IRACSIM: Simulating IRAC Data

Jim Ingalls

(Spitzer Science Center)

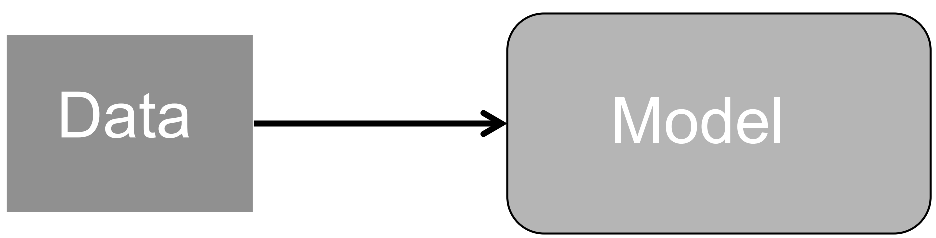
What is a simulation, why do we need it?

- The point of a simulation is insight into a physical system that is too complicated to comprehend in terms of a simple deterministic relationship.
- Simple processes that are relatively easy to explain undergo complex interactions.
- The goal is to explain or elucidate features of the system as a whole by replicating the interactions and controlling their relative magnitudes.
- Randomness/statistical description is usually a key component.

What can we use an IRAC simulator for?

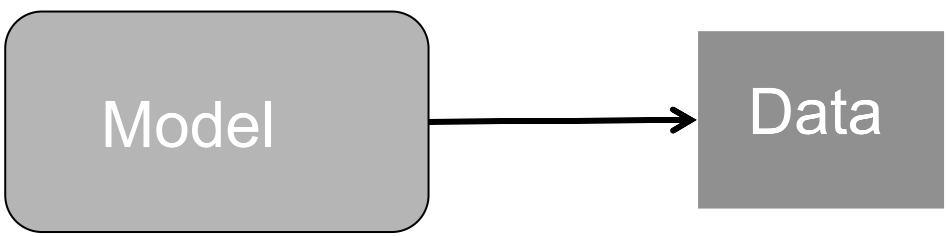
Suzanne Agrain (on planetary imaging):
"Want the planet? Model the instrument"

- High precision requires that we understand something about the instrument
- All techniques described today are ways of modeling the instrument plus a physical system (exoplanet) to solve an inverse problem.
- Use Observations to infer system parameters.



What can we use an IRAC simulator for?

IRAC simulator does the opposite. We use it to examine the system in terms of a forward problem—it starts with a model and predicts data.

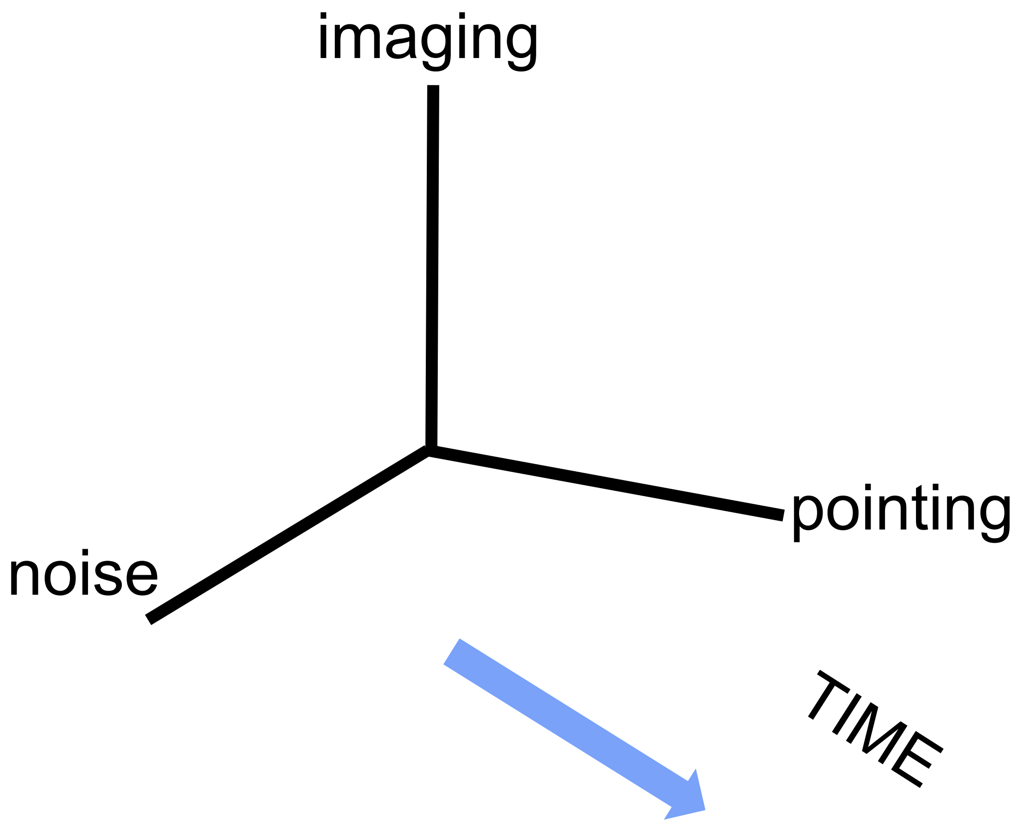


What can we use an IRAC simulator for?

