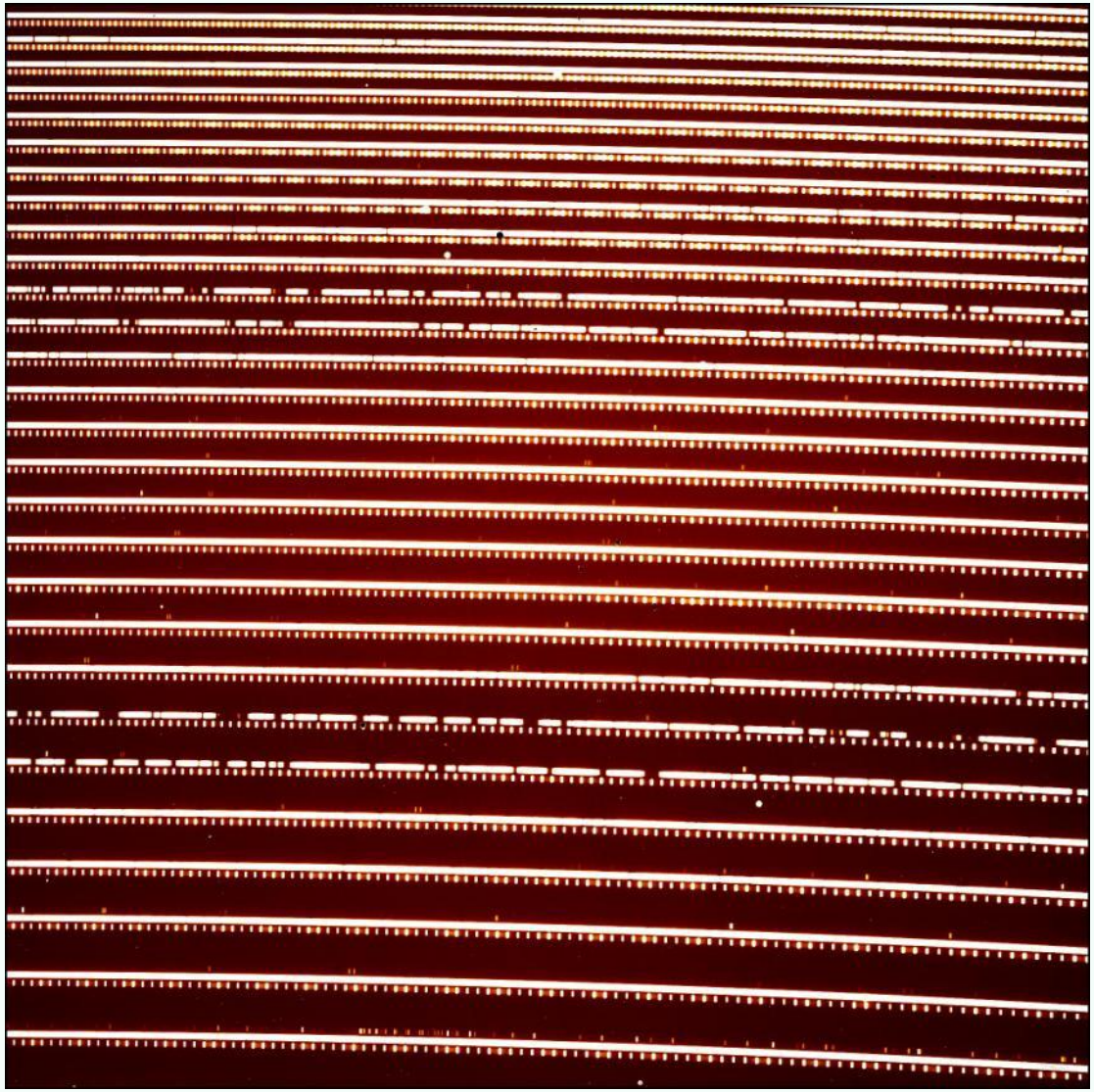


We correct this on a pixel-by-pixel basis

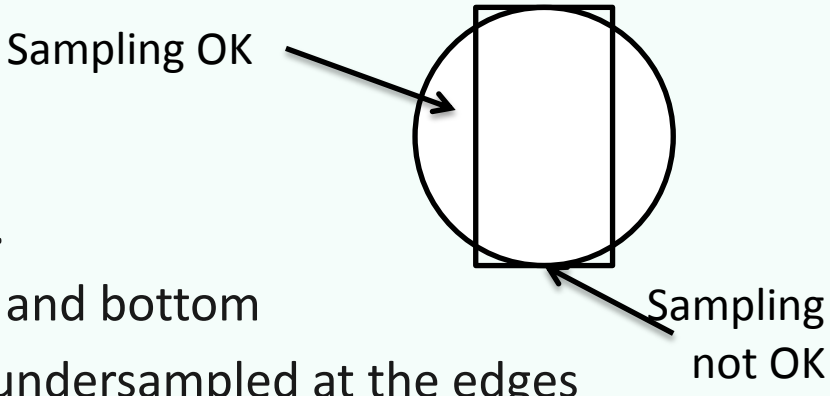
Side Note: This was a computational challenge in our APCP – the model is several GB, and