Global detrending

Current status

James McCormac - University of Warwick - PLATO Data Centre (Richard West)

Getting ready for PLATO meeting - Thursday 14th Sept 2023

Summary

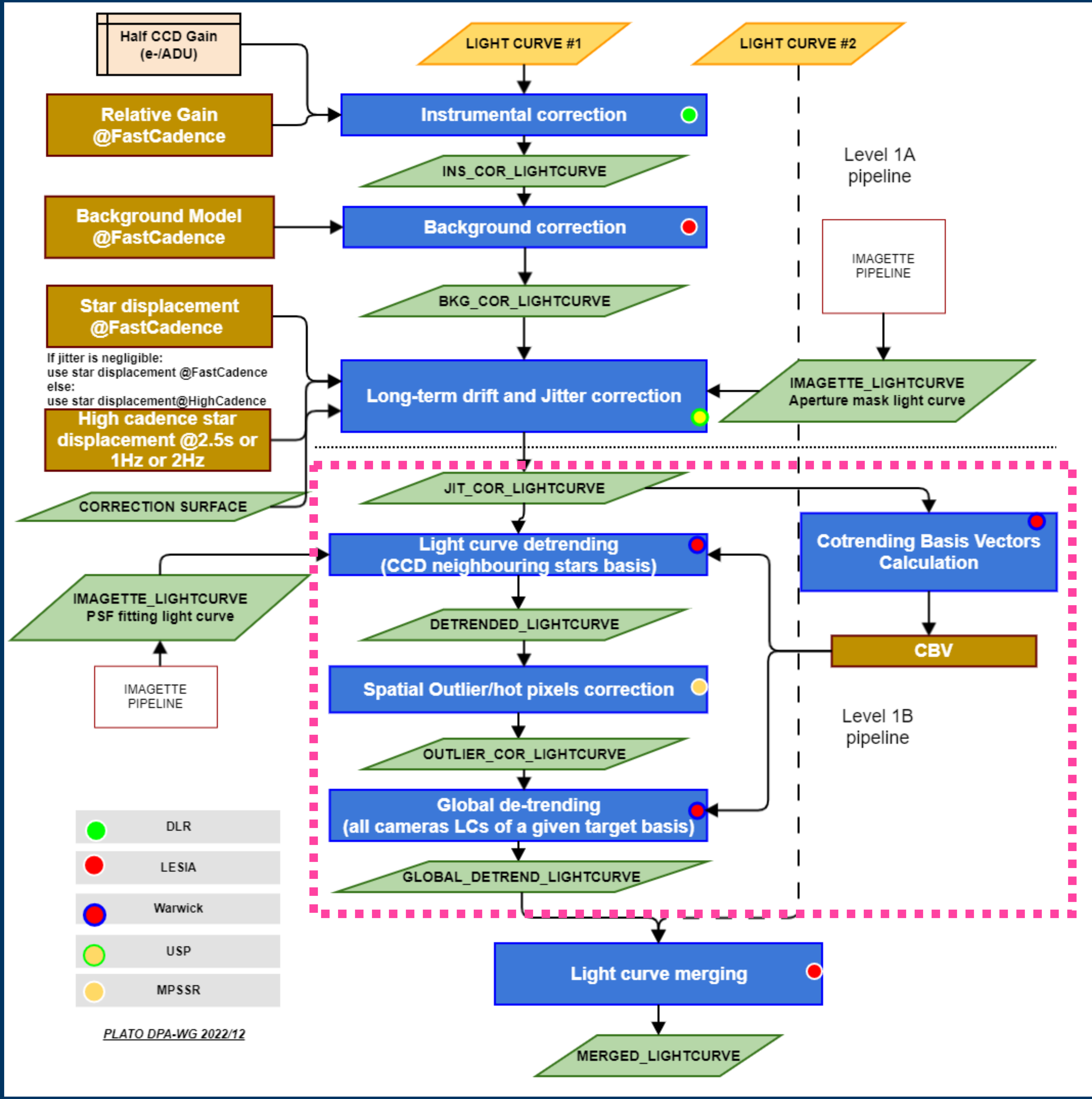
1. Position in the L1 pipeline
2. Summary of detrending methods
3. Existing simulated datasets (PIS/PSLS + PlatoSim)
4. Performance analysis
5. New Docker wrapper for PlatoSim + L1

Position in the L1 Pipeline

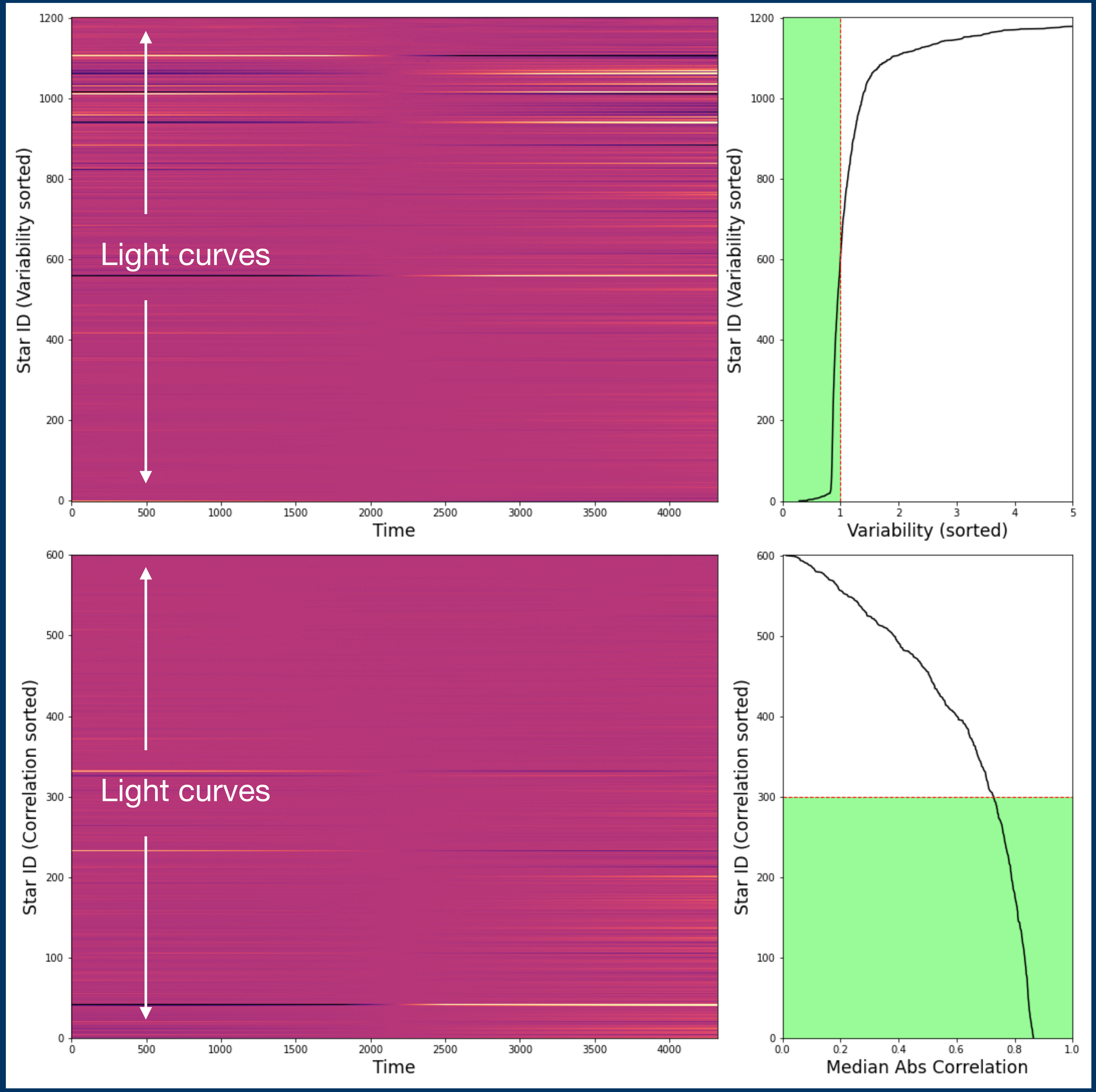
Second last step in L1
After 'long-term drift + jitter correction'
Before 'light curve merging'

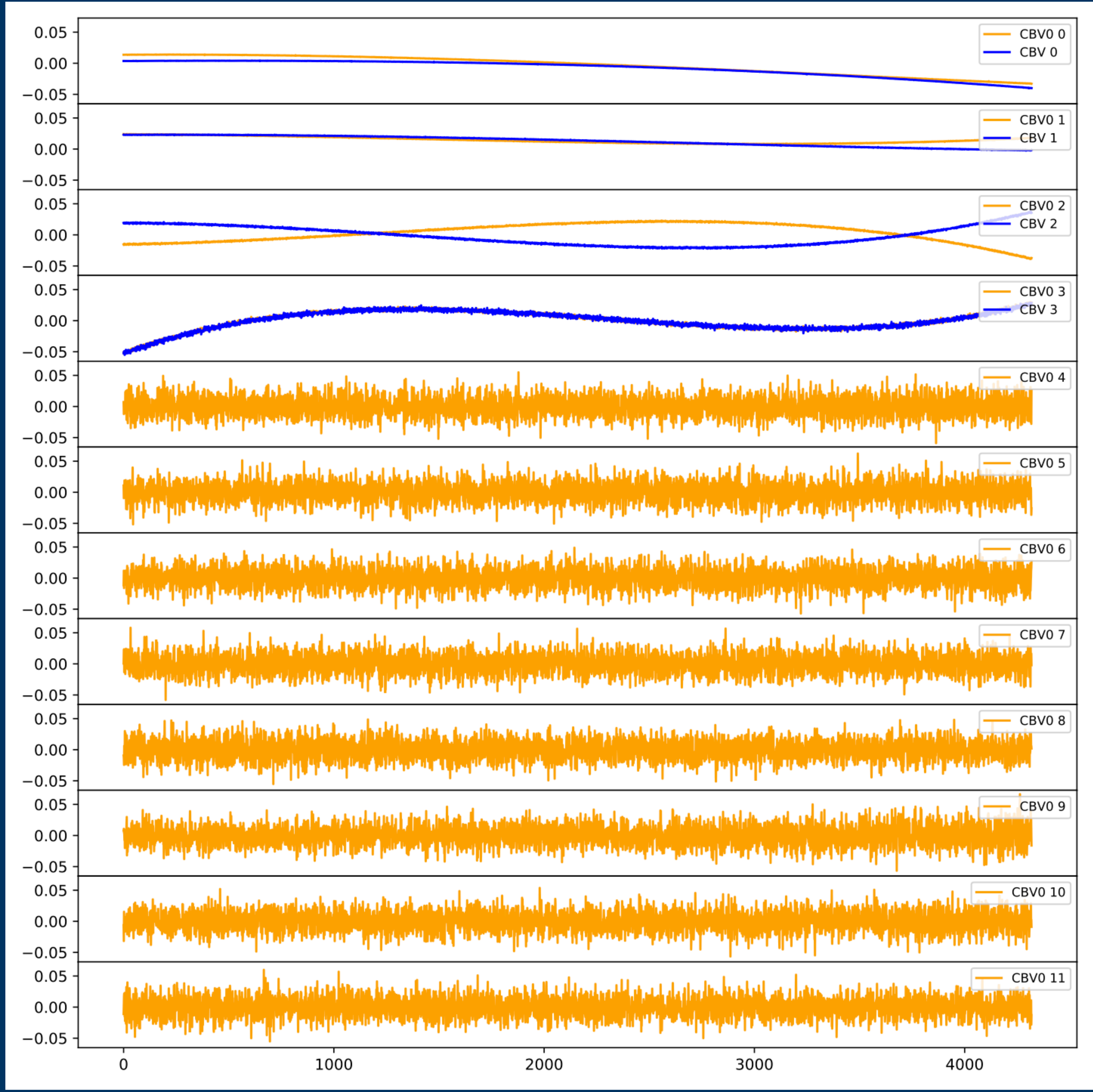
Investigating two methods of detrending:
Kepler PDC-MAP (single camera)
PLATO Republic (multi camera)