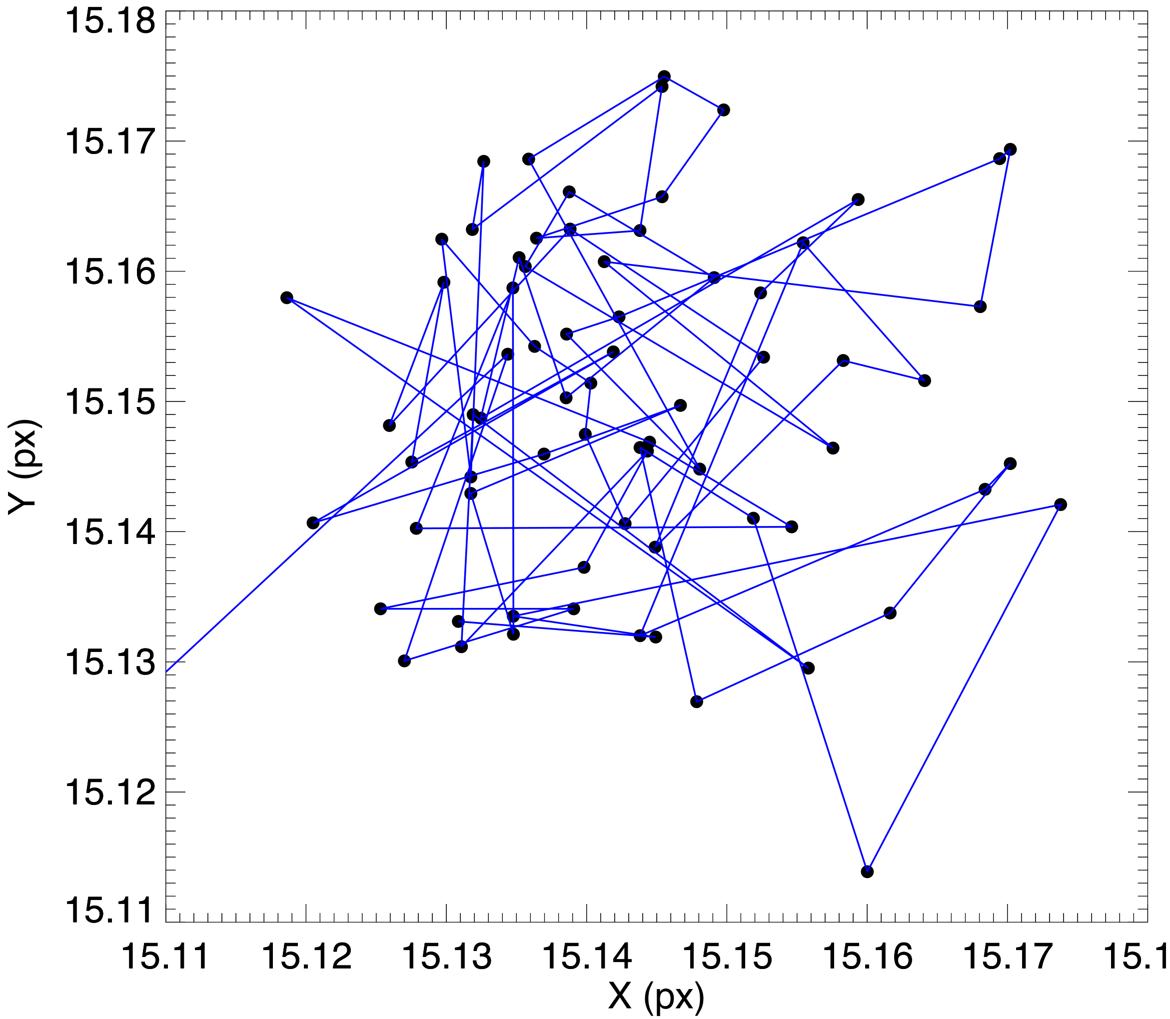
- Centroiding bias
- Intrapixel gain variations coupled with pointing = correlated noise.
- SNR predictions for sources
- Novel observing modes (eg., drift scans)
- **Validate data reduction techniques**

(how do we know we have the right answer if we don't know the answer?)