created an IO problem. Solved with a Julia server that sends linearity data to any process on demand

An HPF Slope Image

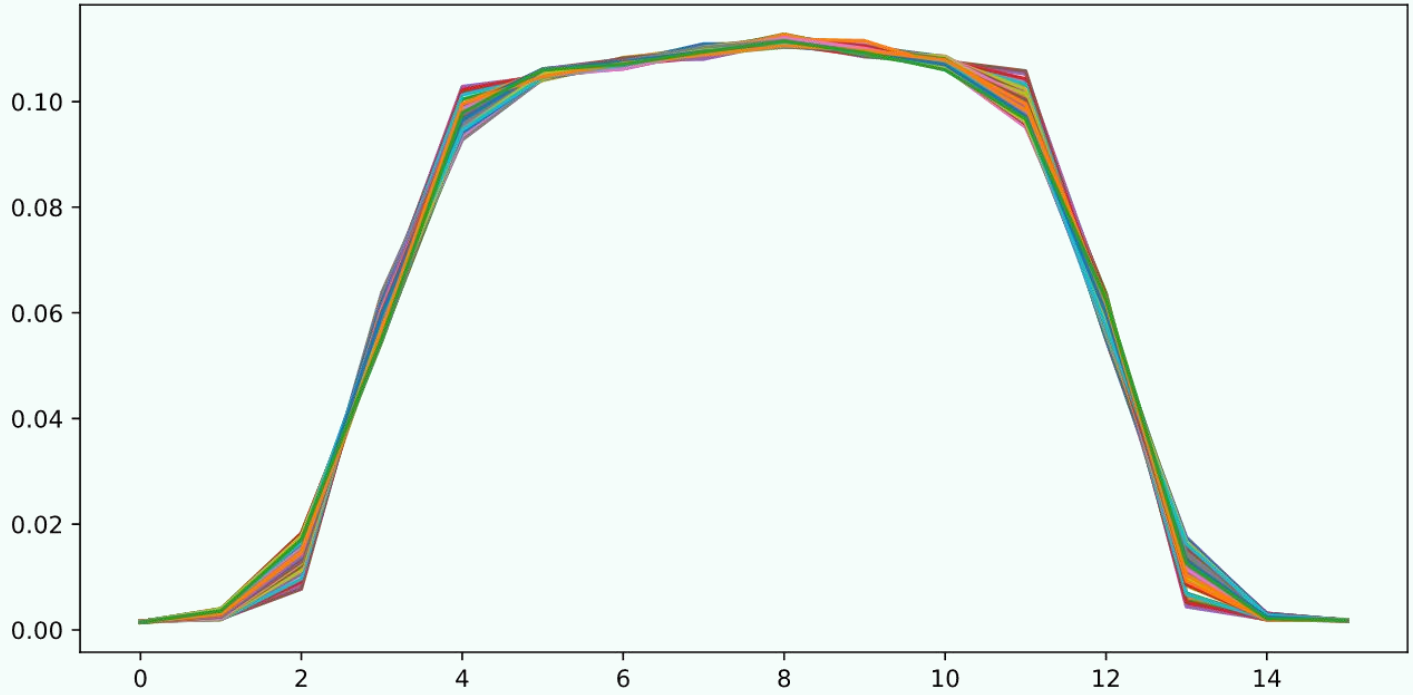
Processed image in units of e-/s (e.g. actual flux)



