Lecture 3: Ensemble photometry and transit detection

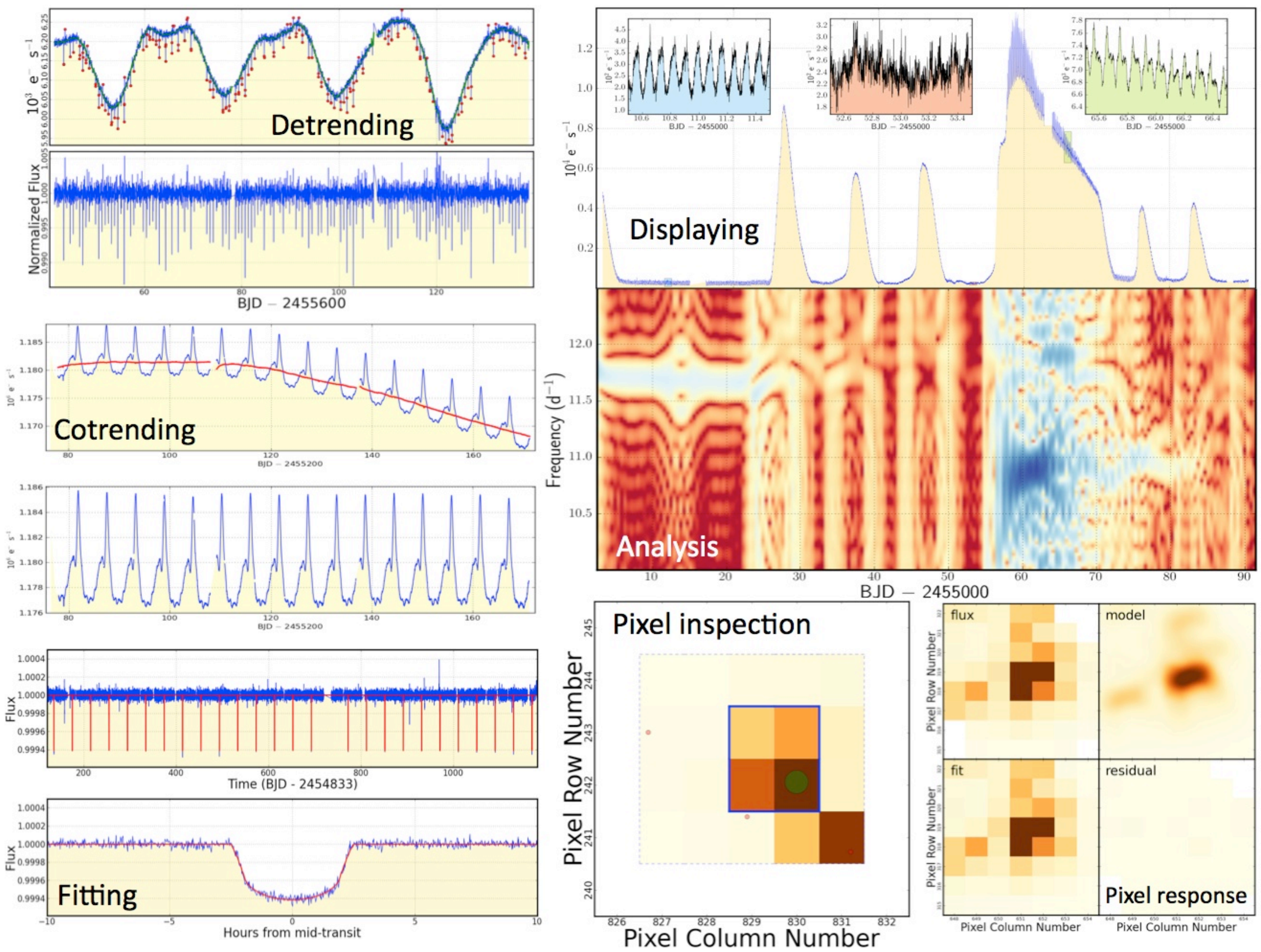
Ensemble differential photometry

Systematics and decorrelation methods

Transit detection by box least-squares

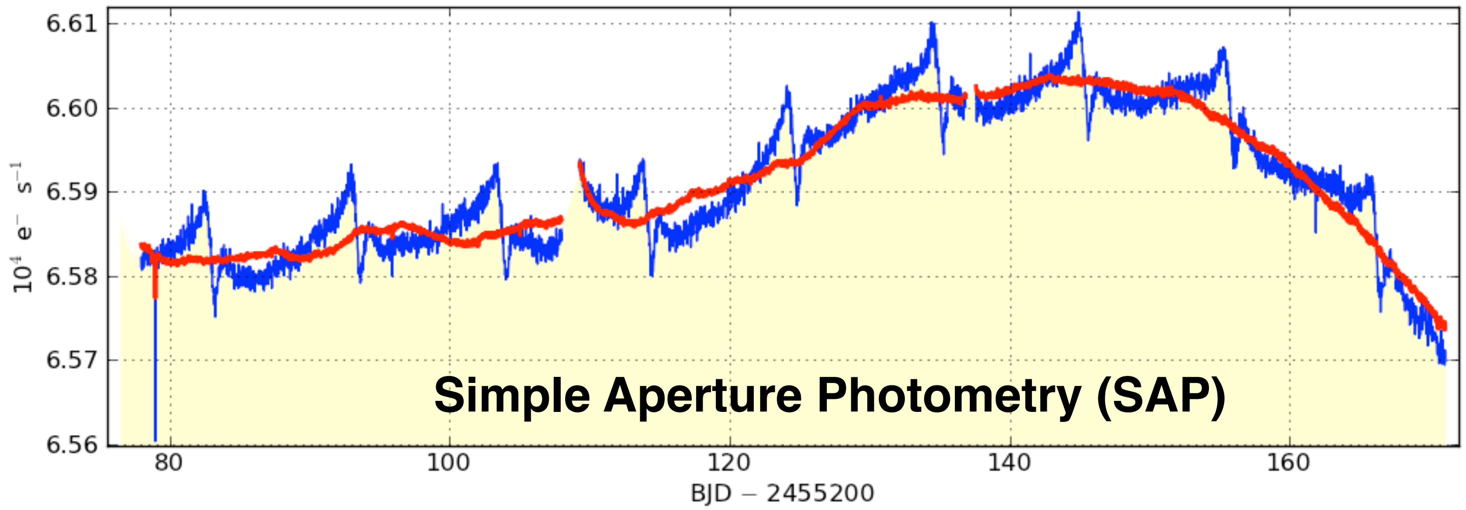
Noise and completeness

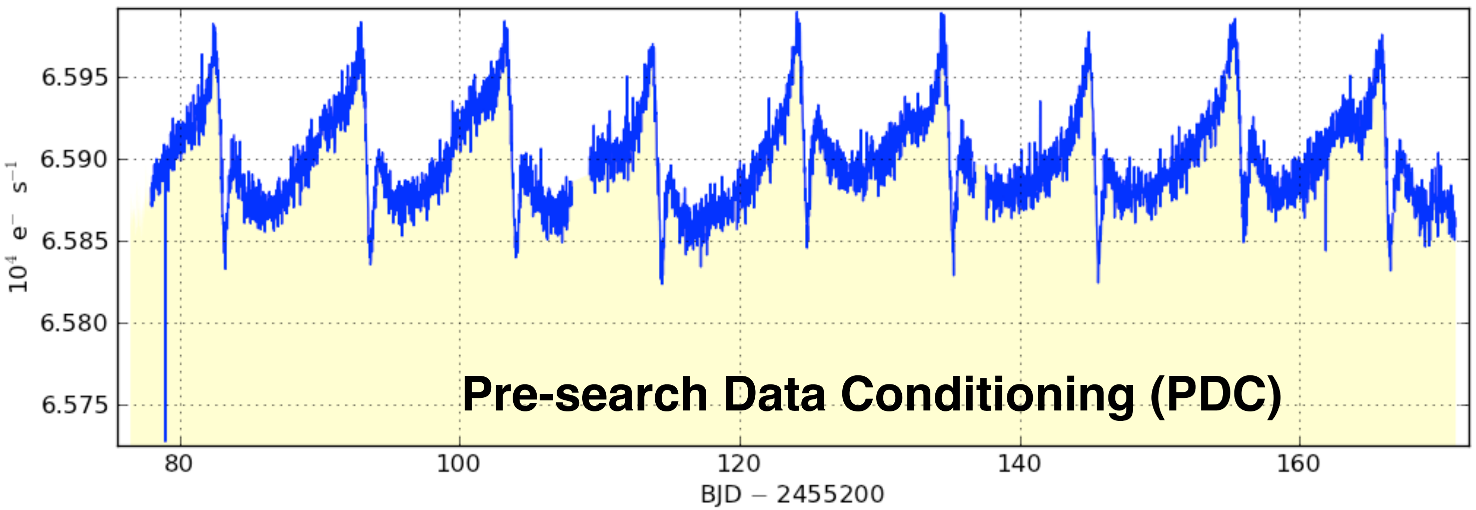
Detrending – the PyKE way



http://keplerscience.arc.nasa.gov/PyKE.shtml

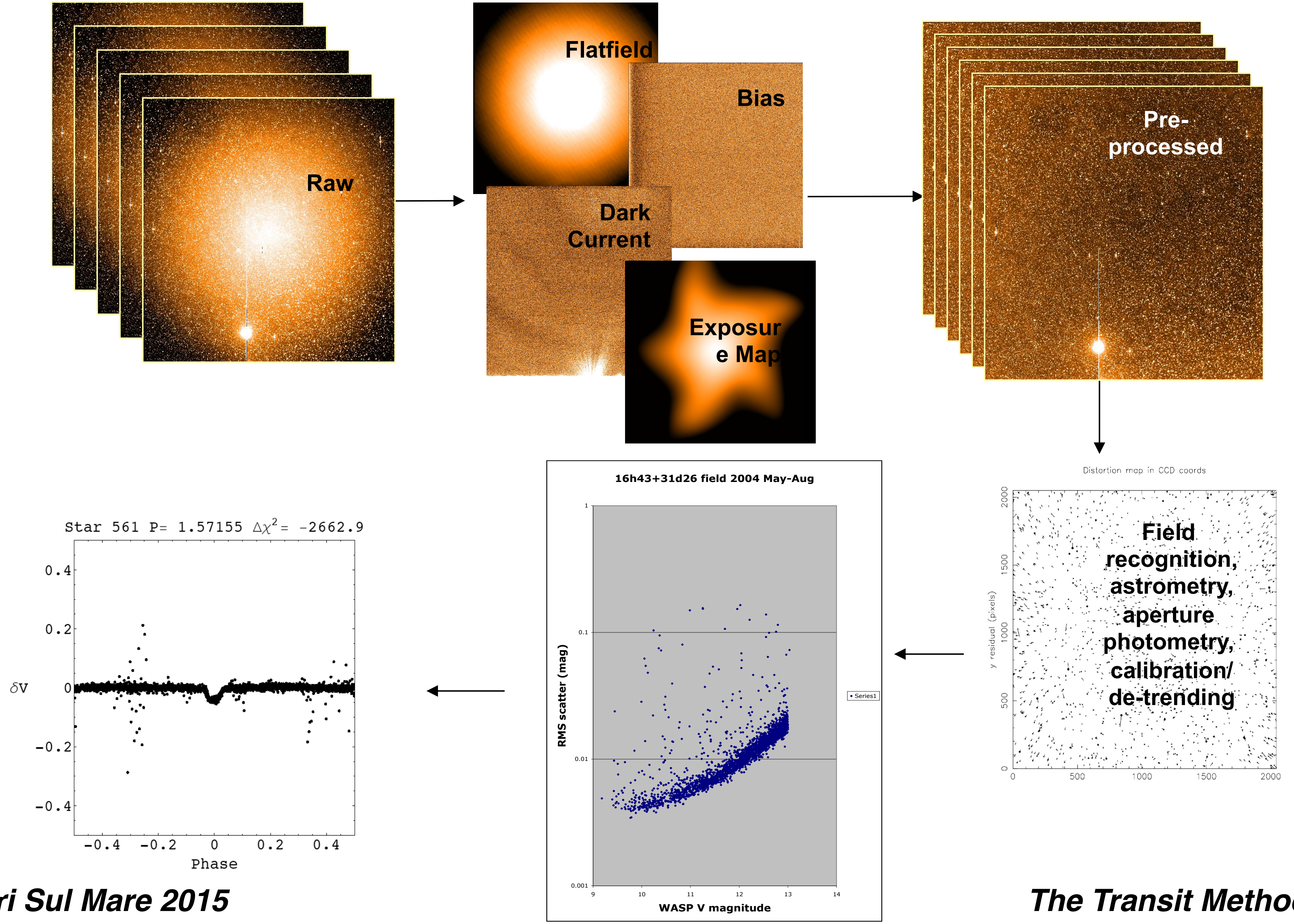
Kepler SAP and PDC-SAP





http://keplerscience.arc.nasa.gov/DataAnalysisProducts.shtml

WASP data reduction pipeline



Catalogue-driven photometry

- **Step 1: Object detection**
  - e.g (S)EXTRACTOR
  - Catalogue all objects on frame to detection threshold
- **Step 2: Establish astrometric solution**
  - Many standard algorithms available, e.g. astrometry.net
  - Cross-match with astrometric catalogue
  - Write world coordinate system to image.
- **Step 3: Aperture photometry**
  - Optimise aperture – many ways to do this!
  - Use frame catalogue to exclude faint objects in sky annulus

Ensemble differential photometry

- **Mean magnitude of star j:**

$$\hat{m}_j = \frac{\sum_i m_{ij} w_{ij}}{\sum_i w_{ij}} \qquad w_{ij} = \frac{1}{\sigma_{ij}^2 + \sigma_{t(i)}^2}$$

Additional intra-frame variance (downgrades poor images)

- **Zero-point correction for frame i:**

$$\hat{z}_i = \frac{\sum_j (m_{ij} - \hat{m}_j) u_{ij}}{\sum_j u_{ij}} \qquad u_{ij} = \frac{1}{\sigma_{ij}^2 + \sigma_{s(j)}^2}$$