Both techniques depend on Cotrending Basis Vectors (CBVs)
CBV calculation split into a preliminary step



Extracting Cotrending Basis Vectors (per camera/CCD)





1. Flux / median(Flux)
Keep low variability objects

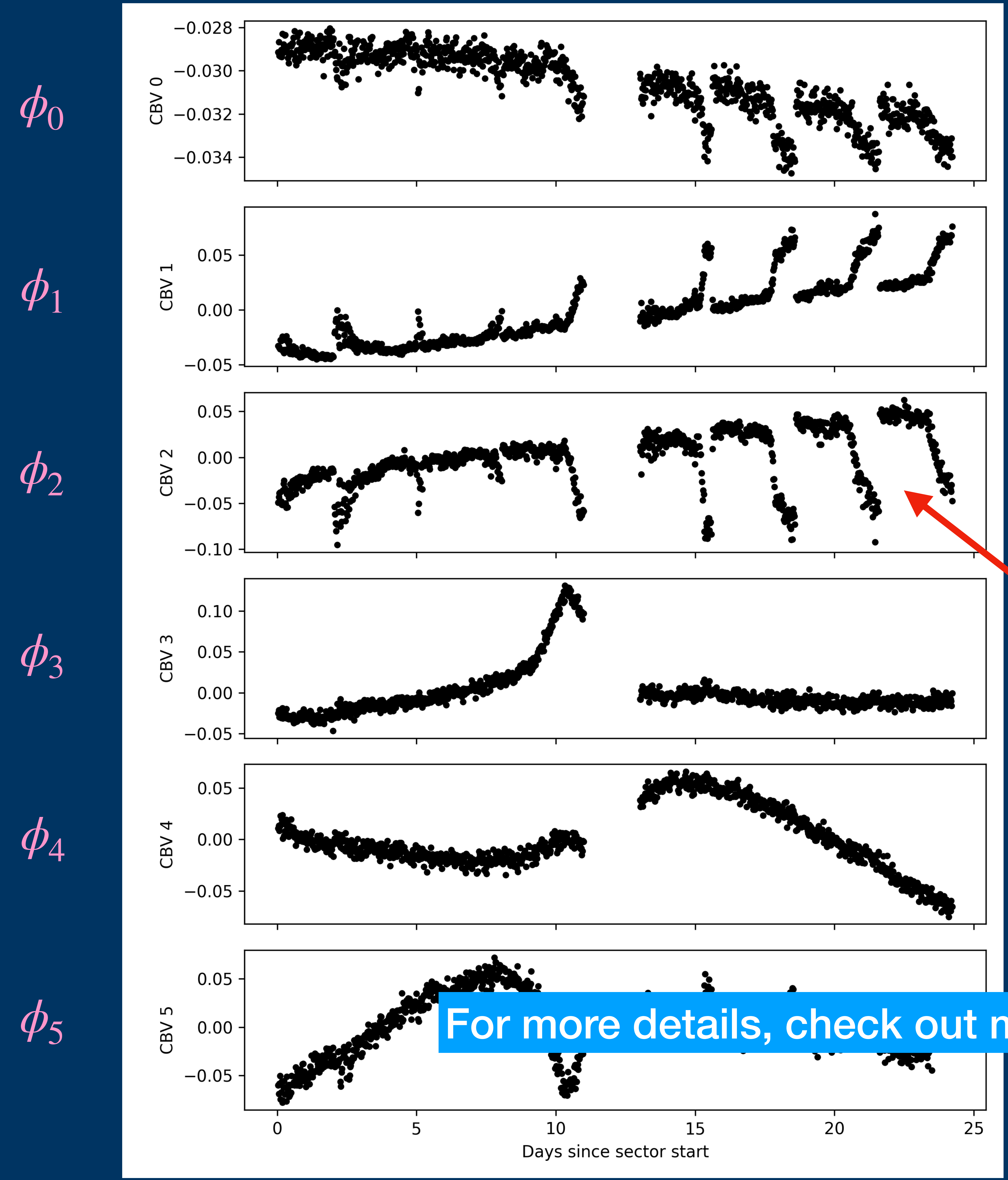
2. Correlate non-variables with each other.
Rank by correlation, keep top 50%

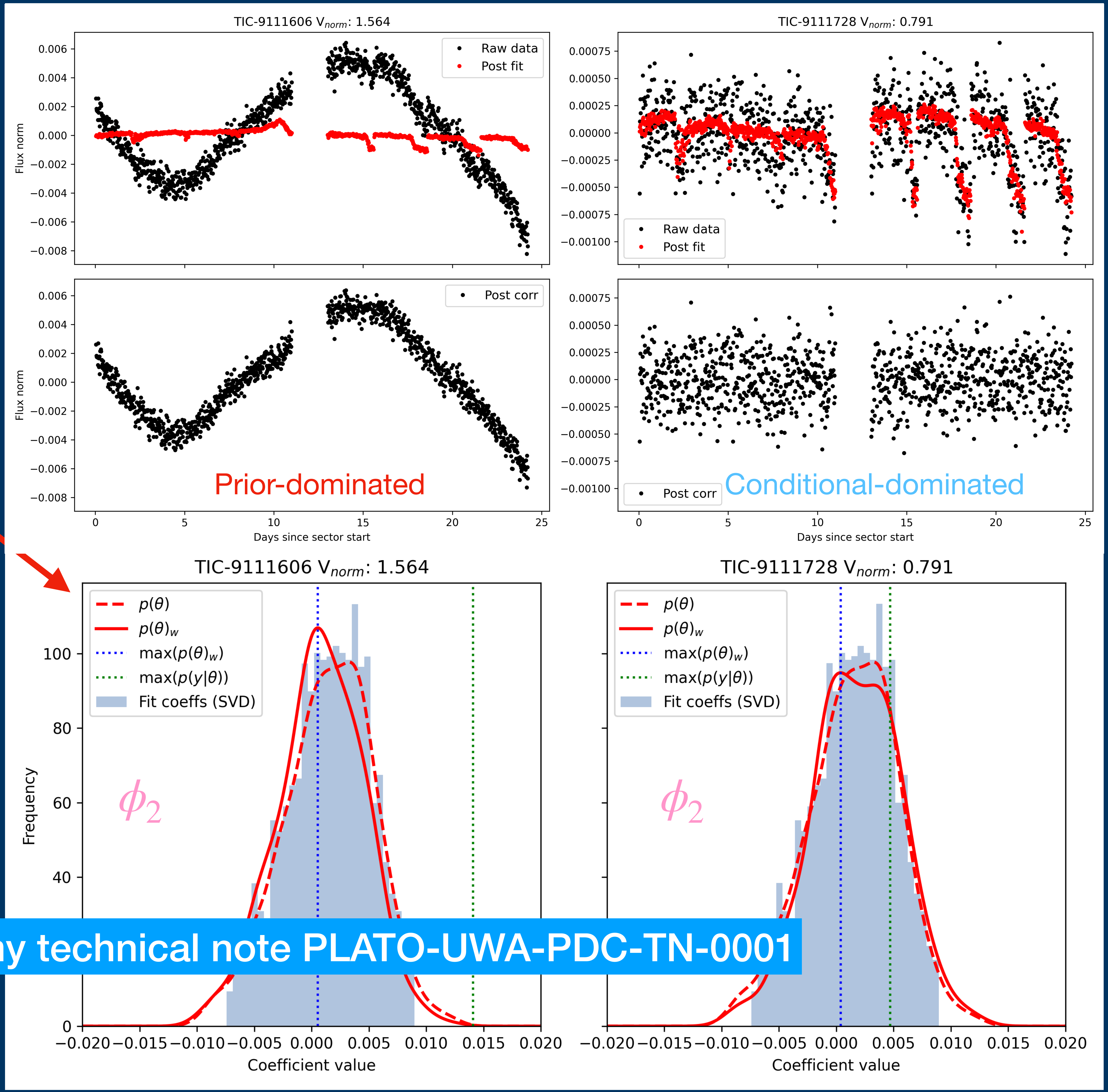
3. Entropy cleaning
SVD on remaining stars

4. Check CBV's SNRs
Keep those > 5dB

MAP application of CBVs - TESS FFI data - Sector 05_1-1

Variable object **= more prior weight** Quiet object **= less/no prior weight**





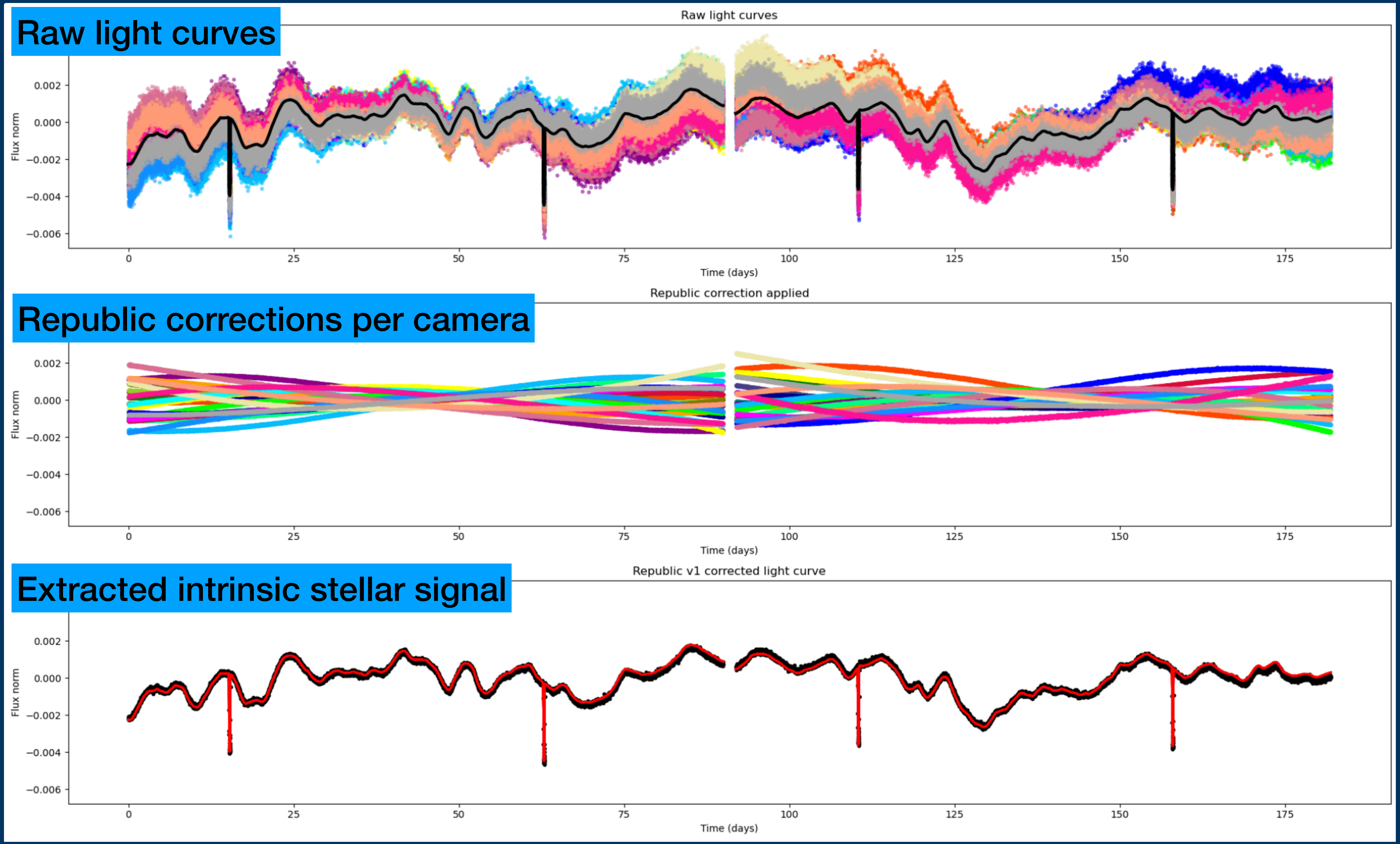
For more details, check out my technical note PLATO-UWA-PDC-TN-0001