HPF Cross-Dispersion Profiles

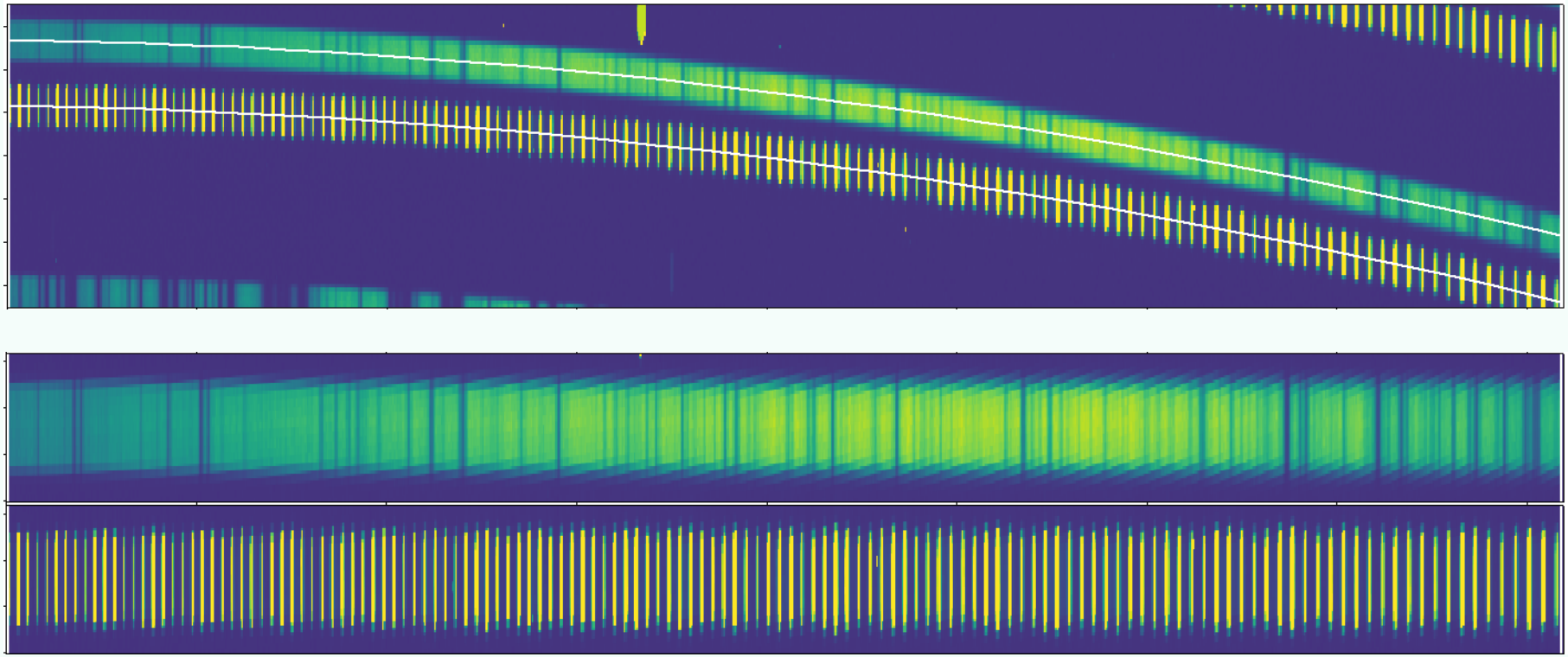


- The HPF fiber has a slit glued on the end.
- This exacerbates a sharp edge at the top and bottom
- XD profiles are very tophat shaped, and undersampled at the edges despite pixel sampling of ~3 pix per res element.



Extraction Aliasing

- This creates a classic aliasing pattern at the beam edges
- Flux is conserved, so sum extraction is OK
- But classic optimal extraction fails because XD smoothness assumption is violated



Extraction Aliasing

Solve this by mapping the aliasing with the flats

Complicated, but effective

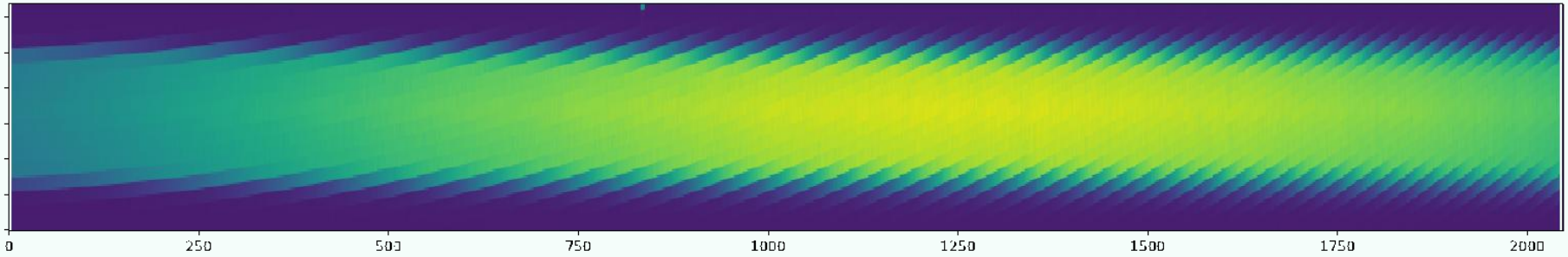
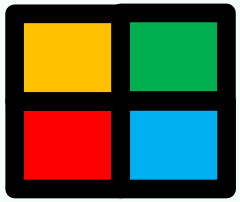
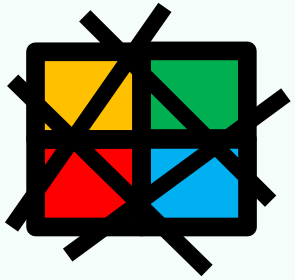