Additional stellar variance (downgrades variable objects)

Frame quality and stellar variability I

- **Define data vector X and model μ:**

$$\boldsymbol{X} = \{m_{ij}, i = 1 \ldots n\} \qquad \boldsymbol{\mu} = \{\hat{m}_j + \hat{z}_i, i = 1 \ldots n\}$$

- **Assuming Gaussian errors:**

$$P(\boldsymbol{X}_i|\boldsymbol{\mu}_i) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2}} \times \exp\left\{-\frac{(m_{ij} - \hat{m}_j - \hat{z}_i)^2}{2\left[\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2\right]}\right\}$$

- **Likelihood of entire data vector for star j:**

$$L(\boldsymbol{\mu}) = (2\pi)^{-n/2} \prod_i \left[\frac{1}{\sqrt{\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2}}\right] \exp\left(-\frac{1}{2}\chi^2\right)$$

- **Where:**

$$\chi^2 = \sum_i \frac{(m_{ij} - \hat{m}_j - \hat{z}_i)^2}{\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2}$$

Frame quality and stellar variability II

- **Maximise likelihood w.r.t. $\sigma^2_{s(j)}$ and $\sigma^2_{t(i)}$:**
  - **Solve iteratively for $\sigma^2_{s(j)}$ holding $\sigma^2_{t(i)}$ constant:**

$$\sum_i \frac{1}{\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2} - \sum_i \frac{(m_{ij} - \hat{m}_j - \hat{z}_i)^2}{\left[\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2\right]^2} = 0$$

  - **Solve iteratively for $\sigma^2_{t(i)}$ holding $\sigma^2_{s(j)}$ constant:**