IRAC axes of behavior

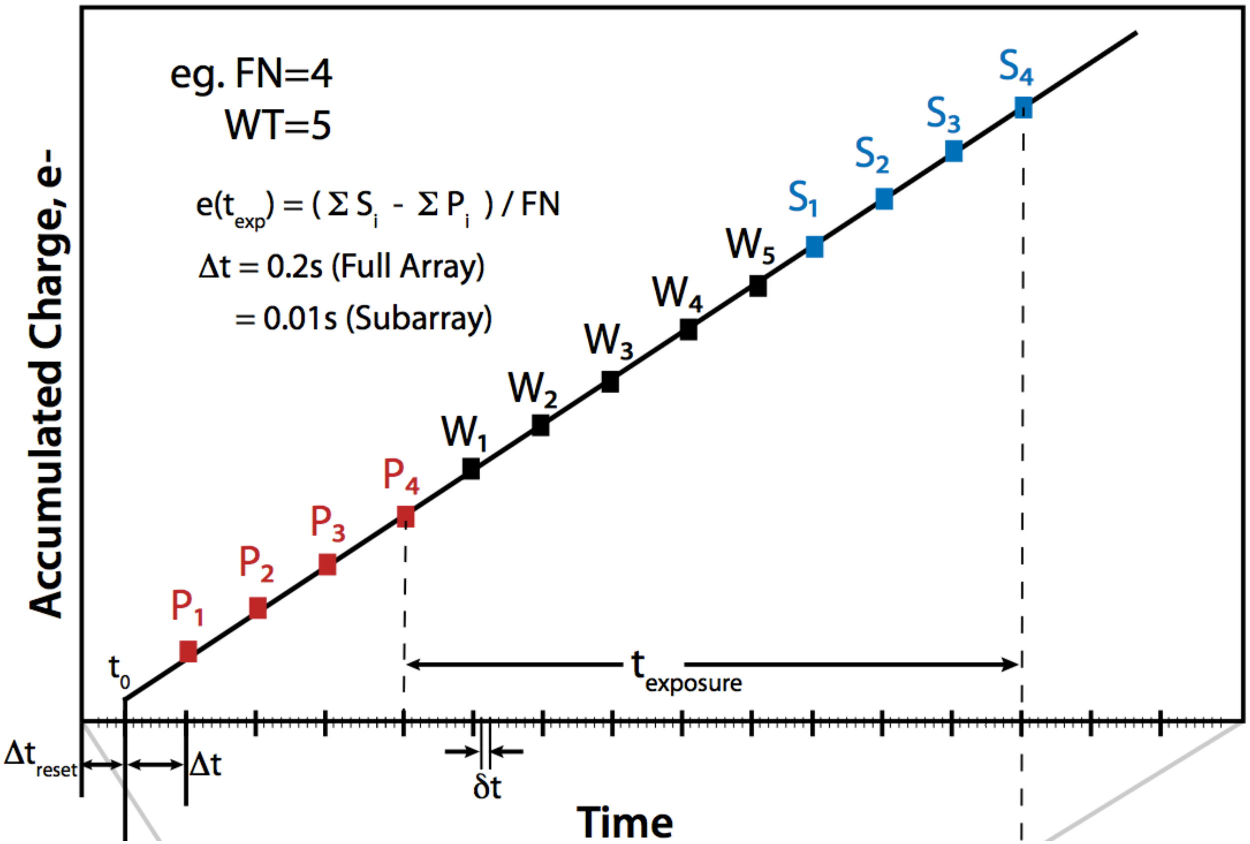


IRAC behavior: Pointing



IRAC behavior: noise

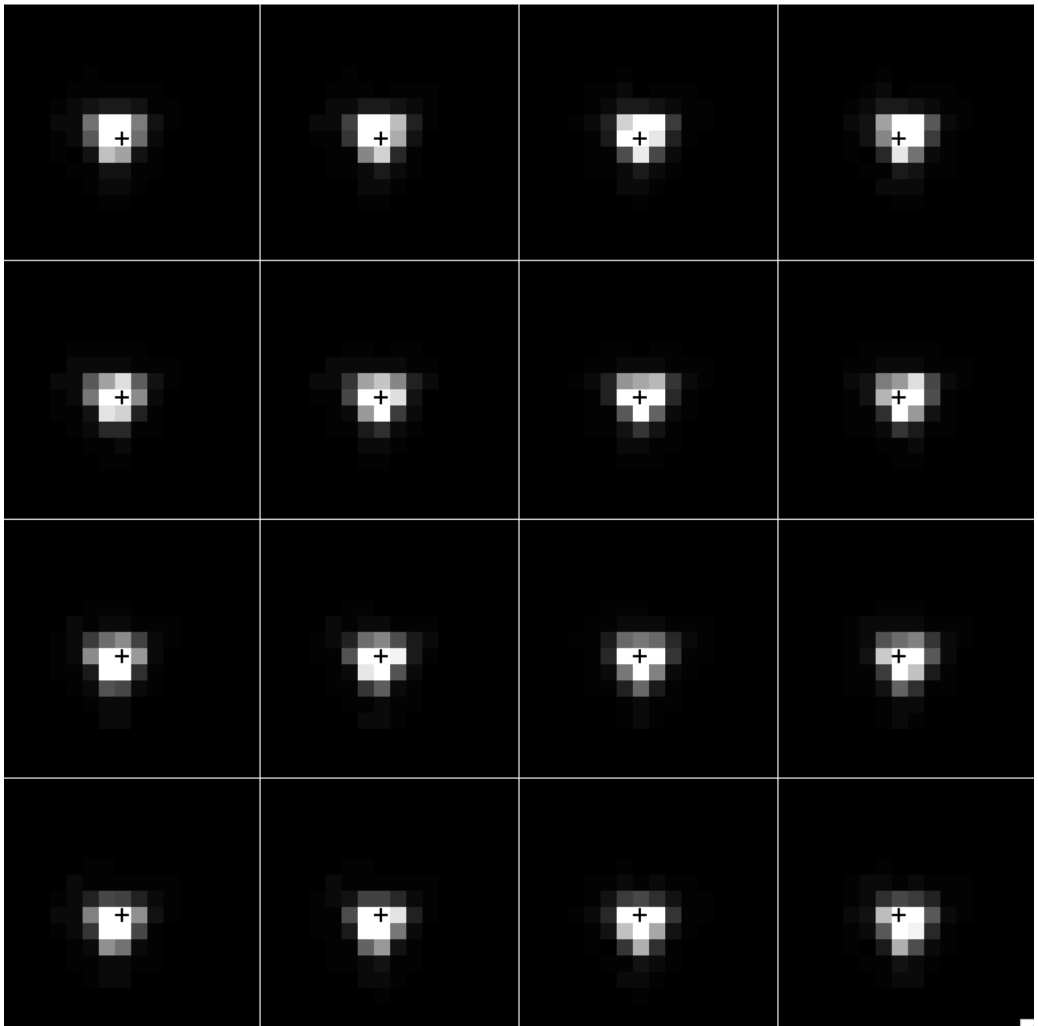
IRAC Fowler sampling



- Non-destructive sampling
- Each signal measurement contains noise of previous samples (correlated)
- Not easy to replicate using single poisson deviate.

IRAC behavior: imaging

IRAC PRF Stellar Realizations

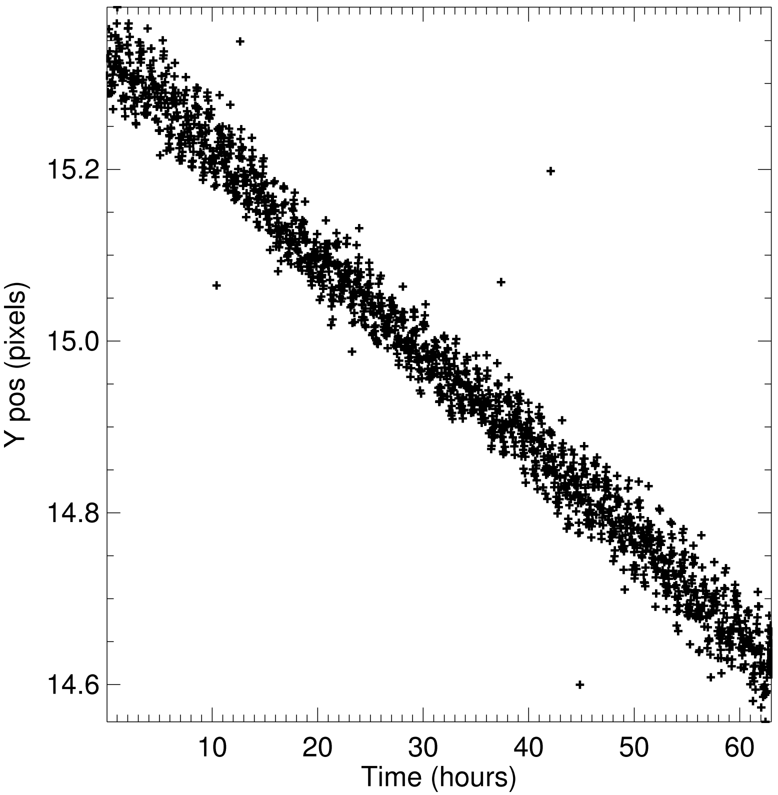


- Undersampled PSF
- Nonuniform pixel response
- Point source appears differently as a function of subpixel position
- Measured centroid is biased

Pointing Model

- Empirical
- IRAC Detector-Based
- Include all known effects
- Most effects specified with an amplitude (vs time) and direction along the detector plane, resulting in projected X and Y pixel behavior
- Allow the user to adjust model parameters, or “turn off” an effect
- User can allow parameters to float and random values within reasonable ranges will be chosen