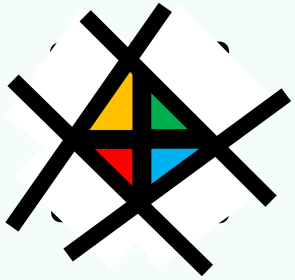
Image rectification done with Polygonal Clipping algorithm (Smith et al. 2007) to ensure absolute flux conservation



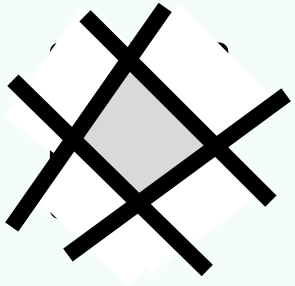
1) Detector pixels



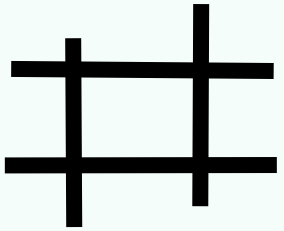
2) Polygon mapping



3) Pixel weights give flux distribution



4) Total up flux in recitified pixel



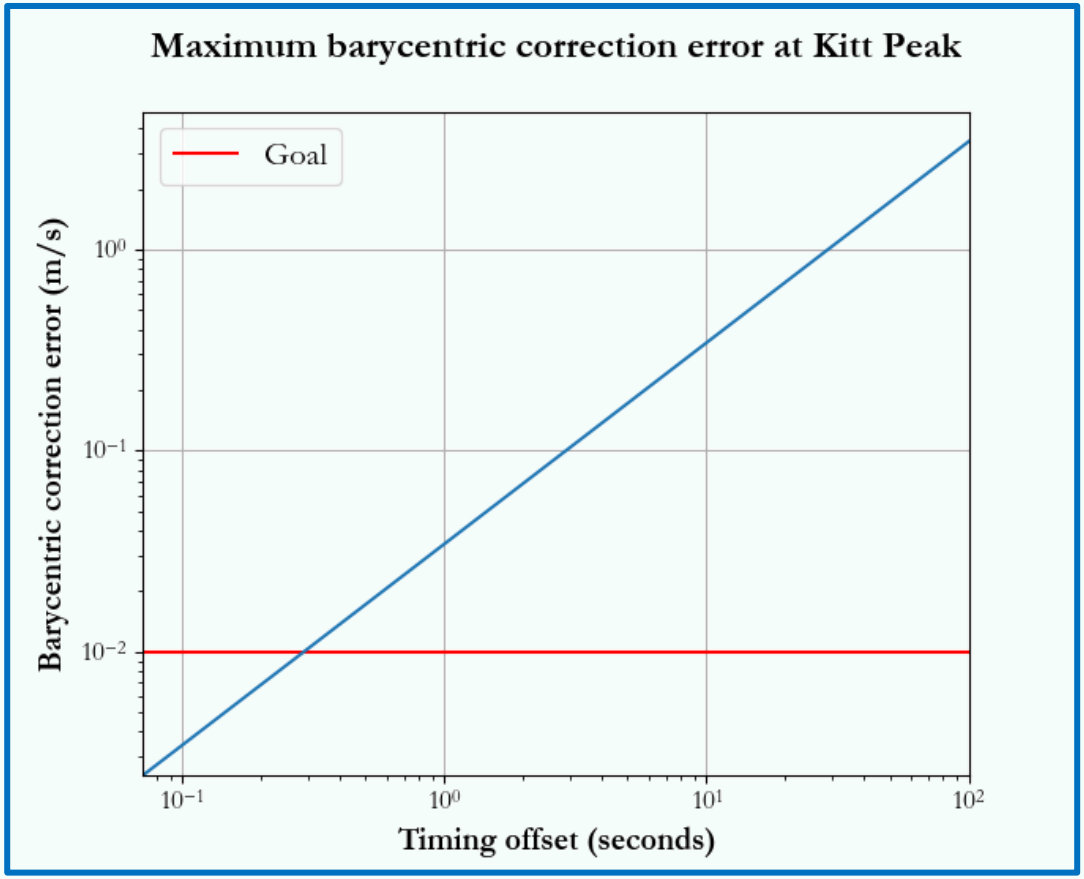
5) Map to rectified image space

For more details, talk to Kyle Kaplan

Barycentric Correction

Accurate center time of the photons you collected is critical for calculating the BC