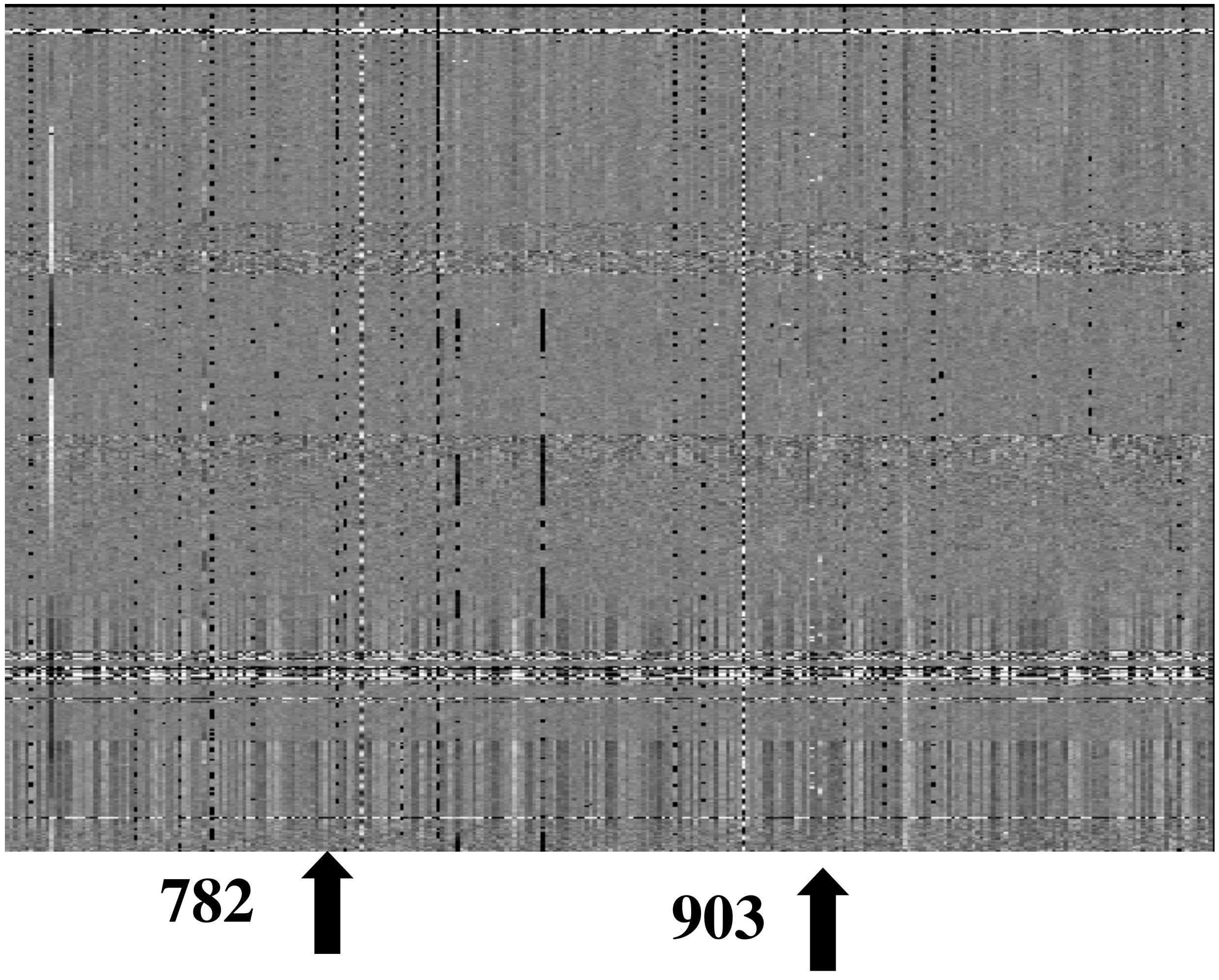
$$\sum_j \frac{1}{\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2} - \sum_j \frac{(m_{ij} - \hat{m}_j - \hat{z}_i)^2}{\left[\sigma_{ij}^2 + \sigma_{t(i)}^2 + \sigma_{s(j)}^2\right]^2} = 0$$

- **Iterate $\hat{m}_j$, $\hat{z}_i$, $\sigma^2_{s(j)}$ and $\sigma^2_{t(i)}$ to convergence.**

Decorrelation/ systematics removal

Tamuz et al 2005

16h30+28 field
300 stars
2549 observations
100 days



Aspen 2005

Systematics and correlated noise

- **References:**
  - **Tamuz, Mazeh & Zucker 2005, MNRAS 355, 1466 (SysRem)**
  - **Kovacs, Bakos & Noyes 2005, MNRAS 356, 537 (TFA)**
  - **Pont, Zucker & Queloz 2006, MNRAS 373, 231**
  - **Carter & Winn 2009, ApJ 704, 51**
  - **Gibson 2014, MNRAS 445, 3401**

SysRem

- **Many patterns of systematics are common to all stars**
  - **Secondary extinction: dependence on stellar colour**
  - **Sky brightness: Dependence on target brightness**
  - **Ambient temperature and focus drift: position dependent?**

- **SysRem: PCA with error bars!**
  - **Construct common temporal basis functions**
  - **Compute optimal scaling coefficient per star**
  - **Remove and repeat.**

SysRem algorithm

- **SysRem produces a corrected magnitude given by:**

$$\tilde{x}_{i,j} = x_{i,j} - \sum_{k=1}^{M} {}^{(k)}c_j \, {}^{(k)}a_i$$

- **Search for the best c_j that minimises:**

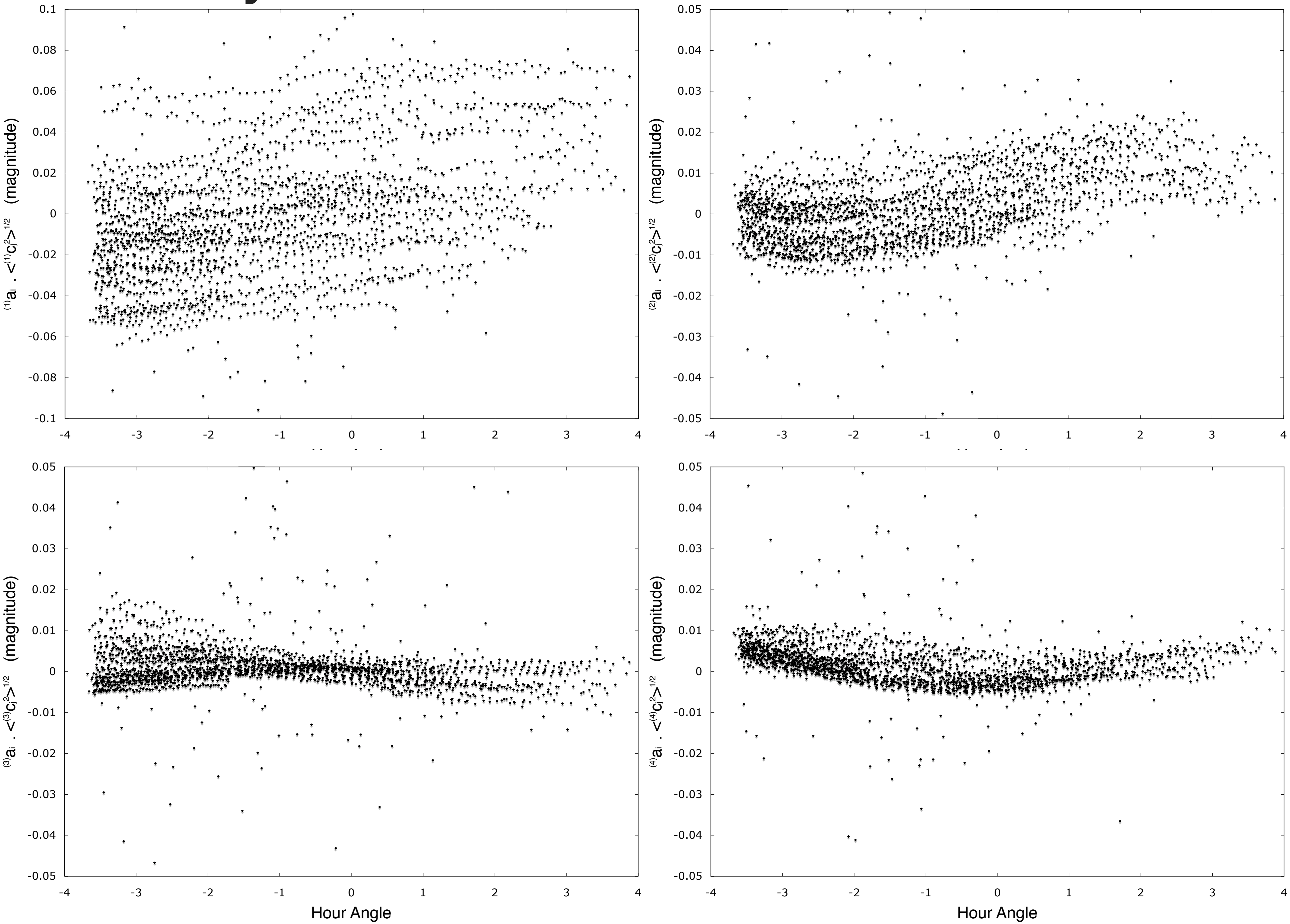
$${}^{(k-1)}S_i^2 = \sum_{i,j} ({}^{(k-1)}\tilde{x}_{i,j} - {}^{(k)}c_i \, {}^{(k)}a_j)^2 w_{i,j}$$

$${}^{(k)}c_j = \sum_i \frac{{}^{(k-1)}\tilde{x}_{i,j} \, {}^{(k-1)}a_i w_{i,j}}{{}^{(k-1)}a_i^2 w_{i,j}}$$

- **Similarly**

$${}^{(k)}a_i = \sum_j \frac{{}^{(k-1)}\tilde{x}_{i,j} \, {}^{(k-1)}c_j w_{i,j}}{{}^{(k-1)}c_j^2 w_{i,j}}$$