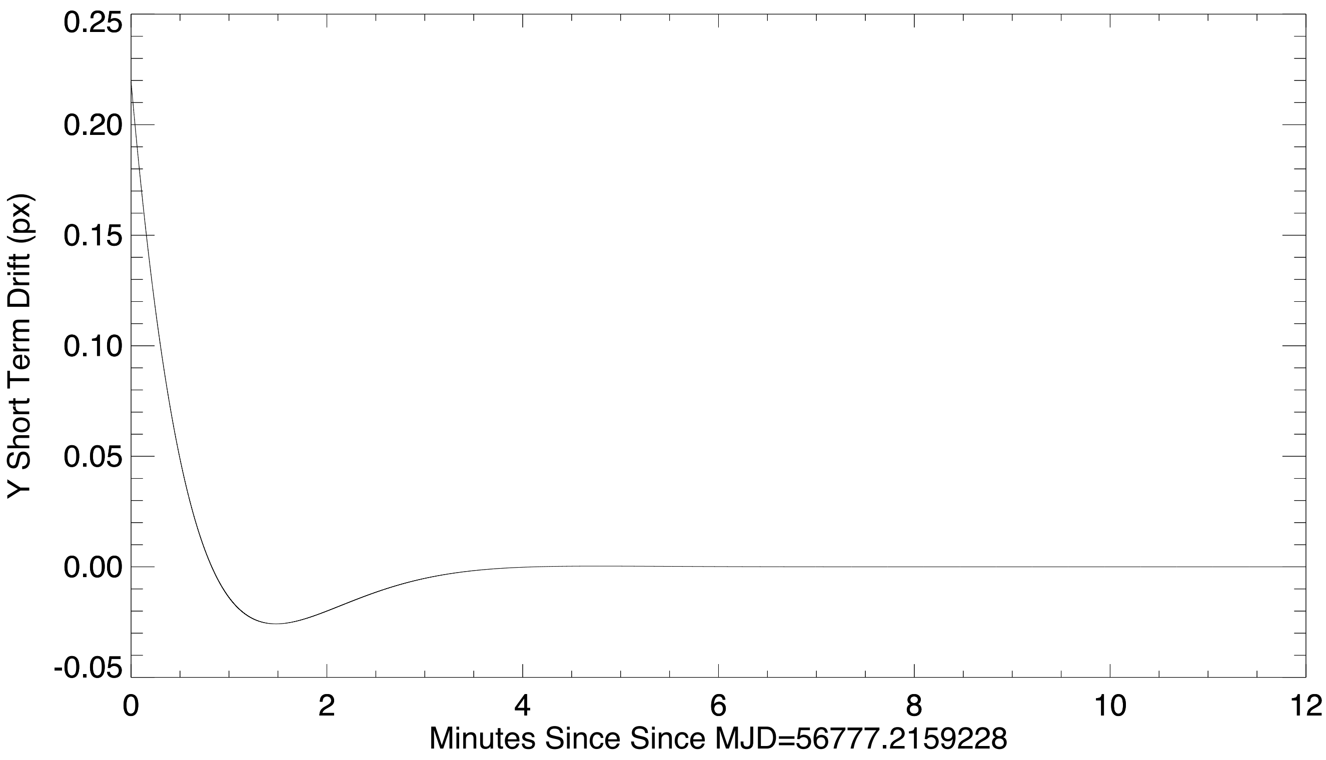
Centroid drift of staring mode observation of XO3



Long term drift < 0.3 arcsec/day

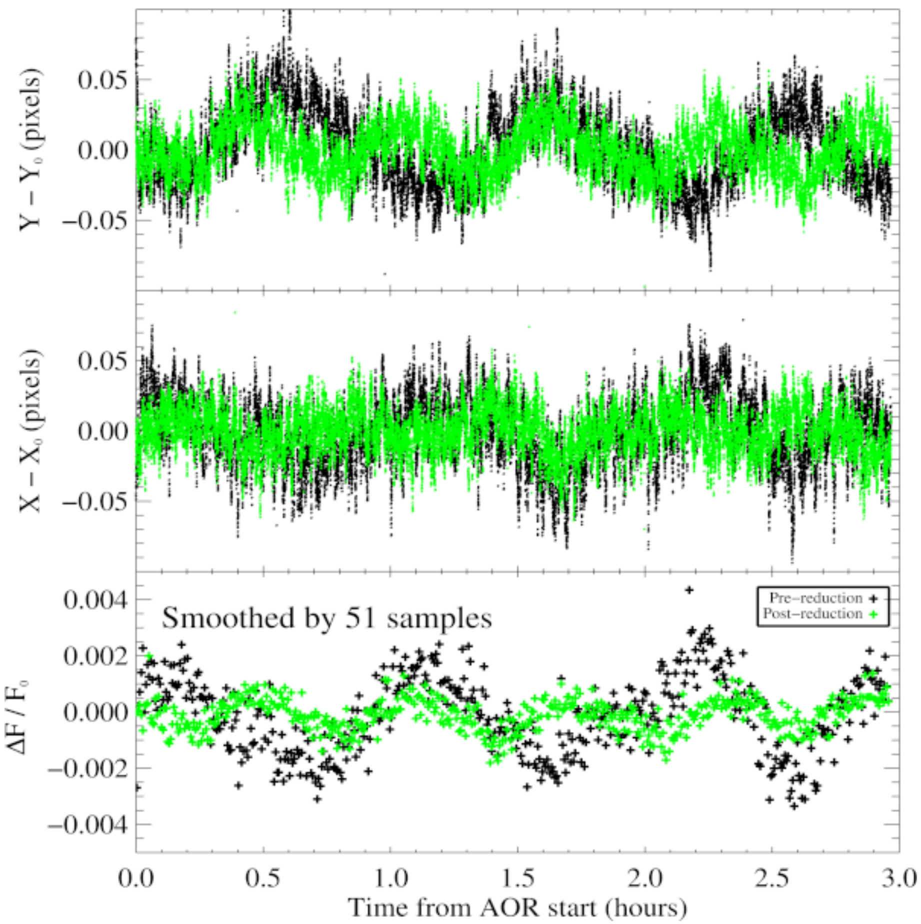
Pointing Model: Short Term "Settling" Drift