Republic v1

linear combination
CBVs*fit_coeffs per camera

Observed flux (per-camera) = Intrinsic Stellar Signal + Systematics Model + Random Noise

Common across cameras



Considers single star light curves and CBVs from all cameras

Fits intrinsic stellar signal as free parameter at each time step and simultaneously determines all CBV fit coefficients

Large set of simultaneous equations solved via matrix inversion

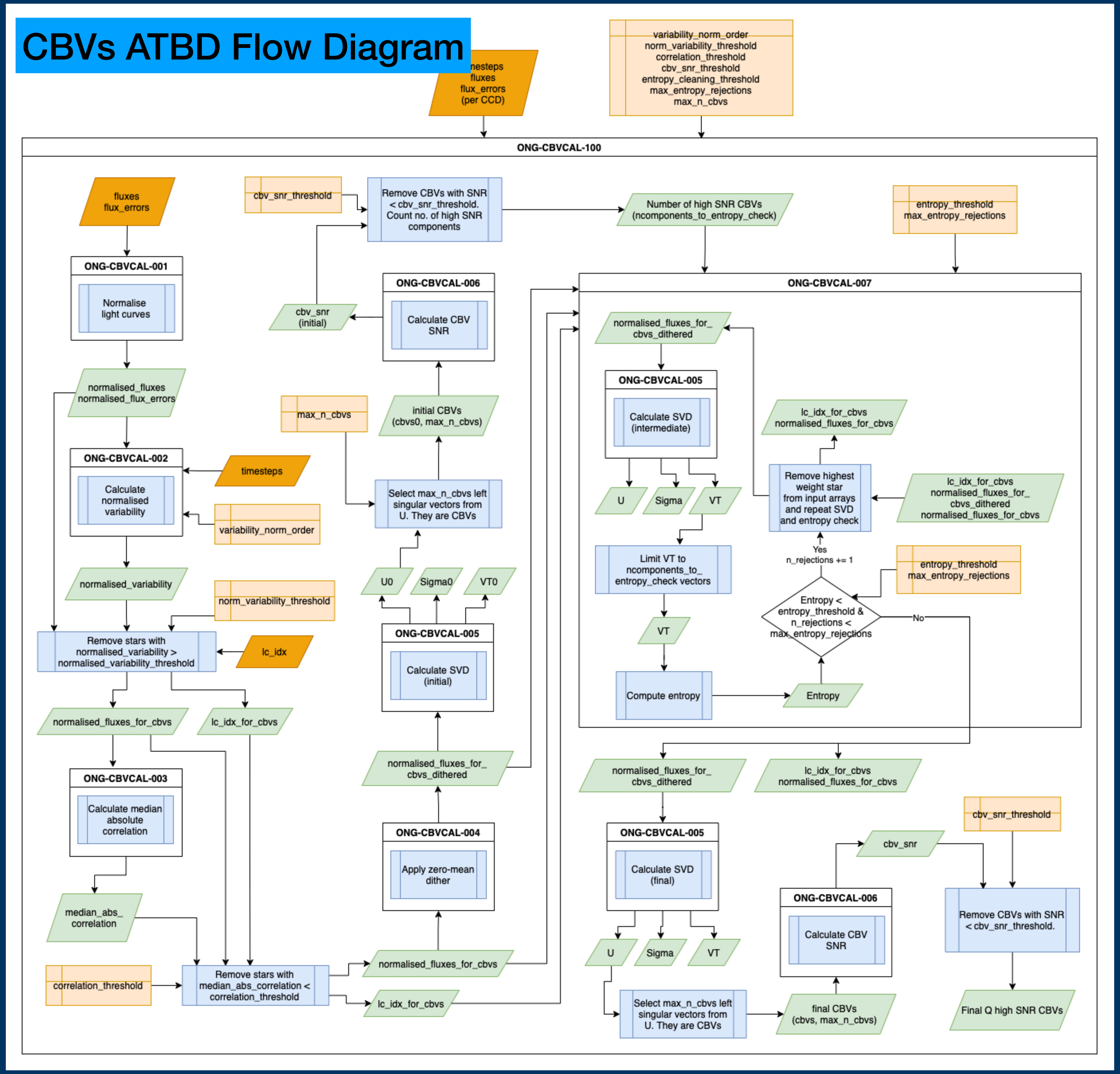
Assumes systematics independent between cameras

Preserves signals common across cameras regardless of origin

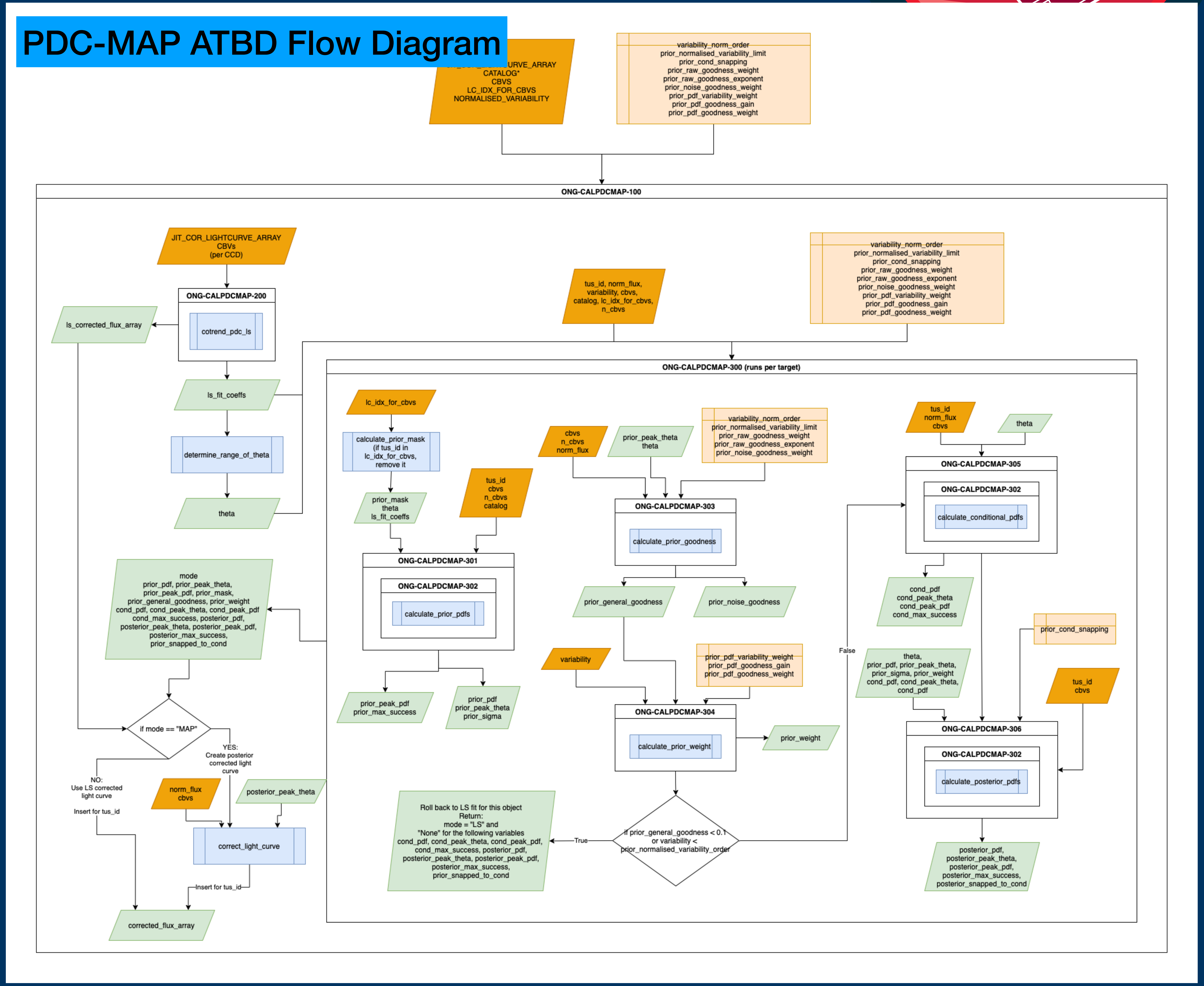
PLATO Algorithm Theoretical Baseline Documents (WP32)

1. Calculation of Cotrending Basis Vectors (PLATO-UWA-PDC-DD-0001)
2. Cotrending using Kepler PDC-MAP (PLATO-UWA-PDC-DD-0002)
3. Cotrending using Republic algorithm (PLATO-UWA-PDC-DD-0003)