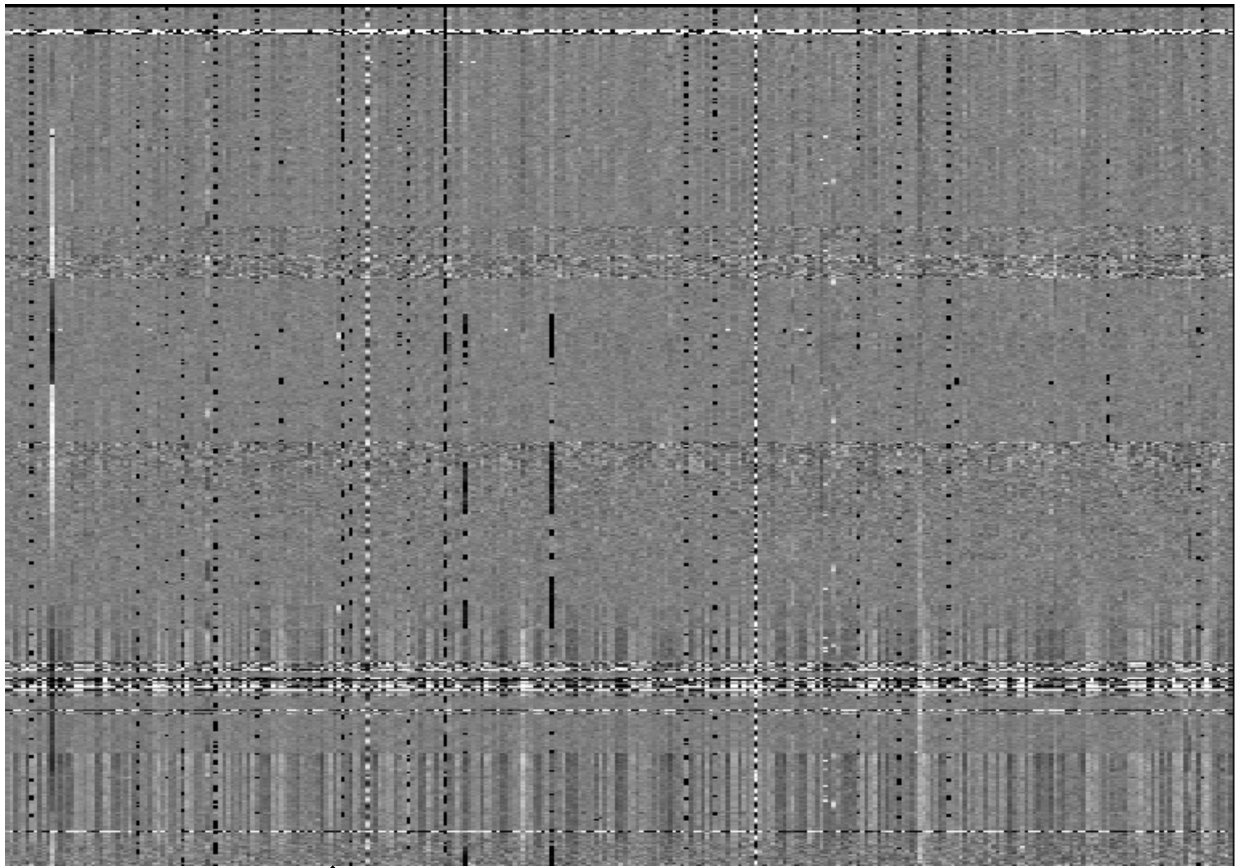
SysRem basis functions vs HA



Decorrelation/ systematics removal

Tamuz et al 2005

16h30+28 field
300 stars
2549 observations
100 days

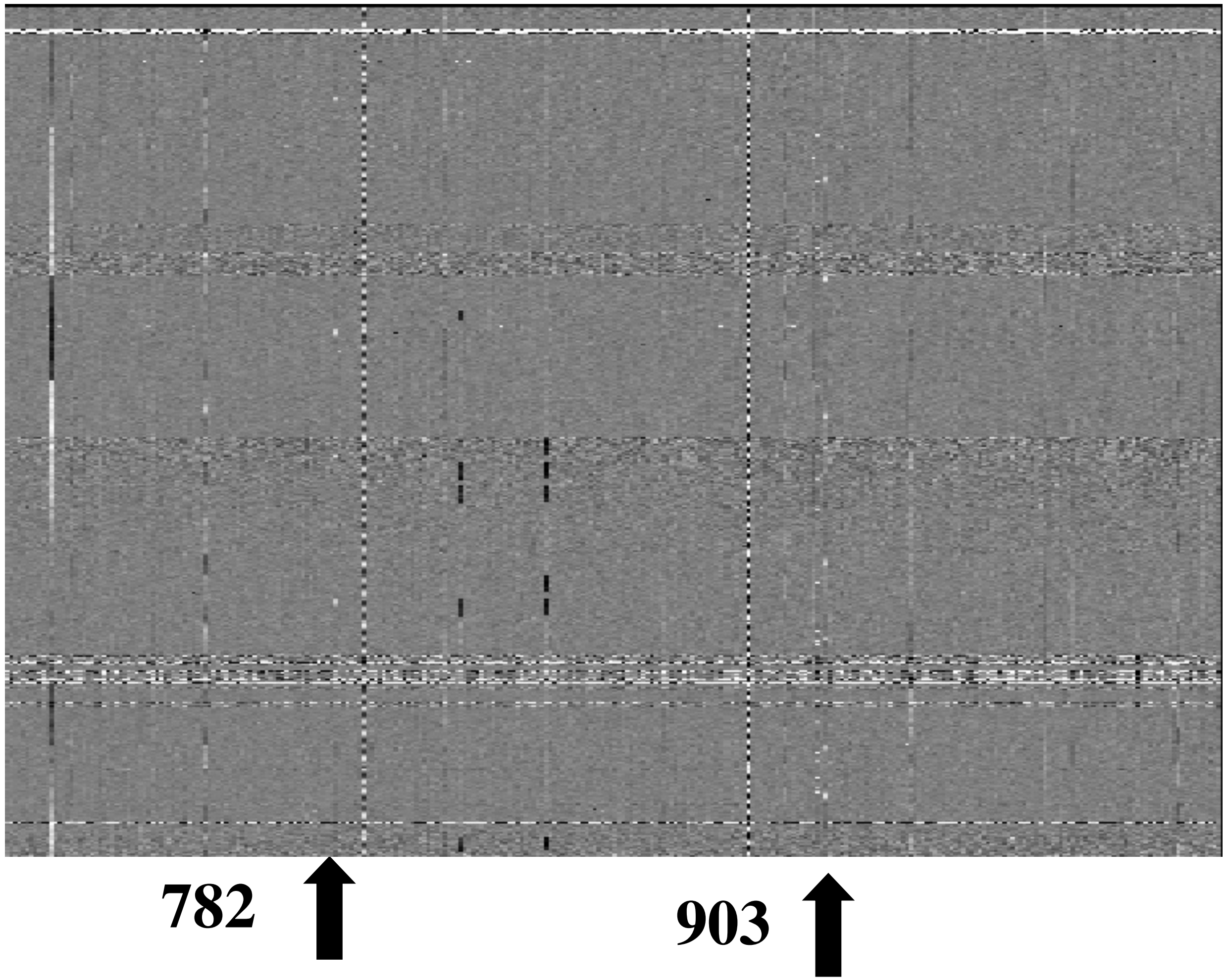


782 903

Decorrelation/ systematics removal

Tamuz et al 2005

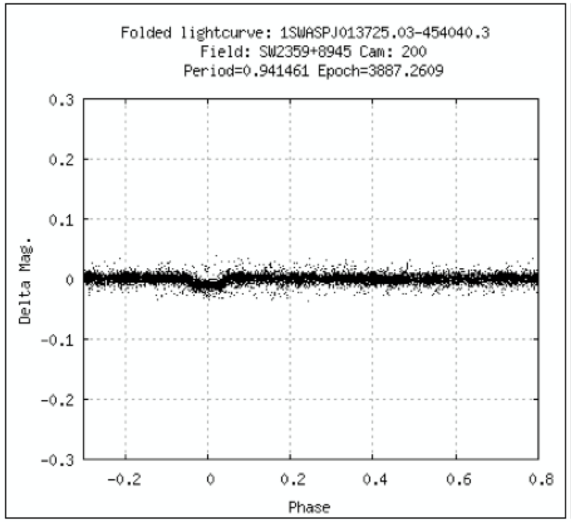
16h30+28 field
300 stars
2549 observations
100 days

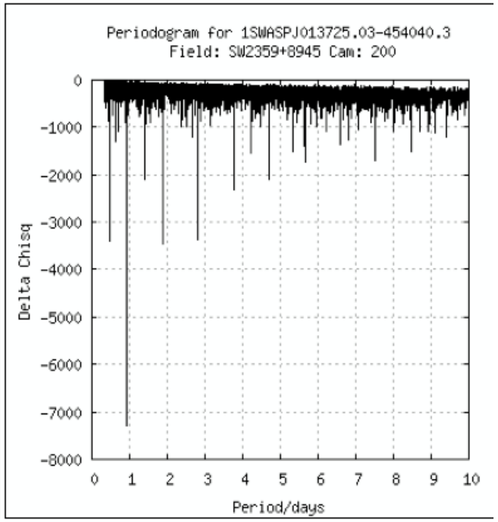


Decorrelation & Detection

- **Decorrelation:**
  - 4-stage SysRem
  - TFA with reconstruction
- **Detection:** Cameron et al 2006 MNRAS 373, 799
  - Accelerated BLS algorithm
  - Coarse grid search
  - Newton-Raphson peak-up on 5 strongest peaks
  - Box width varies with period
- **Example: WASP-1**

i	Teff JH	Teff VK	R_{star} JH	R_{star} VK	V mag	J mag	V-K	J-H	MU_RA (mas/yr)	MU_DEC (mas/yr)	RPMJ	RPMJ diff	Giant?	Dw:Gi	Dil. V	Dil. R
	6410 = F6	6367 = F6	1.32	1.3	9.31	8.41	1.18	0.18	27.4 (± 0.9)	18.6 (± 0.9)	1.01	-19.18	0	72:0	0 %	0 %