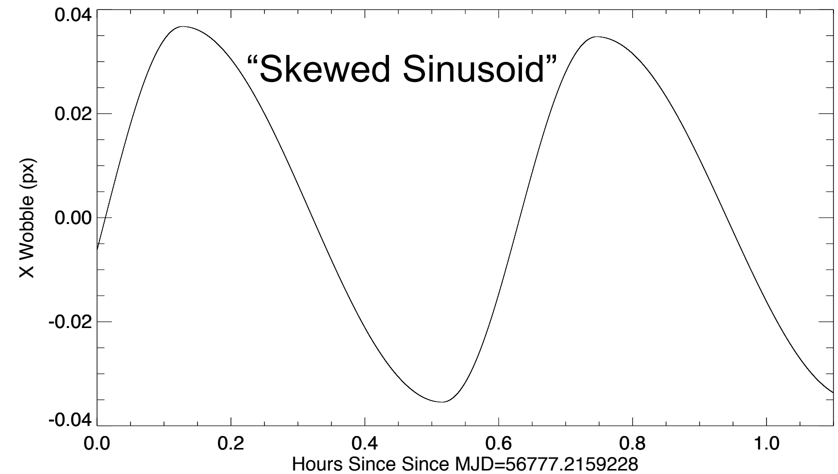
- Rapidly Decaying Sinusoid
- Decay can be positive or negative