- HPF – H2RG UTR mode provides this
- NEID – Secondary exposure meter – low-res. spectrometer provides chromatic timing

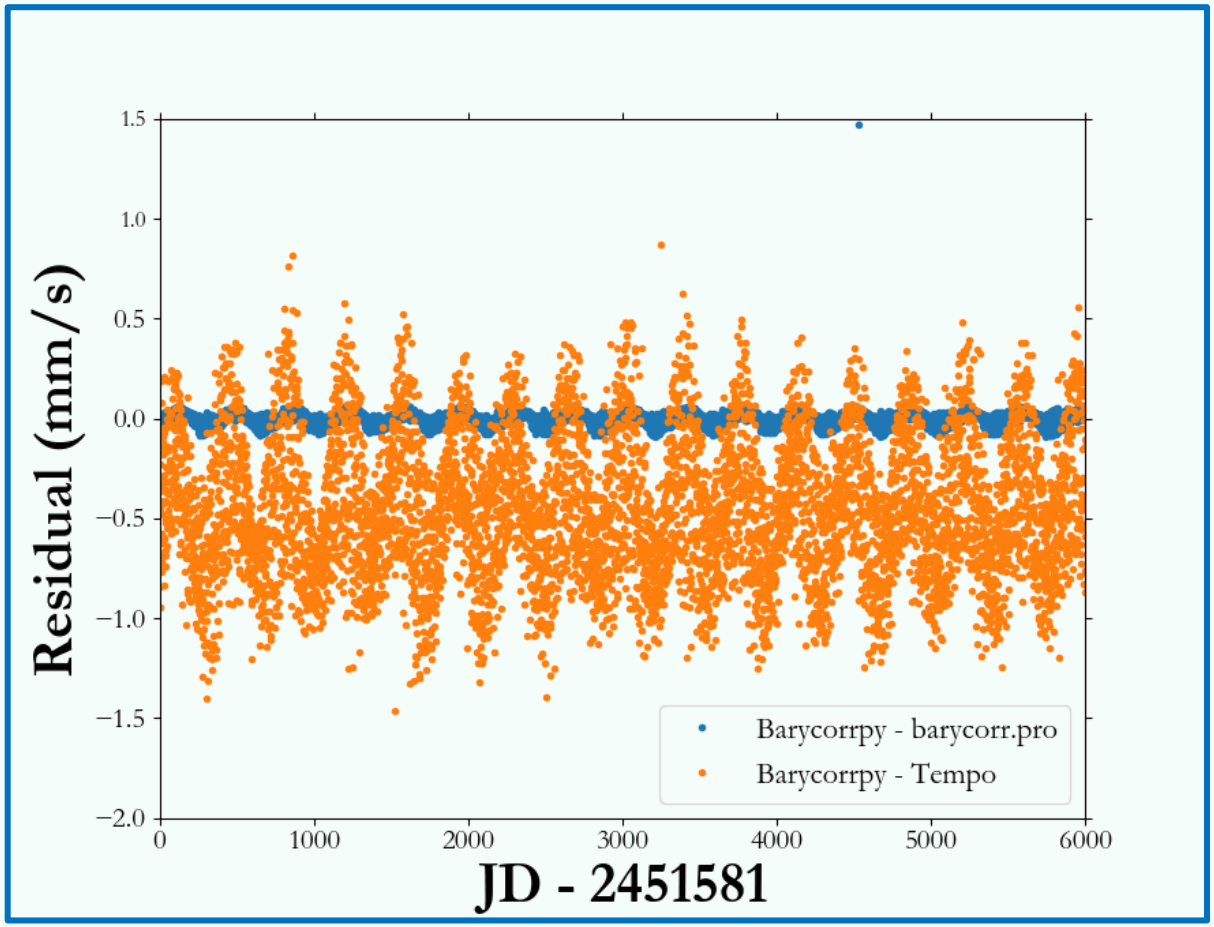


NEID Error budget requires knowing this to ~250 ms chromatically

Barycentric Correction

- But, calculating BC from known exposure time is a solved problem
- Barycorrpy - https://github.com/shbhuk/barycorrpy
  - python code adapted by Kanodia from Wright & Eastman 2014

Yes, that's mm/s!