Period (days): 0.941461 Epoch i: 3887.2609 Re-plot

Pk1 | Pk2 | Pk3 | Pk4 | Pk5 | MCMC | Updated
Px2 | P/2 | Zoom | Transits

Detrended Data | Full archive LC

i	Period	Epoch	Width (hr)	Depth (mag)	Del Chisq	S/N Red	N_tr	Ellip Var	Ellip S/N	Frac. in Tr.	SN Anti	R_{Pl} VK	R_{Pl} JH	sig det eff	H_match i	Clmp	All
	0.941461	3887.2609	2.0050	-0.0116	8762.5723	-18.8831	55	0.0007	5.8654	0.0466	23.8035	1.2	1.22	85.186	200	0	

i	Period	Epoch	Width (hr)	Depth	Impact	R_{star}	M_{star}	R_{Pl}	Prob pl	Prob MS	Prob imp	Chisq_cs	Chisq_ucs	Q	All	
	0.9414507 ±	4213.00681 ±	2.1578	0.0093	0.5329	1.4269	1.4292	1.3367	0.7053	0.7417	0.955	7688.04	7687.28	0.92	Plot 1	Plot 2

Box Least-Squares transit search

- **References**
  - Kovacs, Zucker & Mazeh 2002, A&A 391, 369
  - Aigrain & Irwin 2004, MNRAS 350, 331
  - Collier Cameron et al 2006, MNRAS 373, 799
- **Step 1: Compute inverse variance-weighted mean flux, subtract and precompute χ^2_0, t :**

$$\hat{x} = \frac{\sum_i \tilde{x}_i w_i}{\sum_i w_i} \qquad x_i = \tilde{x}_i - \hat{x}$$

$$\chi_0^2 = \sum_i x_i^2 w_i \qquad t = \sum_i w_i.$$

- **Step 2: set up frequency grid**
  - phase of each observation must change by less than transit duration between adjacent frequencies
- **Step 3: estimate range of transit duration**
  - Use stellar density

Box Least-Squares transit search

- **Step 3: Phase-fold and partition:**
  - **Mean light level in-transit (L):**

$$s = \sum_{i \in \ell} x_i w_i, \quad r = \sum_{i \in \ell} w_i, \quad q = \sum_{i \in \ell} x_i^2 w_i.$$

$$L = \frac{s}{r} \qquad \mathrm{Var}(L) = \frac{1}{r}$$

  - **Mean light level out of transit (H):**

$$H = \frac{-s}{t - r} \qquad \mathrm{Var}(H) = \frac{1}{t - r}$$

  - **Fitted transit depth (δ):**

$$\delta = L - H = \frac{st}{r(t - r)} \qquad \mathrm{Var}(\delta) = \frac{t}{r(t - r)}$$