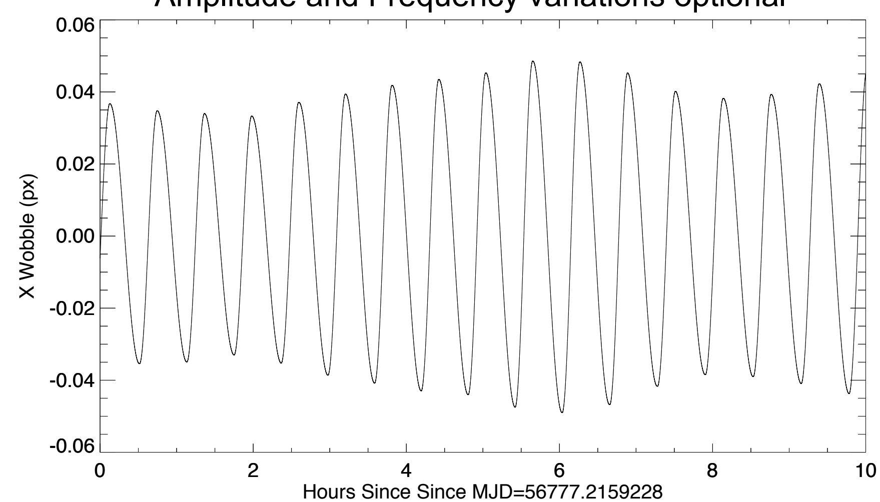
Pointing Model: Heater Wobble

Photometry of HD 158460
Pre (black) and Post (green) heater change

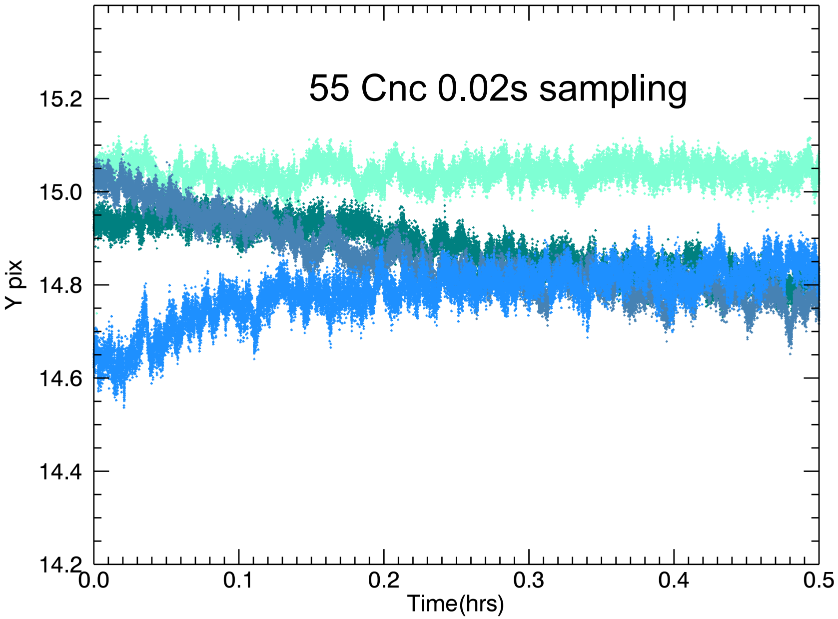


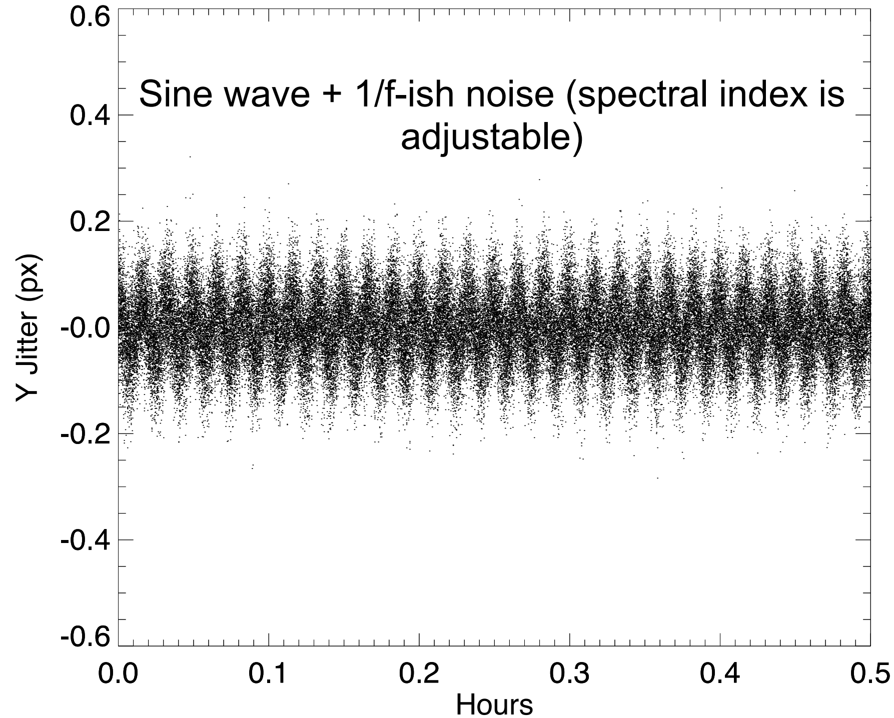


Amplitude and Frequency variations optional



Pointing Model: High Frequency "Jitter"

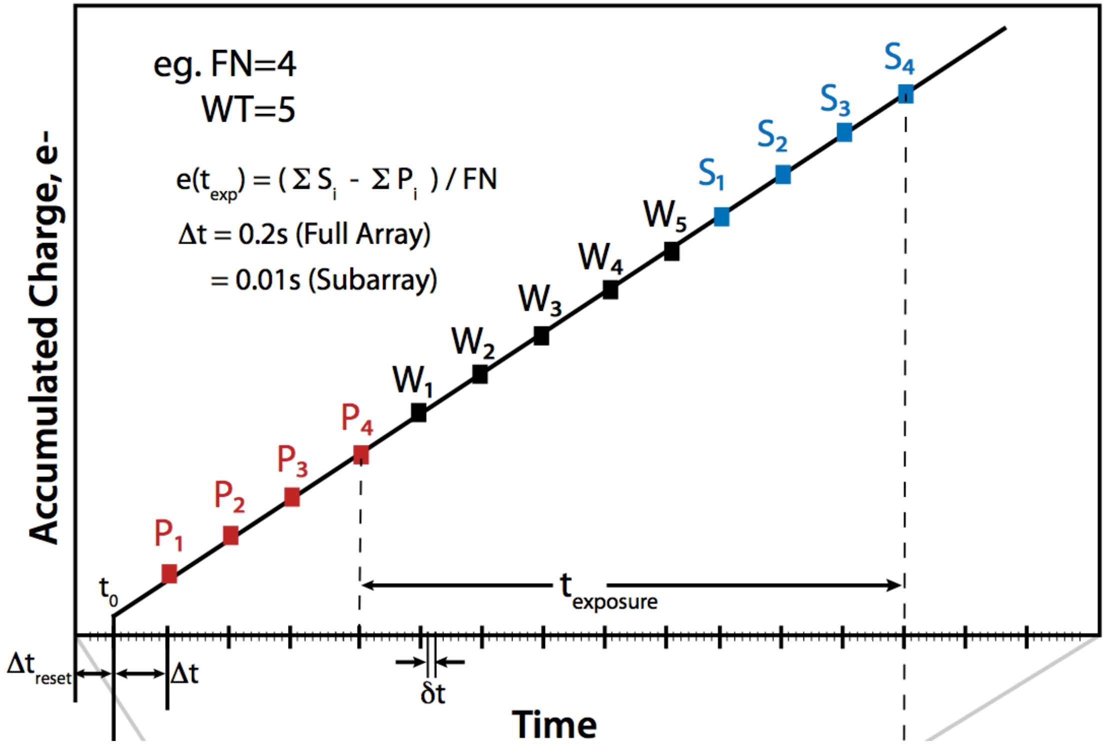




- Fluctuations faster than the Pointing Control System can react

Simulating IRAC Integration

- Pointing model is sampled every 1ms (POINTING)
- IRAC PRF is imaged (e/ms/px) for every (x,y) offset in the pointing model (IMAGING)
- Integrate the 1 ms images to make Fowler samples (IMAGING)
- Shot noise is added per sample based on sample counts (includes noise of previous samples) (NOISE)
- Readout noise is added per read, *after* integration. (NOISE)



Simulator Inputs

- Position(s) of one or more sources (RA,DEC) = (X,Y)
- Date & Time of observation
- Source flux density
- Source light curve
- AOR-type observational parameters (channel, frametime, number of repeats, Full or Subarray)
- Pointing model parameters