See poster by Shubham Kanodia for more details

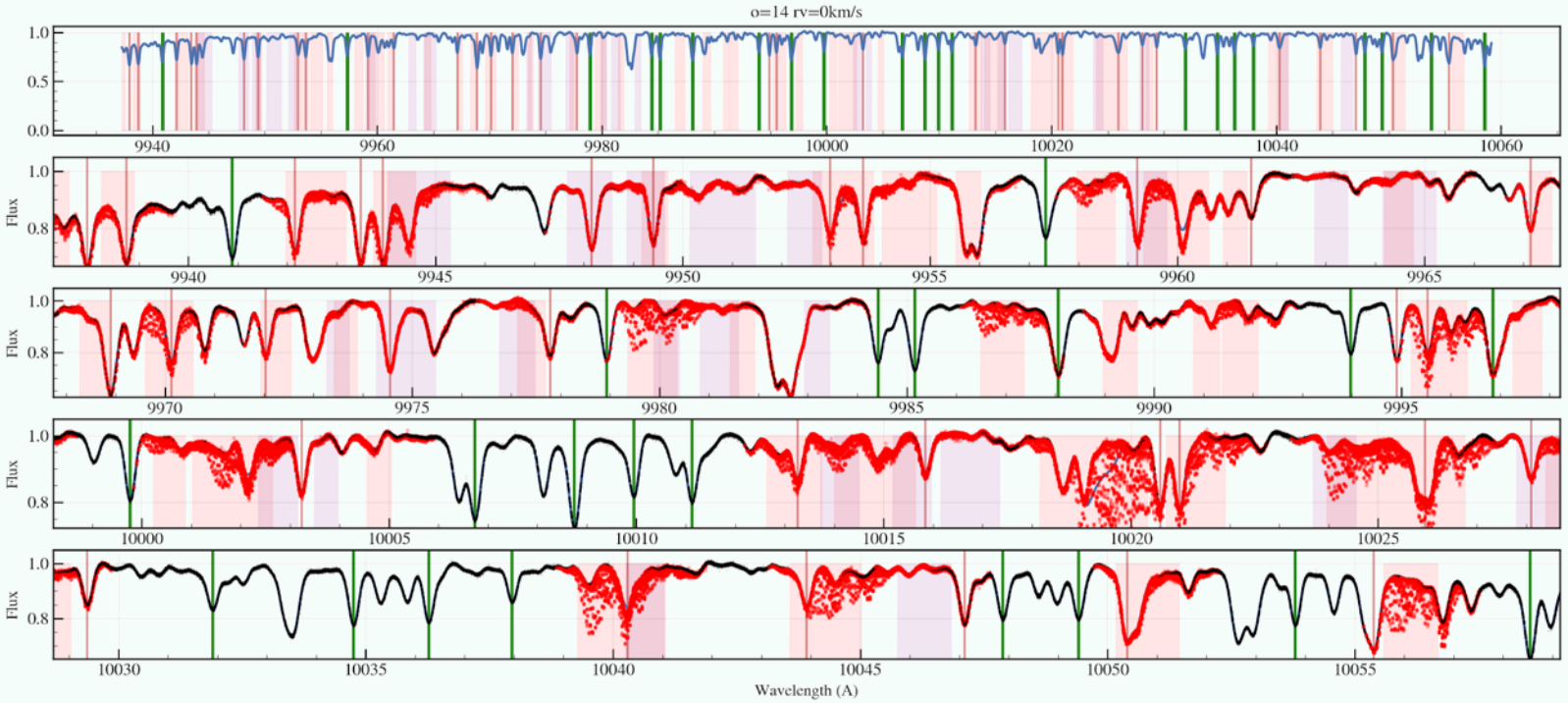
RV Derivation – Template Fitting

GJ699 RV analysis with SERVAL-like template fitting

HPF Order 14: 9940 – 10060 Angstroms

Red: Telluric contamination

Green: Line centers of regions used



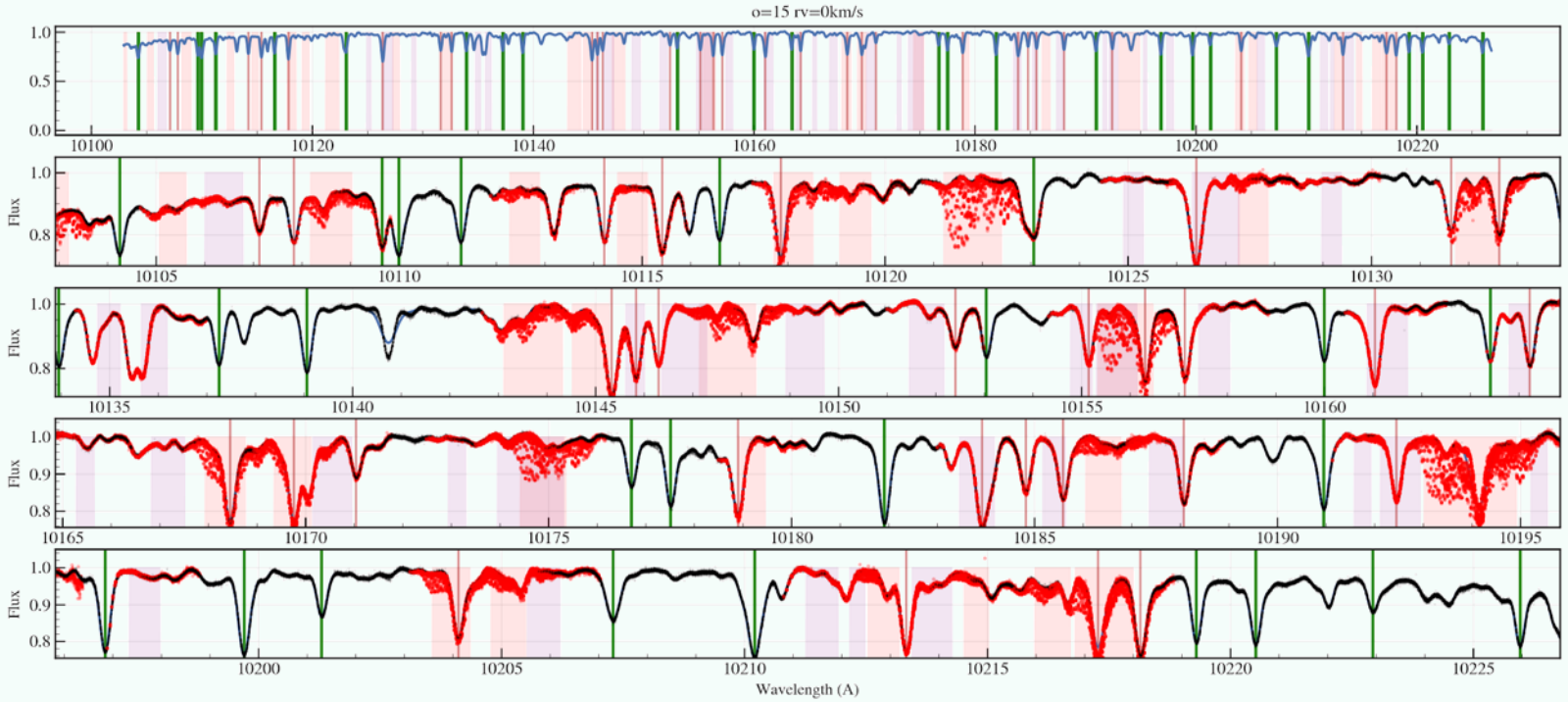
RV Derivation – Template Fitting

GJ699 RV analysis with SERVAL-like template fitting

HPF Order 15: 10100 – 10225 Angstroms

Red: Telluric contamination

Green: Line centers of regions used



For more details, talk to Guðmundur Stefansson

Telluric Correction