CBVs ATBD Flow Diagram



PDC-MAP ATBD Flow Diagram



Summary of PIS/PSLS simulations

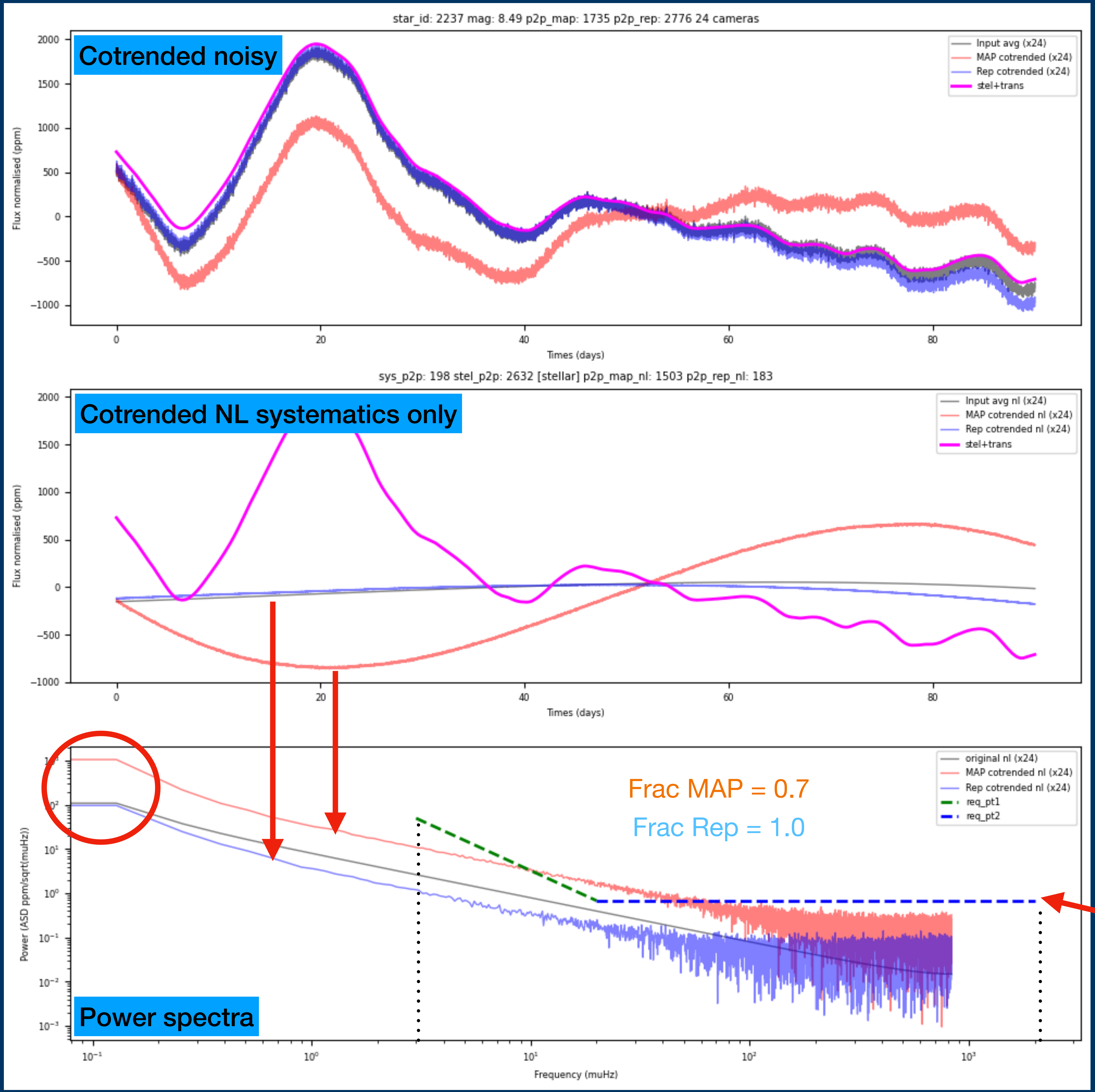
PLATO Image Simulator (PIS)
PLATO Solar-like Lightcurve Simulator (PSLS)
Early simulator used prior to PlatoSim

Simulation	N cams	N qrt	N stars	Cadence	Mask	LCs	Stel	MAP	Rep1	Comments
BOL4_2k_bm	24	8*	2300	10 min	—	Y	Y	Y	Y	MAP p/q; Rep1 multi-q
BOL4_2k_ubm	24	8*	1800	10 min	2w	Y	Y	Y	Y	MAP p/s; Rep1 multi-q
BOL4_2k_psf	24	8*	1700	10 min	—	Y	Y	Y	Y	MAP p/q; Rep1 multi-q

Large set of simulations:
- 24 cameras
- 8 quarters
- ~2k stars each
- 10 mins cadence

- BM: binary mask photometry
- UBM: updating binary mask photometry (P5)
- PSF: PSF photometry (P1)
- Systematics simulated by series of polynomials
- Stellar signals simulated by S. Aigrain (Oxford)

*Multiple quarters simulated by cycling systematics from camera to camera (1->2, 2->3, etc at each quarter boundary)



PIS/PSLS
Power Spectrum Analysis

Calculate per-camera CBVs
Calculate per-target/camera fit coeffs (MAP | Republic)

Cotrend noiseless (NL) light curves using ^
Create 24 camera average

Calculate power spectrum of NL
MAP | Rep corrected lcs (24 cam avg)
Calculate fraction (Frac) of PS frequencies
meeting requirements

Input lc (24cam avg)
MAP corrected lc (24 cam avg)
Rep corrected lc (24 cam avg)
Injected stellar signal

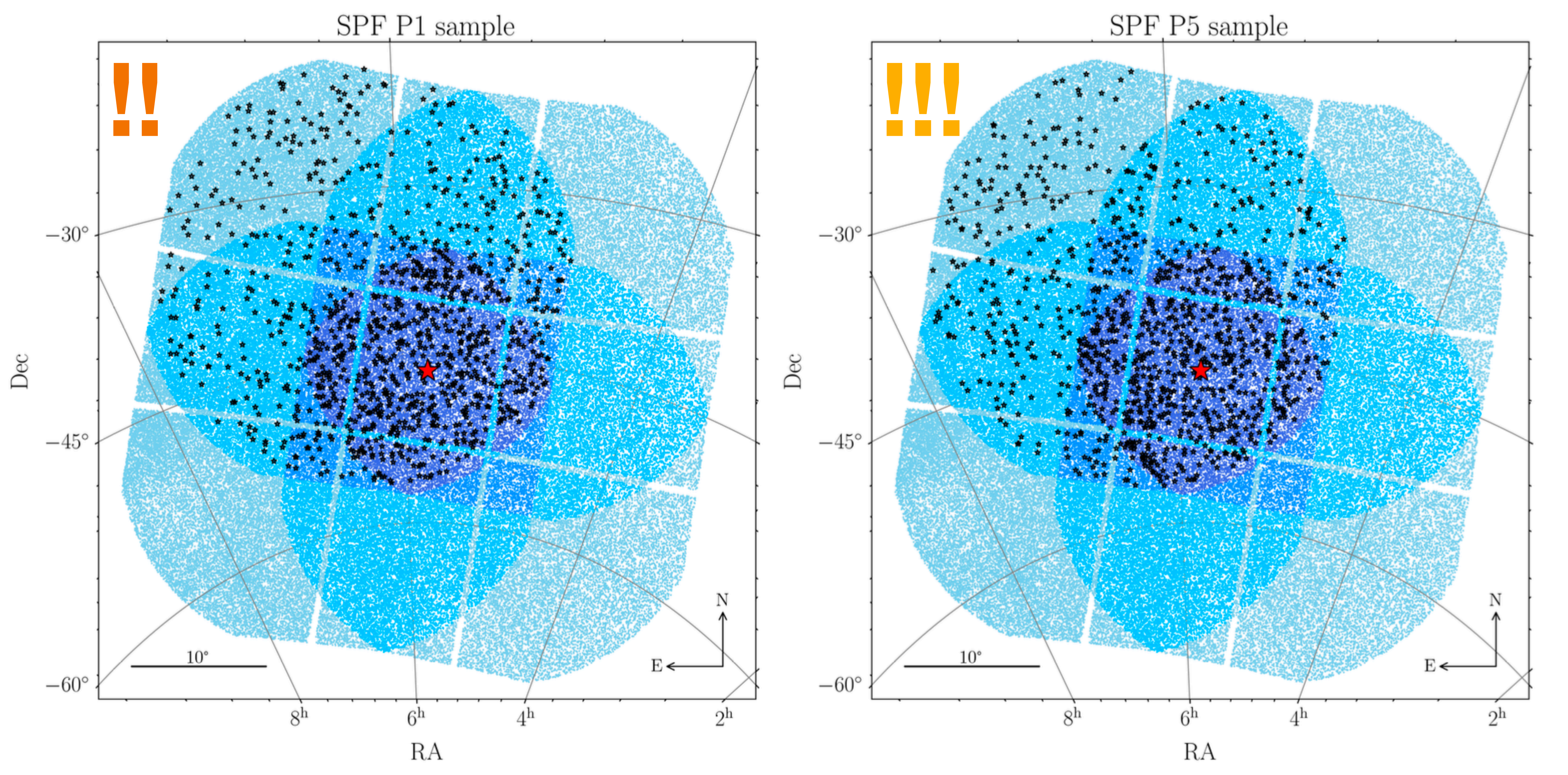
75% stars have
Frac>0.9 (MAP | Rep)

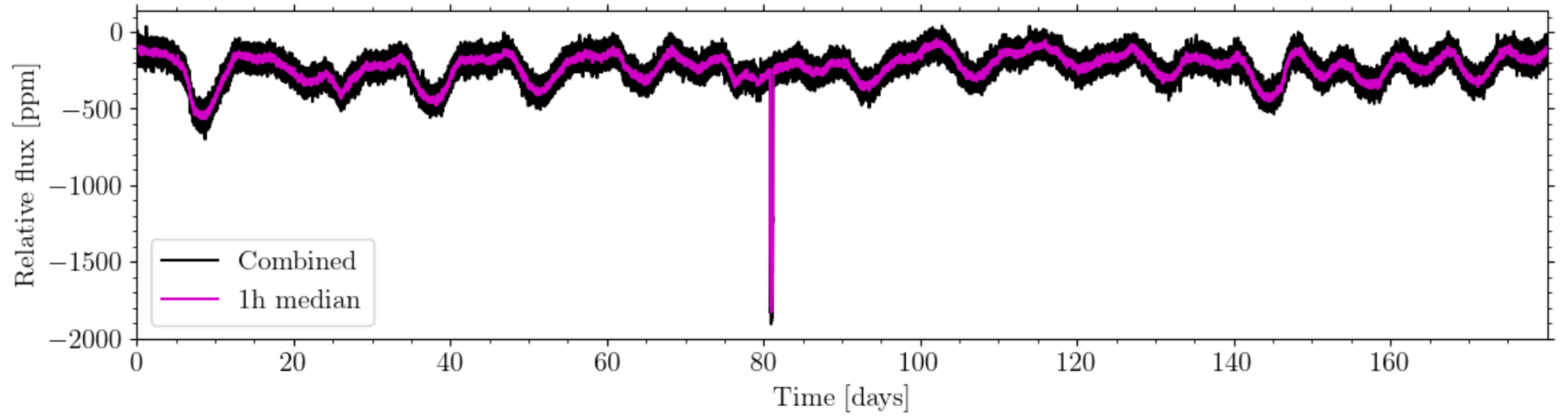
SOC-NF-GDP-GL1-1956	Amplitude of the Spectral Density of the mean residual errors in the L1 pipeline	Assuming the performance of the p… specified in the Payload User Requi… and a payload performance as given… Parameters for Performance Studies (PLATO-DLR-MIS-LI-0002), the ground data processing pipelines (including algorithms and calibration) shall be such that the (double-sided) Amplitude Spectral Density of the mean residual errors in the L1 lightcurves for stars of magnitude mv = 11 in stellar sample 1, after on-board and ground processing, are: - lower than the monotonic rise from 0.68 ppm/√μHz at 20 μHz to 50 ppm/√μHz at 3 μHz, at frequencies below 20 μHz; - lower than 0.68 ppm/√μHz at frequencies between 20 μHz and 40 mHz

Non-functional on-ground data processing L1 requirement

PlatoSim Simulations

- First large scale PlatoSim + L1 run
- PLATO-KUL-PL-TN-0020 (N. Jannsen)
- 1000 P1 & 1000 P5 stars
- Distributed across multiple cameras
- 2 quarters (23 & 24, EOL)
- Best estimates for spacecraft parameters
- Worst case assumed for DKA + TED
- Stellar activity, transits, photometric standards (S. Aigrain et al)
- P1 PSF photometry with L1 (LESIA)**
- P5 aperture photometry with L1 (LESIA)***