Box Least-Squares transit search

- **Step 4: Badness-of-fit:**
  - **Signal-to-noise ratio of transit depth**

$$\mathrm{S/N} = s\sqrt{\frac{t}{r(t-r)}}$$

  - **Improvement in fit relative to constant model is $(\mathrm{S/N})^2$:**

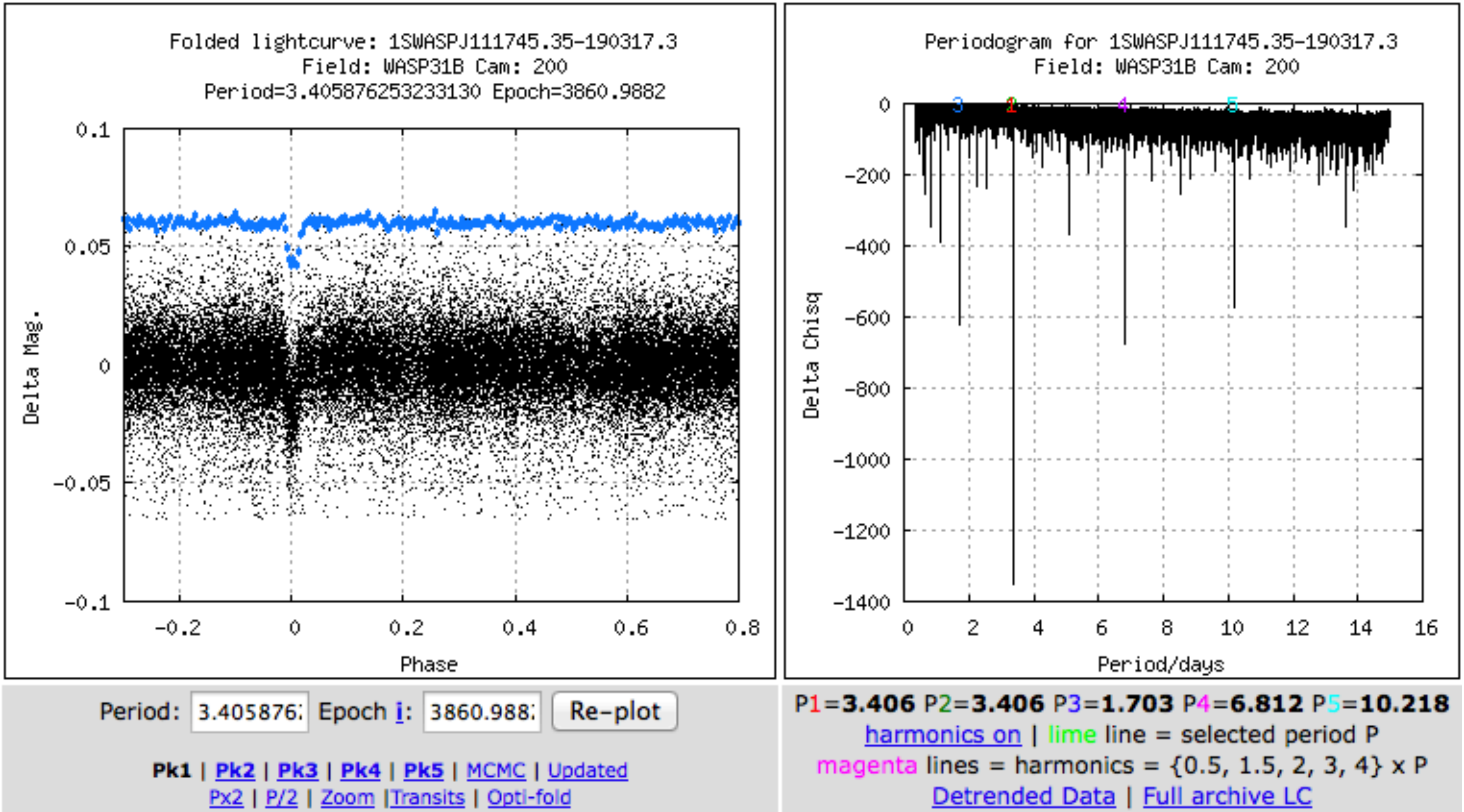
$$\Delta\chi^2 = \frac{s^2 t}{r(t-r)}$$

  - **Goodness of fit outside transit:**

$$\chi_h^2 = \chi_0^2 - \frac{s^2}{(t-r)} - q$$

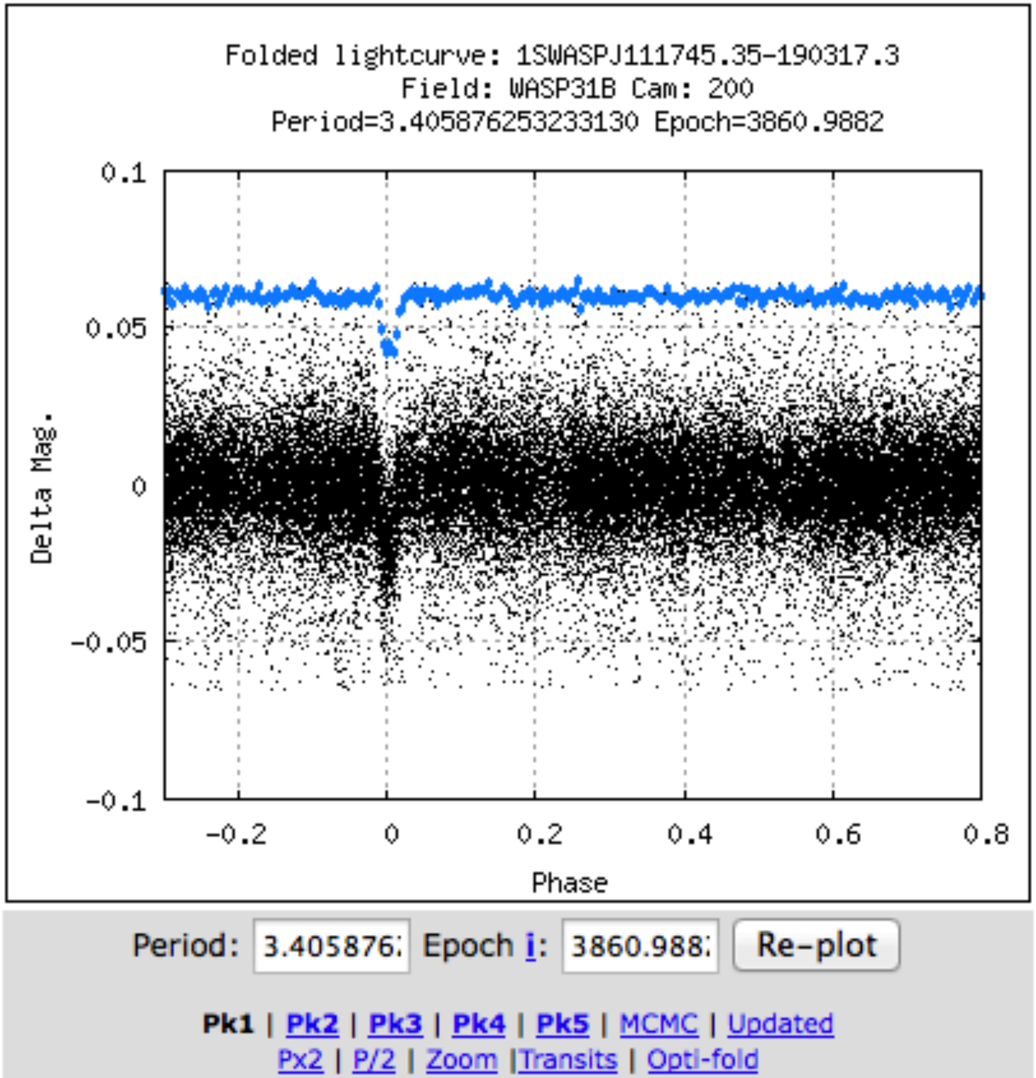
Box Least-Squares transit search

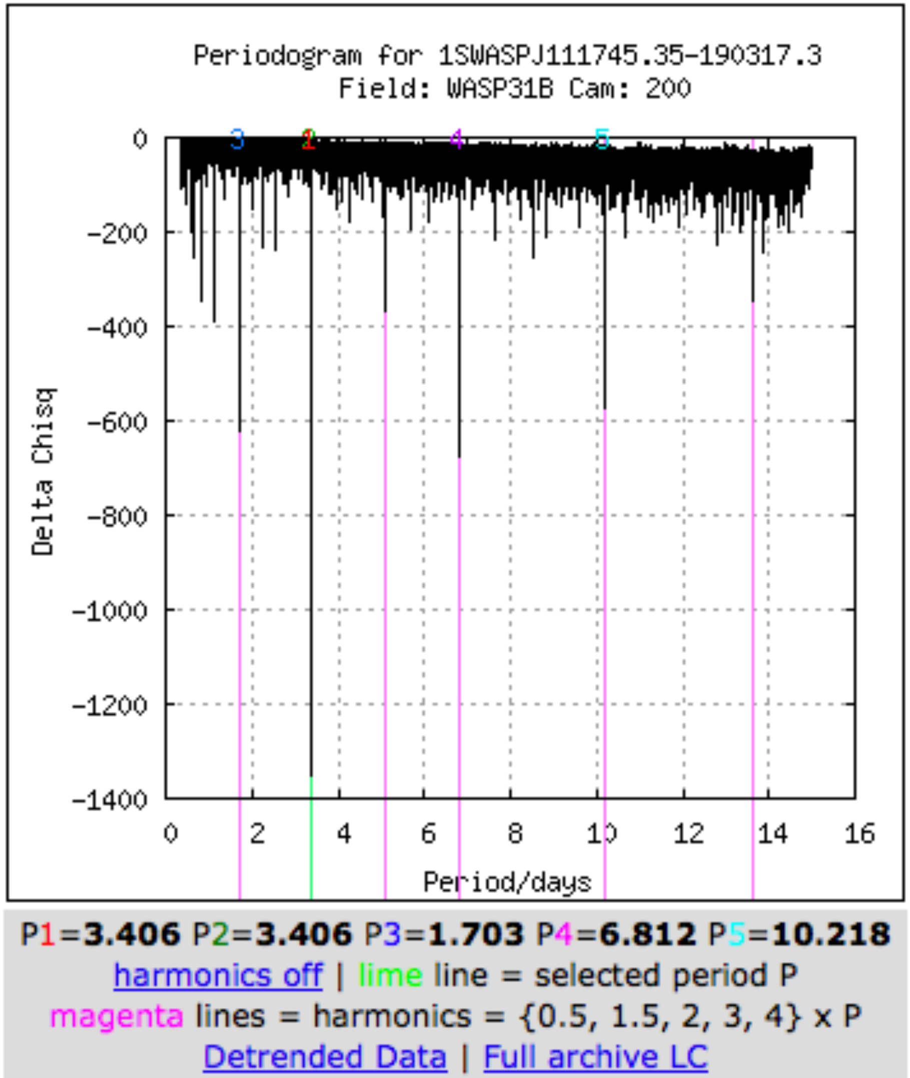
- **Step 5: BLS periodogram**



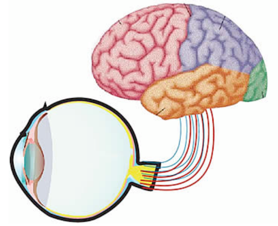
Box Least-Squares transit search

- **Step 5: BLS periodogram**

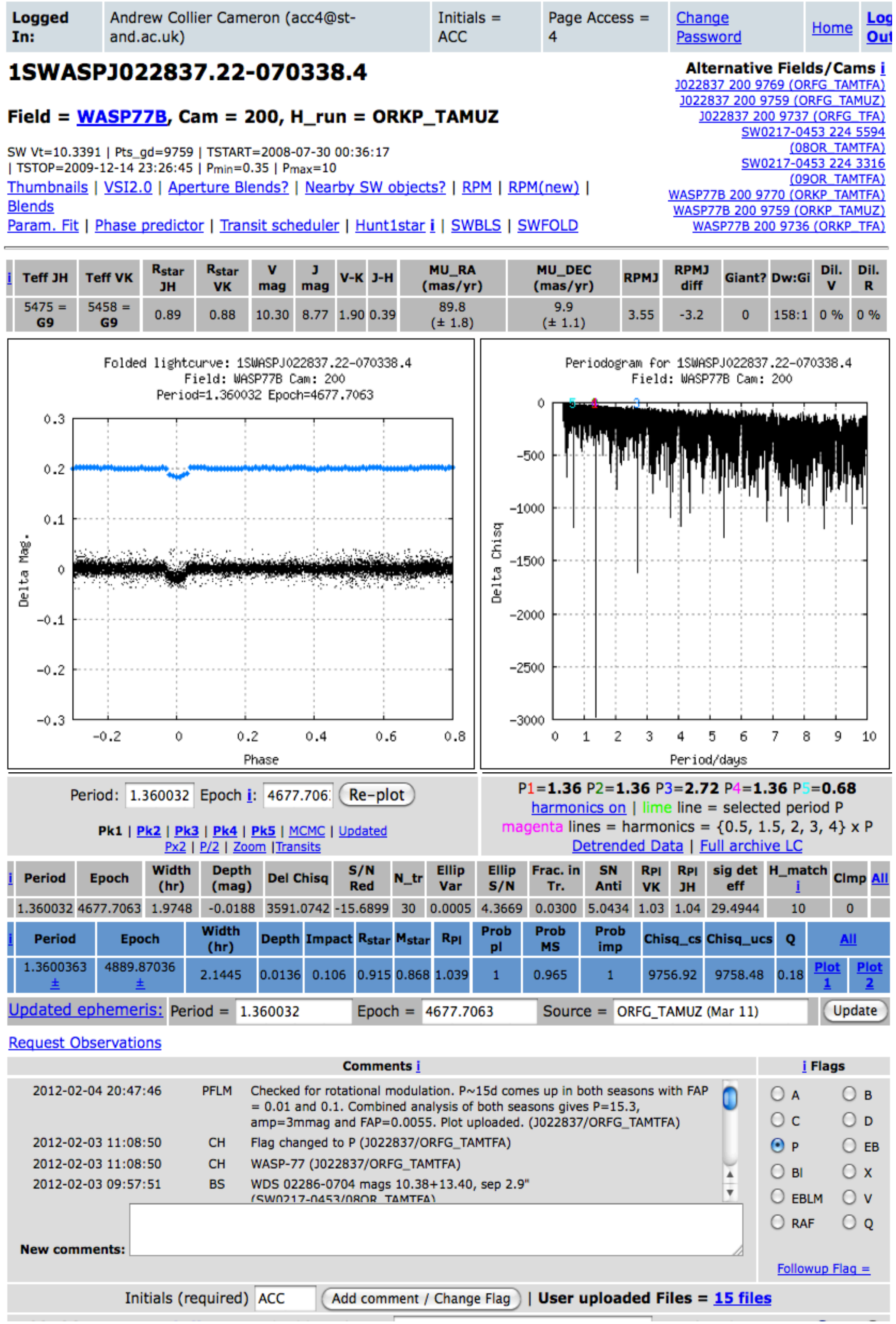




Planet search



- **SysRem+BLS**
  - **Tamuz, Mazeh & Zucker 2005**
  - **Kovacs, Zucker & Mazeh 2002**
- **Skype-linked network of human neuro-visual processors.**
- **Each processor allocated several thousand candidate light curves + associated statistics from main database.**
- **Motivated by promises of food, postdoc employment, fame, glory, etc.**