Our current RVs are limited by tellurics: we are not utilizing a lot of the spectrum!

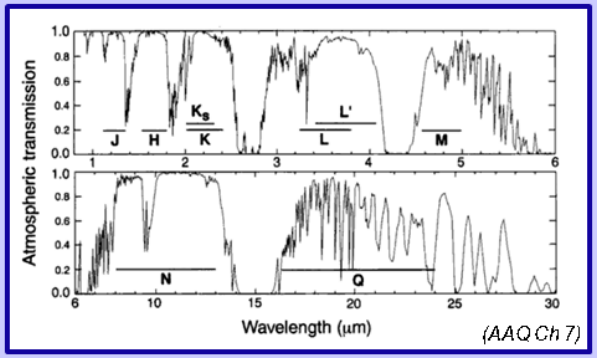
But this is not a new problem.

We've been working on it for a long time.

Title slide from EPRV I in 2010

Advances in Telluric Characterization for Precision Spectroscopy

Chad Bender
Penn State University



Collaborators:
Brandon Botzer, Sara Gettel (PSU)
The PSU Pathfinder Team: Suvrath Mahadevan, Larry Ramsey, Steven Redman, Ryan Terrien (PSU)
John Carr (NRL)
The NIST Laser Comb Team: Scott Diddams, Frank Quinlan, Gabe Ycas (NIST), Steve Osterman (CASA)

We've been working on it for a long time.