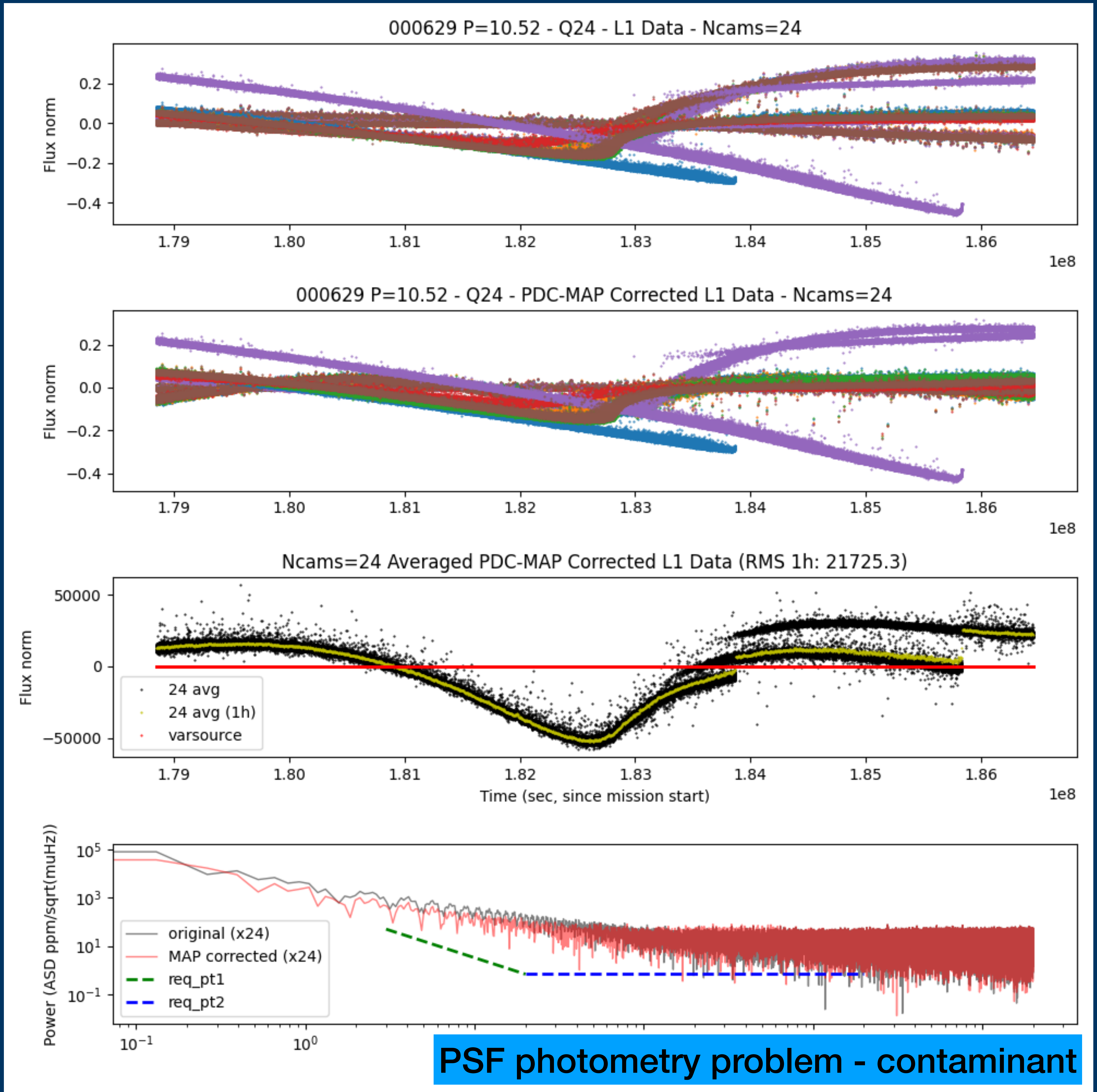


Input signals	No.
Standards: “Constant”	3
Standards: roAp	7
Granulation, oscillations	100
Granulation, oscillations, activity	445
Granulation, oscillations, activity, exoplanet	445

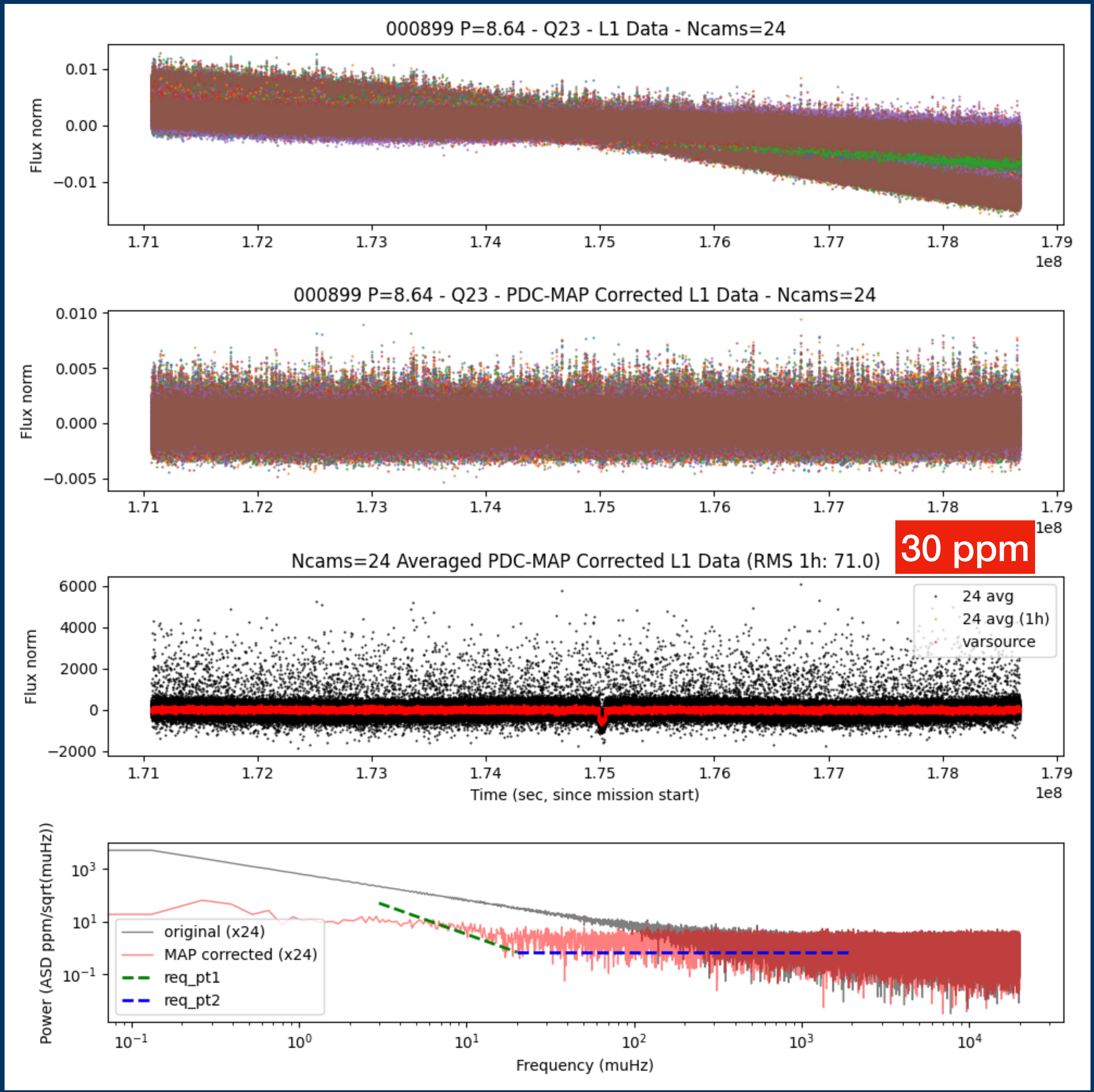
	P1 sample		*P5 sample*	
N-Cams	Targets	Time series	Targets	Time series
6	100	600	100	600
12	200	2400	200	2400
18	100	1800	100	1800
24	600	14400	600	14400
Sum	1000	19200	1000	19200

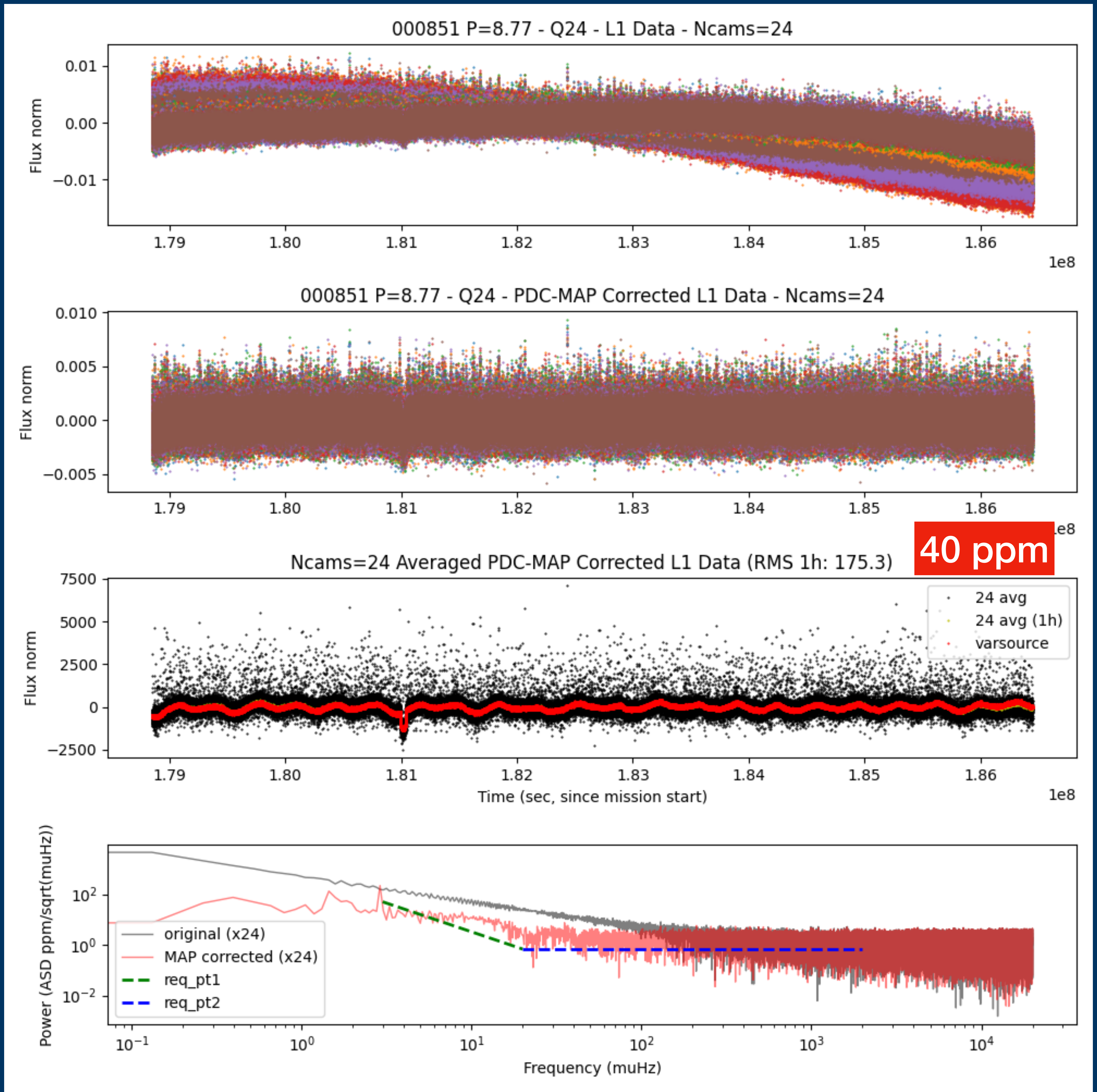
Teething Problems with PlatoSim + L1 simulations

1. Outlier rejection not applied
    1. Detected but not addressed
2. ~50% of 24 camera P1 data missing
    1. Multiprocessing issue on KU Leuven cluster
3. CBVs extracted at 25s cadence didn't capture low amplitude trends
    1. Binned to 600s, extracted CBVs and then interpolated CBVs to 25s
4. P5 data mask updates not applied
    1. Redoing P5 sims
5. PlatoSim & L1 pipeline still under development
    1. Adding features, fixing bugs etc