Logged In:	Andrew Collier Cameron (acc4@st-and.ac.uk)	Initials = ACC	Page Access = 4	Change Password	Home	Log Out

1SWASPJ022837.22-070338.4

Field = WASP77B, Cam = 200, H_run = ORKP_TAMUZ

SW Vt=10.3391 | Pts_gd=9759 | TSTART=2008-07-30 00:36:17 | TSTOP=2009-12-14 23:26:45 | Pmin=0.35 | Pmax=10

Thumbnails | VSI2.0 | Aperture Blends? | Nearby SW objects? | RPM | RPM(new) | Blends

Param. Fit | Phase predictor | Transit scheduler | Hunt1star i | SWBLS | SWFOLD

Alternative Fields/Cams i
J022837 200 9769 (ORFG_TAMTFA)
J022837 200 9759 (ORFG_TAMUZ)
J022837 200 9737 (ORFG_TFA)
SW0217-0453 224 5594 (08OR_TAMTFA)
SW0217-0453 224 3316 (09OR_TAMTFA)
WASP77B 200 9770 (ORKP_TAMTFA)
WASP77B 200 9759 (ORKP_TAMUZ)
WASP77B 200 9736 (ORKP_TFA)

i	Teff JH	Teff VK	Rstar JH	Rstar VK	V mag	J mag	V-K	J-H	MU_RA (mas/yr)	MU_DEC (mas/yr)	RPMJ	RPMJ diff	Giant?	Dw:Gi	Dil. V	Dil. R
	5475 = G9	5458 = G9	0.89	0.88	10.30	8.77	1.90	0.39	89.8 (± 1.8)	9.9 (± 1.1)	3.55	-3.2	0	158:1	0 %	0 %

Period: 1.360032 Epoch i: 4677.706 Re-plot

Pk1 | Pk2 | Pk3 | Pk4 | Pk5 | MCMC | Updated
Px2 | P/2 | Zoom | Transits

P1=1.36 P2=1.36 P3=2.72 P4=1.36 P5=0.68
harmonics on | lime line = selected period P
magenta lines = harmonics = {0.5, 1.5, 2, 3, 4} x P
Detrended Data | Full archive LC

i	Period	Epoch	Width (hr)	Depth (mag)	Del Chisq	S/N Red	N_tr	Ellip Var	Ellip S/N	Frac. in Tr.	SN Anti	RPJ VK	RPJ JH	sig det eff	H_match i	Clmp	All
	1.360032	4677.7063	1.9748	-0.0188	3591.0742	-15.6899	30	0.0005	4.3669	0.0300	5.0434	1.03	1.04	29.4944	10	0	

i	Period	Epoch	Width (hr)	Depth	Impact	Rstar	Mstar	RPJ	Prob pl	Prob MS	Prob imp	Chisq_cs	Chisq_ucs	Q	All	
	1.3600363 ±	4889.87036 ±	2.1445	0.0136	0.106	0.915	0.868	1.039	1	0.965	1	9756.92	9758.48	0.18	Plot 1	Plot 2

Updated ephemeris: Period = 1.360032 Epoch = 4677.7063 Source = ORFG_TAMUZ (Mar 11) Update

Request Observations

Comments i

2012-02-04 20:47:46	PFLM	Checked for rotational modulation. P~15d comes up in both seasons with FAP = 0.01 and 0.1. Combined analysis of both seasons gives P=15.3, amp=3mmag and FAP=0.0055. Plot uploaded. (J022837/ORFG_TAMTFA)
2012-02-03 11:08:50	CH	Flag changed to P (J022837/ORFG_TAMTFA)
2012-02-03 11:08:50	CH	WASP-77 (J022837/ORFG_TAMTFA)
2012-02-03 09:57:51	BS	WDS 02286-0704 mags 10.38+13.40, sep 2.9" (SW0217-0453/08OR_TAMTFA)

New comments:

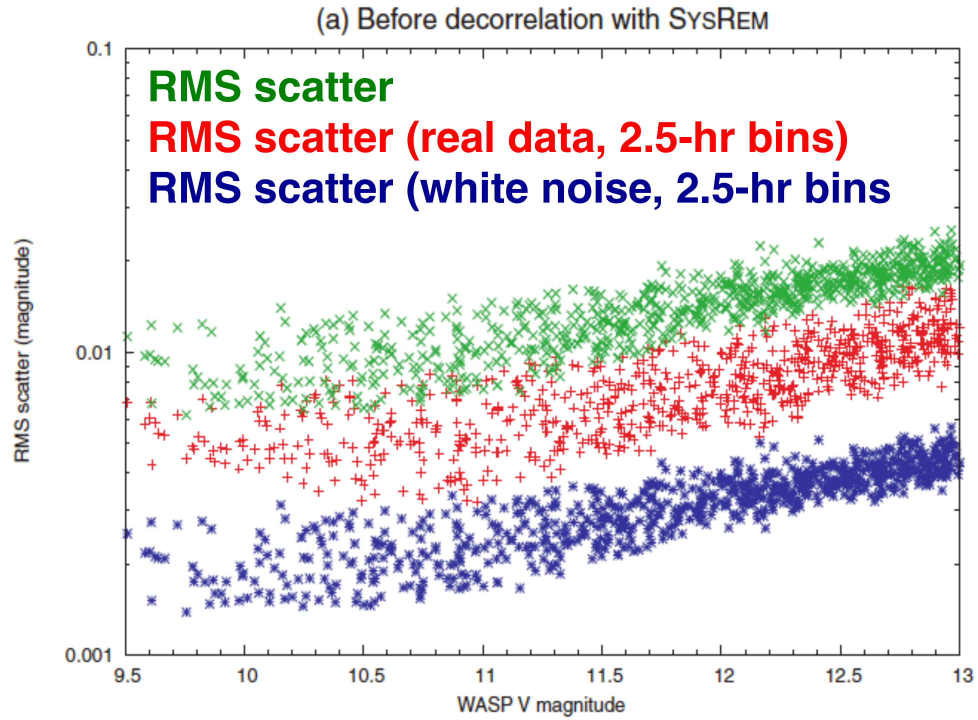
i Flags: A, B, C, D, P, EB, BI, X, EBLM, V, RAF, Q

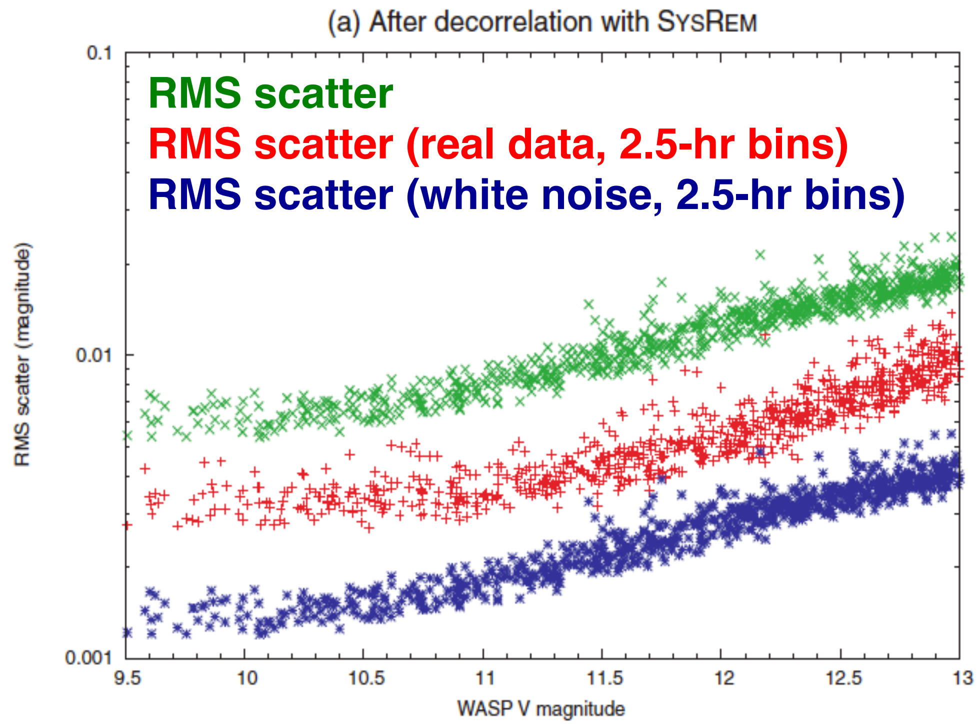
Followup Flag =

Initials (required) ACC | Add comment / Change Flag | **User uploaded Files = 15 files**