TERRASPEC Correction

- Similar to other forward modeling routines that use LBLRTM from AER
- What works well:
  - Correction of most weak (micro - ~40% depth) lines to < 0.5%
  - Updated HITRAN 2016 H_2O & O_2 linelist has improved
- What doesn't:
  - TERRASPEC is slow (CPU hours – days per spectrum)
    - We have a plan to solve this by eliminating real-time calls to LBLRTM
  - Automated version does not correct as well as by hand
    - Incorporate better initial guess parameters and some quality checks
    - Use existing HPF data to build better stellar templates
  - No sky emission
    - Add Kaplan OH code
  - Line profile is a voigt
    - Limitation of LBLRTM – Need a F90 programmer to help!
  - Remainder of HITRAN2016 not included
    - Working with HITRAN folks to update lists outside of AER update cycle

NEID Pipeline and Data Products

- NEID Pipeline run at NExScI within 24 hours
- Pipeline and data products governed by DMP and ICD
- Pipeline development tracked via Git with branches and versions
- Continued development post delivery (nominally 5 years)

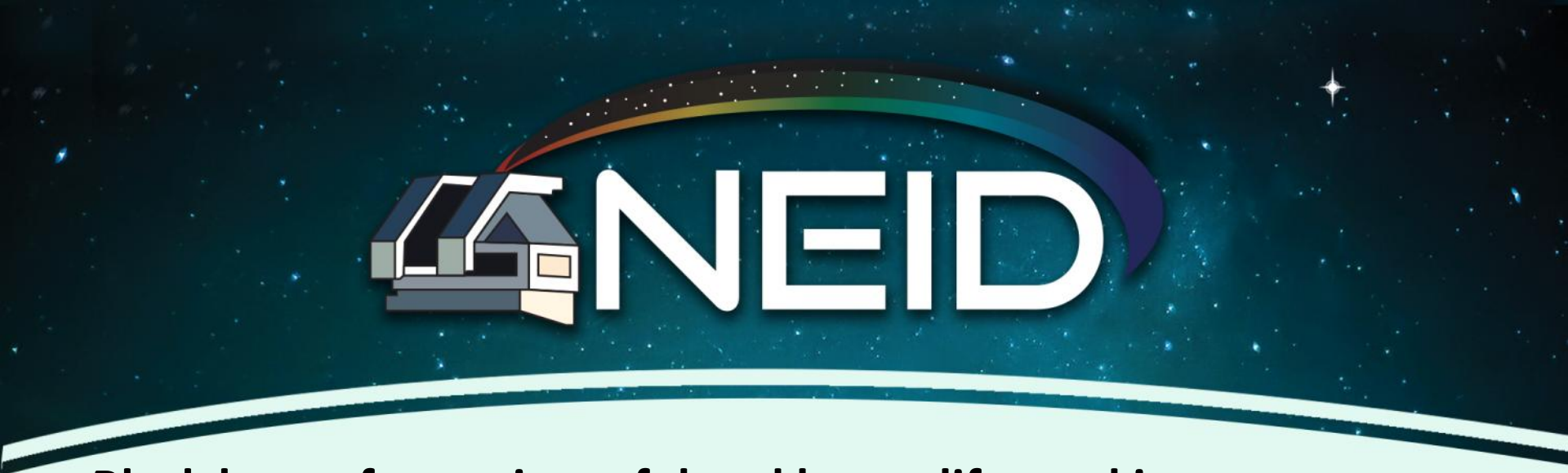
We are providing the following:

Level 0 Data:
- Raw echellogram images
- header metadata (~400 keywords))
- exposure meter time series
- guider camera image
- CFB datacube

Level 1 Data:
- Extracted Sci, Sky, Cal spectra
- Pixel-by-pixel variances
- Drift corrected wavelength solutions

Level 2 Data:
- Telluric Model
- Sky Model
- Scatted Solar Model
- CCFs
- Activity Indicator Parameters

Master Products:
- Pixel Masks
- Linearity Map
- Flats
- Master wavelength solutions
- CTI & Crosstalk maps
- Stitch Boundary maps



Black box software is useful and has poliferated in recent years

Open the black box and understand at least a little of what is going on inside

Mostly you won't get bit

Want to join the NEID software team?

Hiring a Post-Doc position at University of Arizona

Start date this fall!

Ad appearing soon