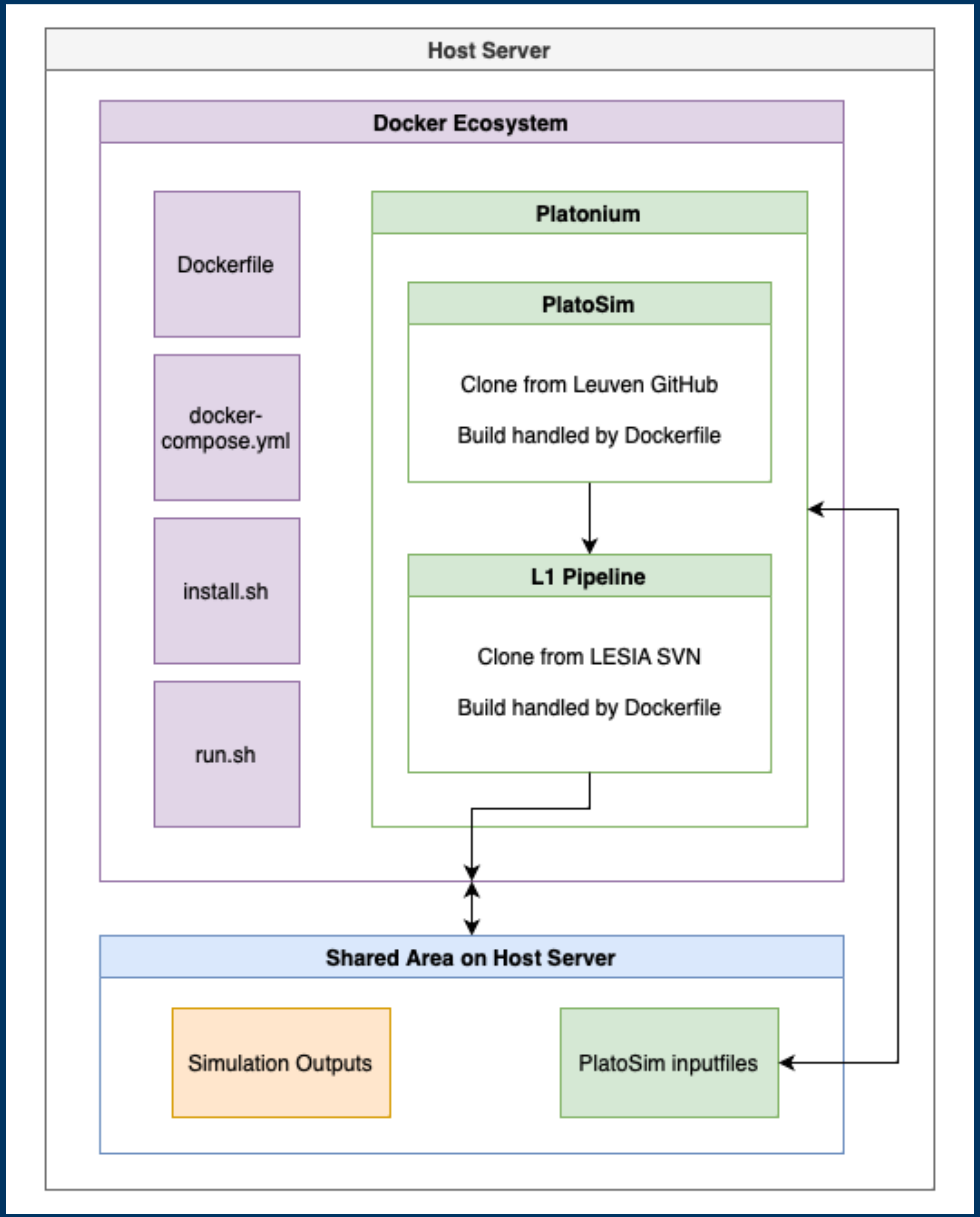


PlatoSim Power Spectrum Analysis - P1 sample





Docker Wrapper for PlatoSim + L1 Pipeline



Building a Docker image

1. Clone my docker_ecosystem repo, then inside:
    1. Clone PlatoSim3 repo
    2. Checkout L1 pipeline LESIA SVN
2. Edit docker-compose.yml
    1. Add PATH for shared area to mount inside container
    2. PlatoSim + L1 pipeline products saved here
3. Run install.sh to build Docker image
    1. Ubuntu image
    2. Poetry python install matching PlatoSim + L1 requirements
    3. Linux library requirements
    4. All environment variables set

Running a simulation with L1 photometry

1. Configure your PlatoSim inputfile.yaml
2. Platonium (N. Jannsen):
    1. Calls PlatoSim to generate imagettes
    2. Calls L1 pipeline for PSF (P1) or Aperture (P5) photometry
3. Results saved to mounted area on Docker host.

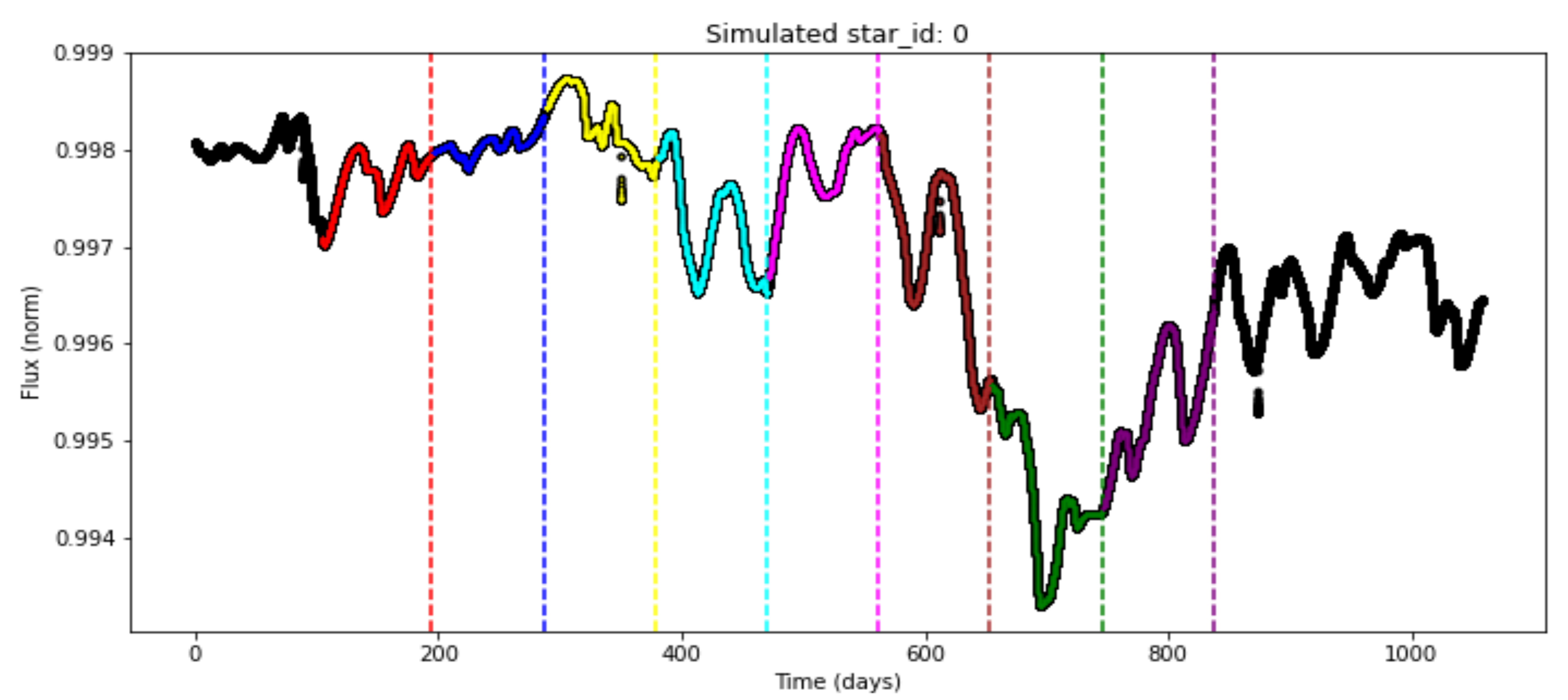
$> platonium <star_id> <cam_id> <group_id> <quarter_id> <project> <sample> <varfile>

Thanks, questions?

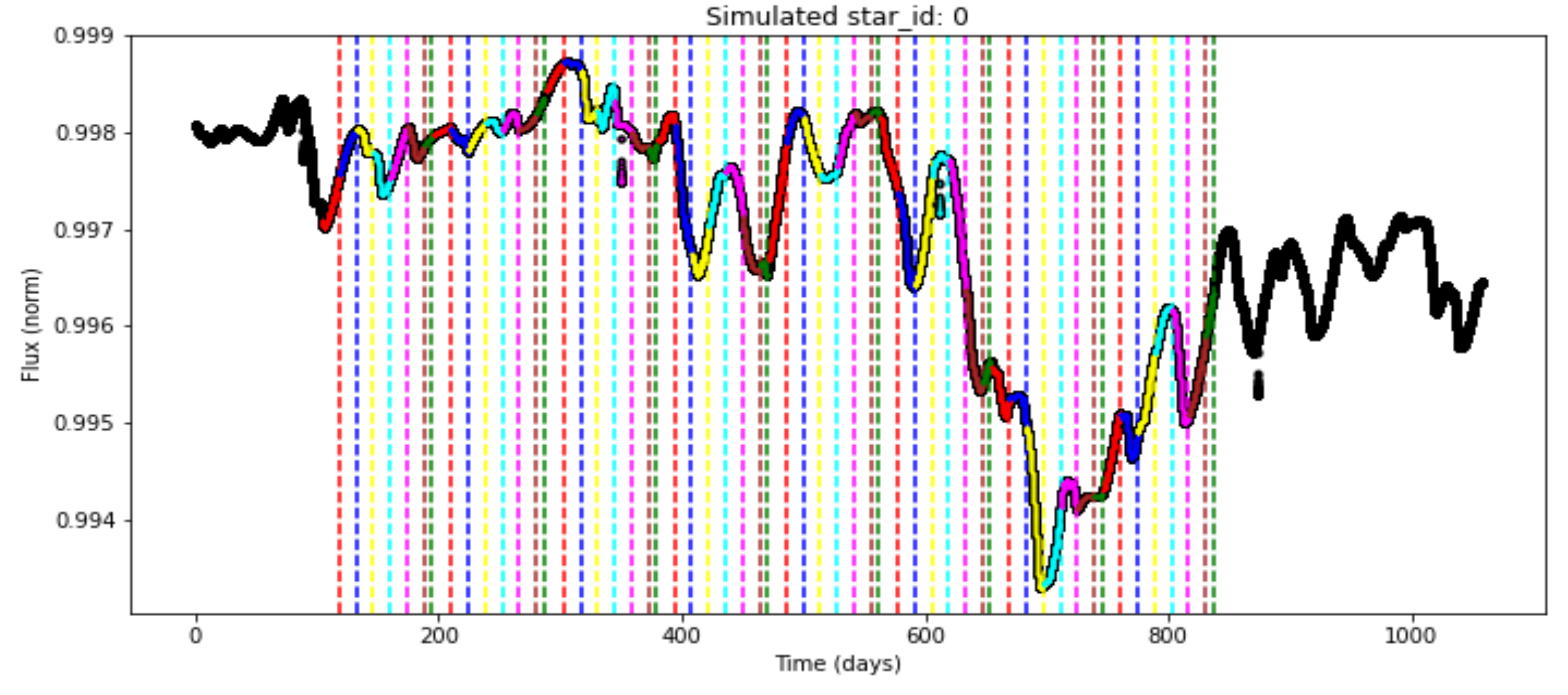
Multi-quarter Stellar & Transit Signal Injection

- Signals simulated by S. Aigrain (Oxford)
- 1250 stars
- 3 years duration
- 10 mins cadence
- 1-3d gaps between quarters
- Real stellar signals from modelling active regions etc.
- Many injected transit signals

Randomly selected slices to inject into PIS/PSLS light curves



bm injection



ubm injection

PDC-MAP vs PLATO Republic

Exploit the multi-camera approach in PLATO knowing that **the stellar signal is common across all the cameras**, while the systematics in each camera are *expected* to be different.