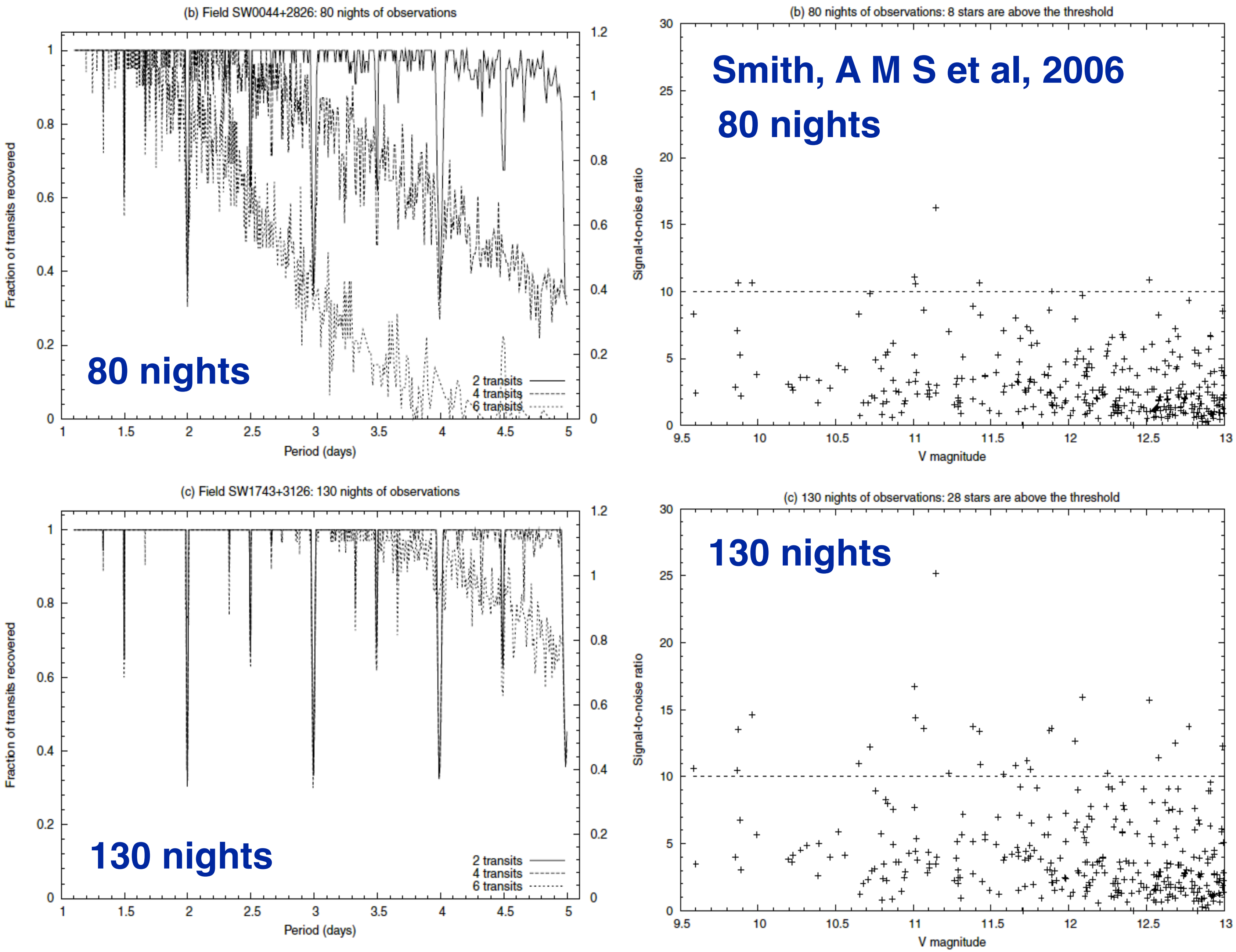
Systematics and red noise

- **Systematics:**
  - Secondary extinction
  - Temperature-dependent focus
  - Sky brightness-dependent bias in background subtraction
  - SysRem: Tamuz et al 2005
  - TFA: Kovacs et al 2005
- **Red noise:**
  - Pont et al 2006
  - Smith et al 2006

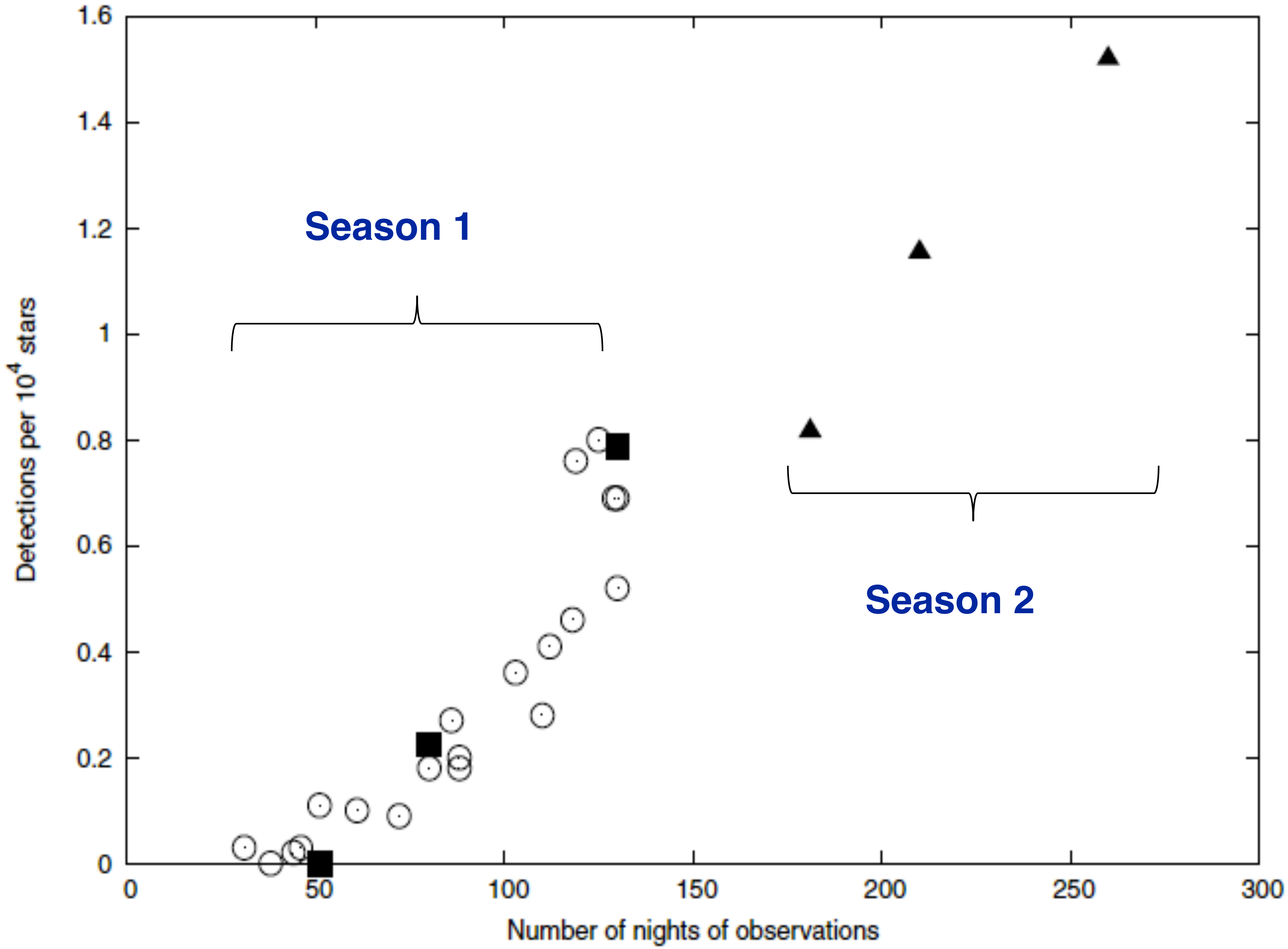




Red noise and detection threshold



Red noise and planet catch



Summary

- **Systematics and astrophysical variability mask transits**
- **Not all systematics are understood**
- **PCA-like methods (e.g. SysRem, PDC-MAP) remove systematics while preserving astrophysical variability**
- **More aggressive detrending (e.g. TFA) removes astrophysical variability too**
- **Box least-squares method is efficient**
- **Correlated noise must be taken into account when assessing detection thresholds.**