How PDC-MAP works

- Treat each camera individually
- Compare every star to every other star
- Assumes systematics dominate the variance in the data of quiet stars
- Tends to remove any systematics on timescales similar to the duration of a segment.

Whereas Republic

- Treats each star individually
- Compare the light curves from different cameras
- Model the stellar signal alongside the systematics
- Signals common across the multiple cameras for a given star will tend to be preserved (regardless of timescale)
- Several flavours of Republic, focused on version 1 so far

Republic developed by Suzanne Aigrain and Oscar Barragán, University of Oxford

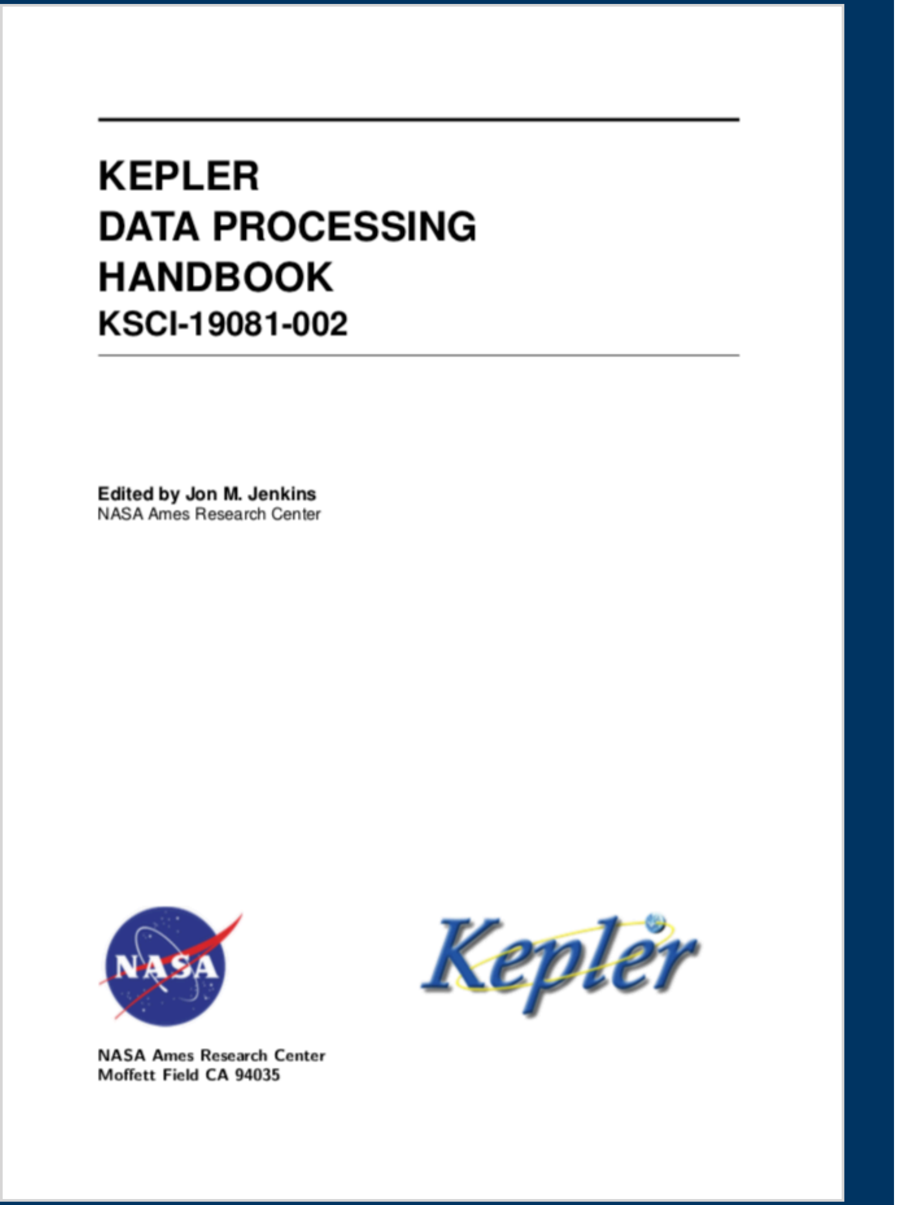
PLATO P-samples

Sample	target V mag	noise ppm in 1 h	type	no targets	sampling sec	data
P1	≤ 11	≤ 50	F5-K7	$\geq 15\ 000$	25	imagettes
P2	≤ 8.5	≤ 50	F7-K7	$\geq 1\ 000$	25	imagettes
P4	≤ 16		M	$\geq 5\ 000$	25	lightcurve/imagettes
P5 including:	≤ 13		F5-K7	$\geq 245\ 000$ ≥ 10 % ≥ 5 % ≥ 9000	600 50 50 25	lightcurve lightcurve centroids imagettes

Table 3 Summary of PLATO stellar sample requirements on dwarf and sub-giant targets.

Detrending: removal of low frequency trends via an arbitrary filter, leaving only high-frequency content

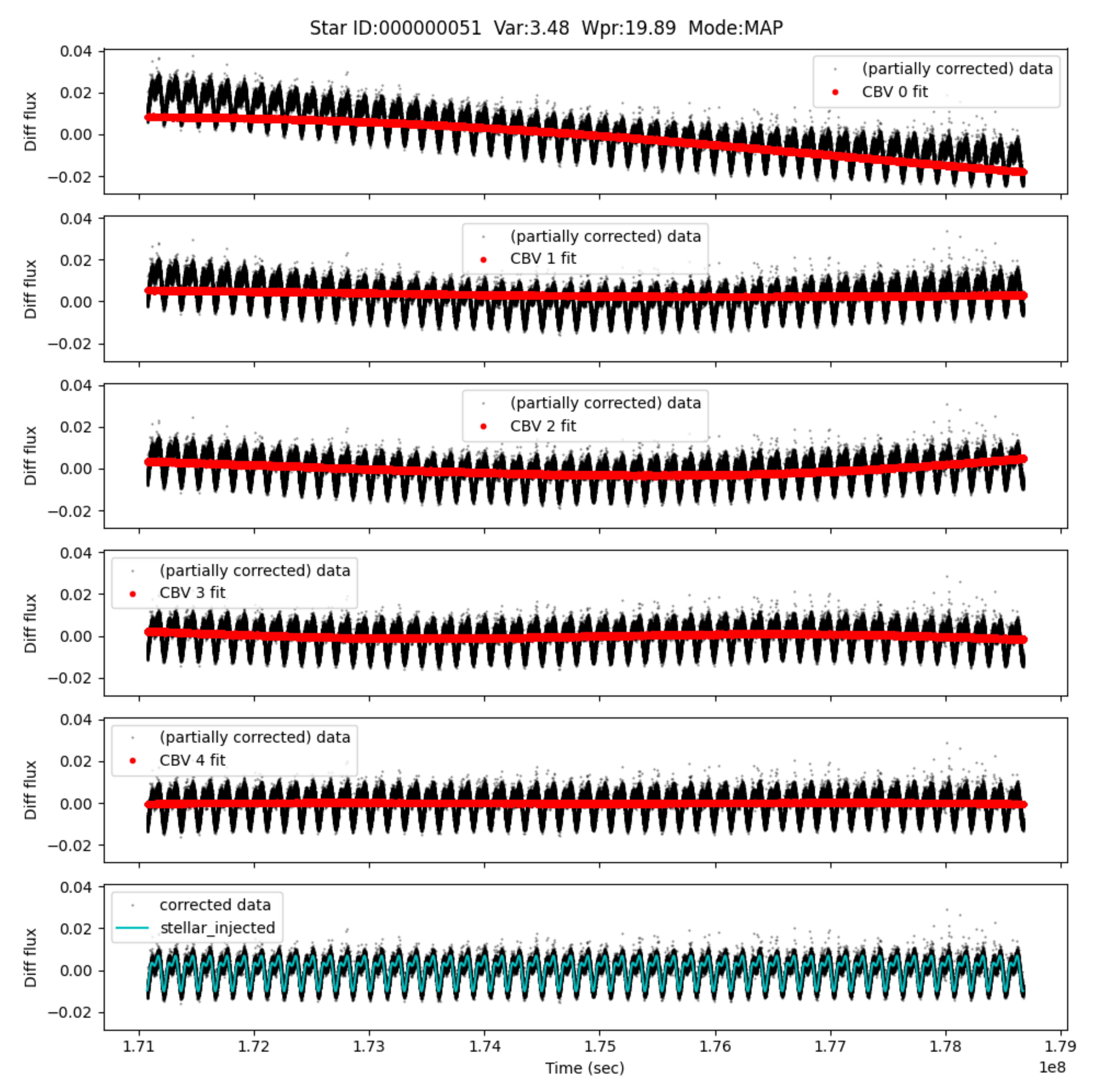
Cotrending: the projection of systematic effects' time series from a target light curve algebraically


KEPLER
DATA PROCESSING
HANDBOOK
KSCI-19081-002

Edited by Jon M. Jenkins
NASA Ames Research Center

NASA Ames Research Center
Moffett Field CA 94035

RoAP Stable Calibrator

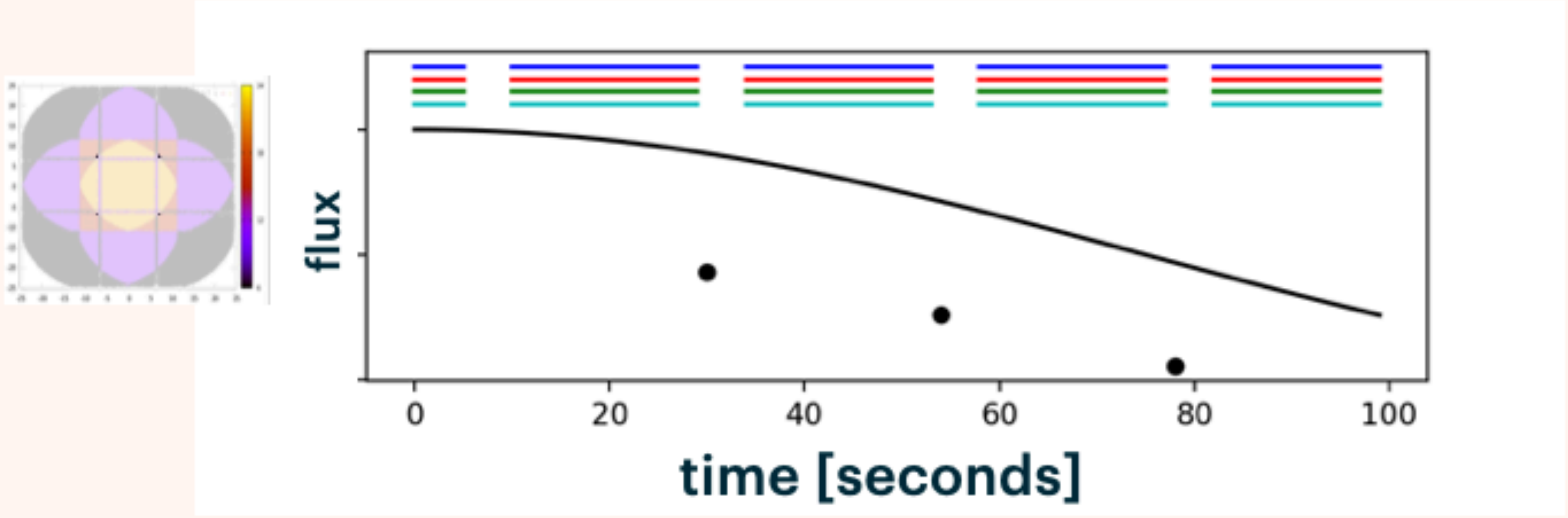


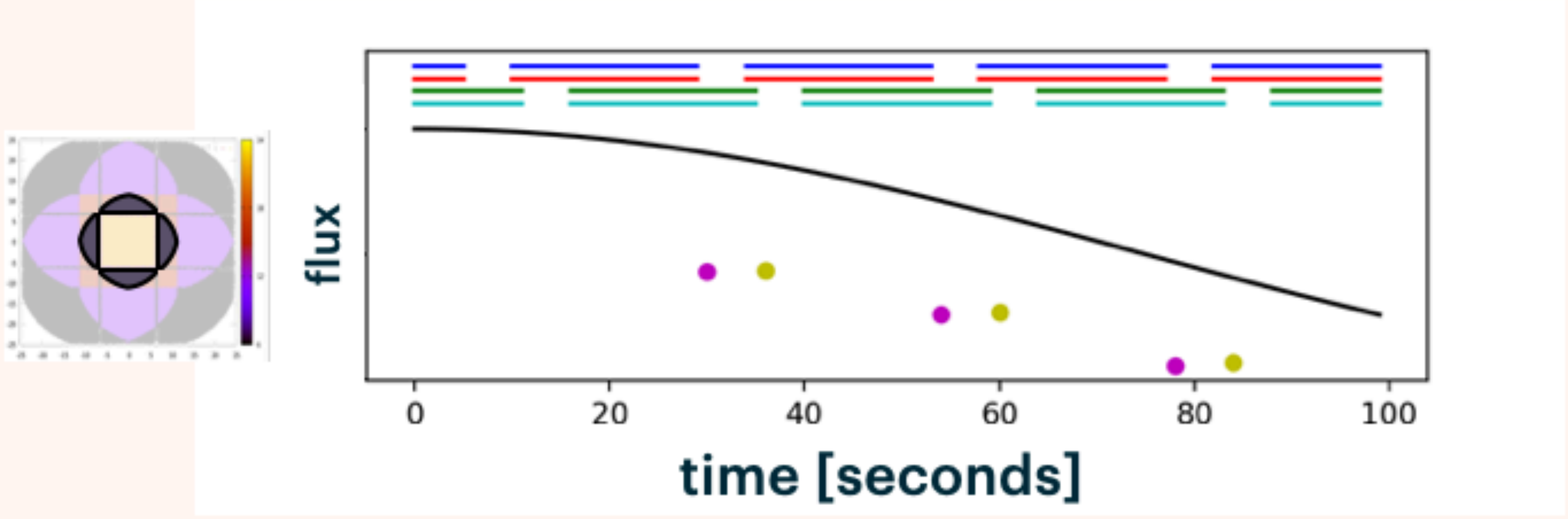
MPS

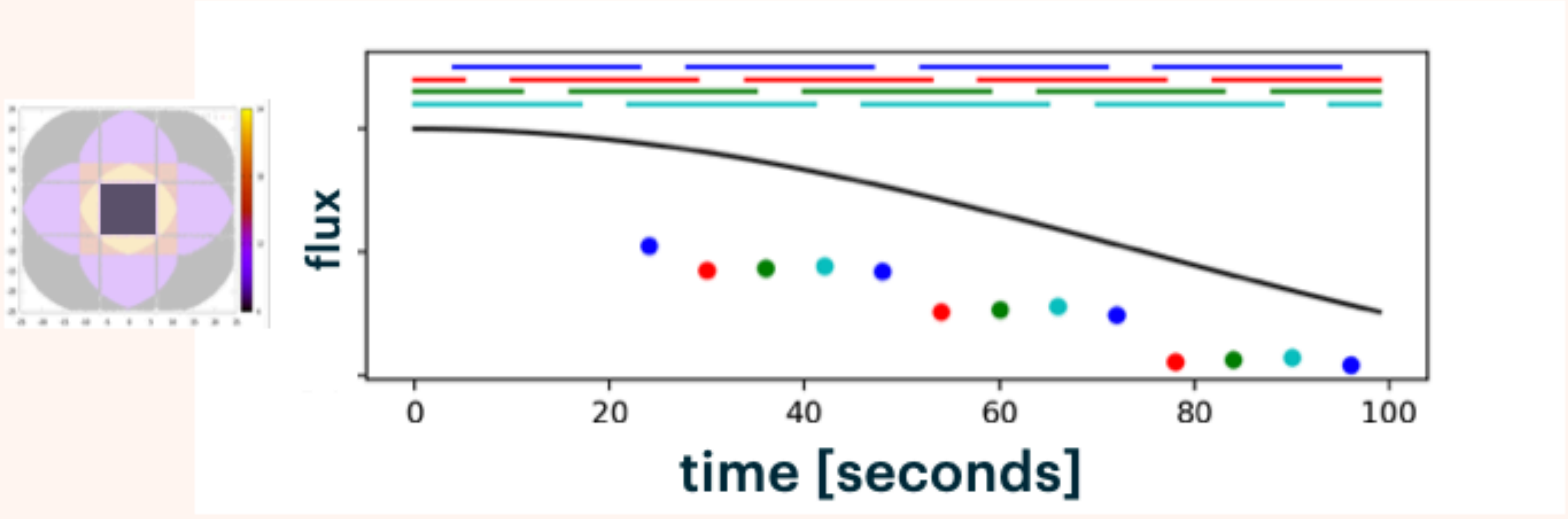
MAX PLANCK INSTITUTE
FOR SOLAR SYSTEM RESEARCH

MERGED LIGHT CURVES

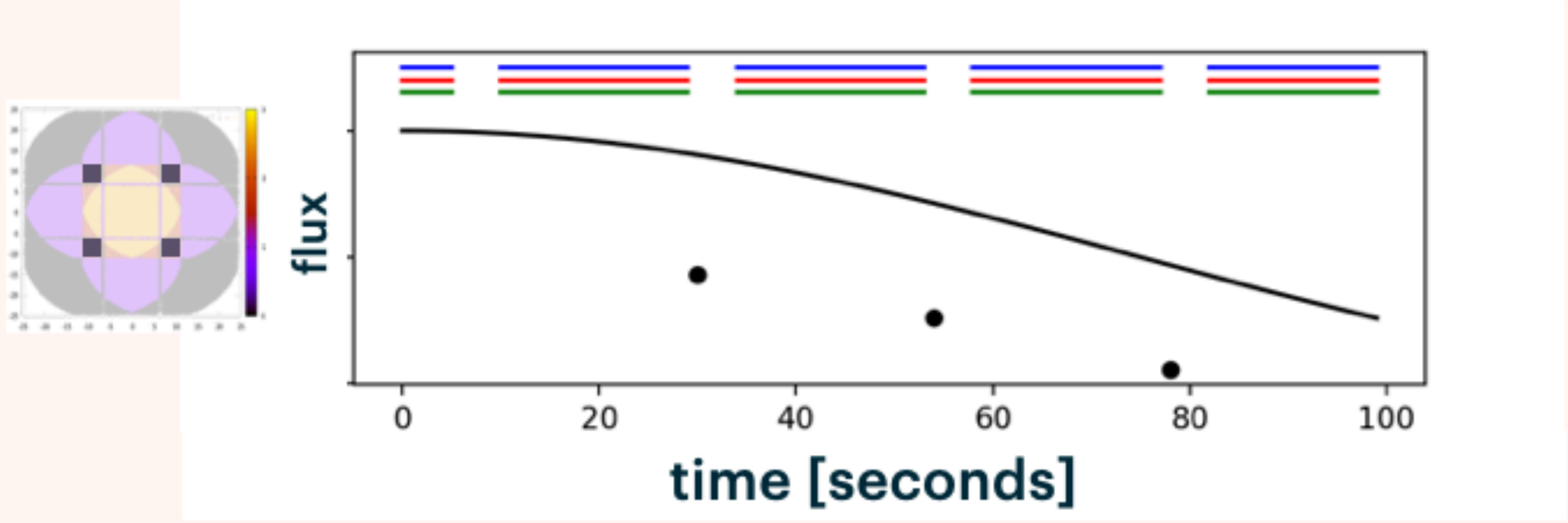
24 cameras

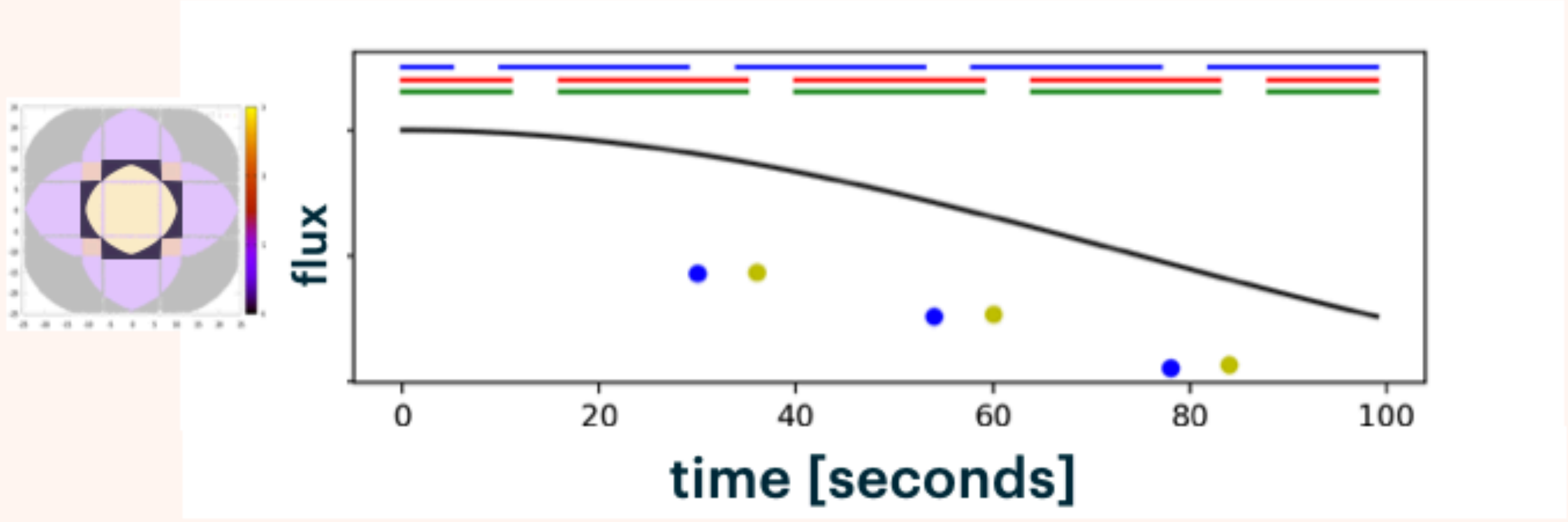






18 cameras





12 cameras