Simulator Outputs

- BCD image files
- Uncertainty files
- Can produce results in DN or electrons
- Realistic Fits header (should fool your software)
- Header lists average x,y position over integration from pointing model

```
SIMPLE  =                    T / Written by IDL:  Thu May 15 15:57:58 2014
BITPIX  =                  -32 / Number of bits per data pixel
NAXIS   =                    3 / Number of data axes
NAXIS1  =                   32 /
NAXIS2  =                   32 /
NAXIS3  =                   64 /
DATE    = '2014-04-29'         / Creation UTC (CCCC-MM-DD) date of FITS header
COMMENT FITS (Flexible Image Transport System) format is defined in 'Astronomy
COMMENT and Astrophysics', volume 376, page 359; bibcode 2001A&A...376..359H
COMMENT
COMMENT *******************THESE ARE SIMULATED DATA*******************
COMMENT
ORIGIN  = 'Spitzer Science Center (FAKE DATA)' /Organization generating this FIT
TELESCOP= 'Spitzer (FAKE DATA)' /SPITZER Space Telescope
INSTRUME= 'IRAC (FAKE DATA)'   /SPITZER Space Telescope instrument ID
CHNLNUM =                    2 /1 digit instrument channel number
EXPTYPE = 'sci     '           /Exposure Type
REQTYPE = 'AOR     '           /Request type (AOR, IER, or SER)
AOT_TYPE= 'IracMapPC'          /Observation template type
AORLABEL= 'WASP52_sim_ch2'     /AOR Label
FOVID   =                   77 /Field of View ID
FOVNAME = 'IRAC_Center_of_4.5umSub-Array' /Field of View Name

          / TIME AND EXPOSURE INFORMATION

DATE_OBS= '2014-04-30T05:10:55.731' /Date & time (UTC) at DCE start
MJD_OBS =        56777.2159228 /[days] MJD in UTC at DCE start (,JD-2400000.5)
HMJD_OBS=        56777.2159228 /[days] Corresponding Helioc. Mod. Julian Date
BMJD_OBS=        56777.2159228 /[days] Solar System Barycenter Mod. Julian Date
SCLK_OBS=        1083301855.73 /[sec] SCLK time (since 1/1/1980) at DCE start
AORTIME =              2.00000 /[sec] Frameset selected in IRAC AOT
SAMPTIME=            0.0100000 /[sec] Sample integration time
FRAMTIME=              2.00000 /[sec] Time spent integrating (whole array)
COMMENT Photons in Well = Flux[photons/sec/pixel] * FRAMTIME
EXPTIME =        1.92000000000 /[sec] Effective integration time per pixel
COMMENT DN per pixel = Flux[photons/sec/pixel] / GAIN * EXPTIME
AINTBEG =        21833809.2939 /[Secs since IRAC turn-on] Time of integ. start
ATIMEEND=        21833937.7959 /[Secs since IRAC turn-on] Time of integ. end
AFOWLNUM=                    8 /Fowler number
AWAITPER=                  184 /[0.01 sec] Wait period
AREADMOD=                    1 /Full (0) or subarray (1)
HDR_MODE=                    F /DCE taken in High Dynamic Range mode

          / TARGET AND POINTING INFORMATION

OBJECT  = 'WASP-52 b'          /Target Name
CRVAL1  =        348.494767764 /[deg] RA at CRPIX1,CRPIX2 (using ptg model)
CRVAL2  =        8.76141566878 /[deg] DEC at CRPIX1,CRPIX2 (using ptg model)
RA_HMS  = '23h13m58.7s'        /[hh:mm:ss.s] CRVAL1 as sexagesimal
DEC_HMS = '008d45m41s'         /[dd:mm:ss] CRVAL2 as sexagesimal
RADESYS = 'ICRS    '           /International Celestial Reference System
EQUINOX =              2000.00 /Equinox for ICRS celestial coord. sys.
CD1_1   =   -0.000337893987308 /CD matrix element
CD1_2   =        0.00000000000 /CD matrix element
CD2_1   =        0.00000000000 /CD matrix element
CD2_2   =    0.000337737990775 /CD matrix element
CTYPE1  = 'RA---TAN-SIP'       /RA---TAN with distortion in pixel space
CTYPE2  = 'DEC--TAN-SIP'       /DEC--TAN with distortion in pixel space
CRPIX1  =              16.5000 /Reference pixel along axis 1
CRPIX2  =              16.5000 /Reference pixel along axis 2
PXSCAL1 =       -1.21641835431 /[arcsec/pix] Scale for axis 1 at CRPIX1,CRPIX2
```

Simulator Outputs

```
        / SIMULATION INFORMATION

XLOCAT   =         15.3082349651 /X Position of Star 0 (simulated, average)
YLOCAT   =         15.0917280049 /Y Position of Star 0 (simulated, average)
XYMEAN01= '(15.337853,15.309308)' /(X,Y) Position of Star 0 on subframe 1
XYMEAN02= '(15.338775,15.291182)' /(X,Y) Position of Star 0 on subframe 2
XYMEAN03= '(15.326275,15.266002)' /(X,Y) Position of Star 0 on subframe 3
XYMEAN04= '(15.328428,15.250069)' /(X,Y) Position of Star 0 on subframe 4
XYMEAN05= '(15.304004,15.228810)' /(X,Y) Position of Star 0 on subframe 5
XYMEAN06= '(15.322533,15.215123)' /(X,Y) Position of Star 0 on subframe 6
XYMEAN07= '(15.292161,15.181839)' /(X,Y) Position of Star 0 on subframe 7
XYMEAN08= '(15.313308,15.156205)' /(X,Y) Position of Star 0 on subframe 8
```

Exoplanet Wrapper

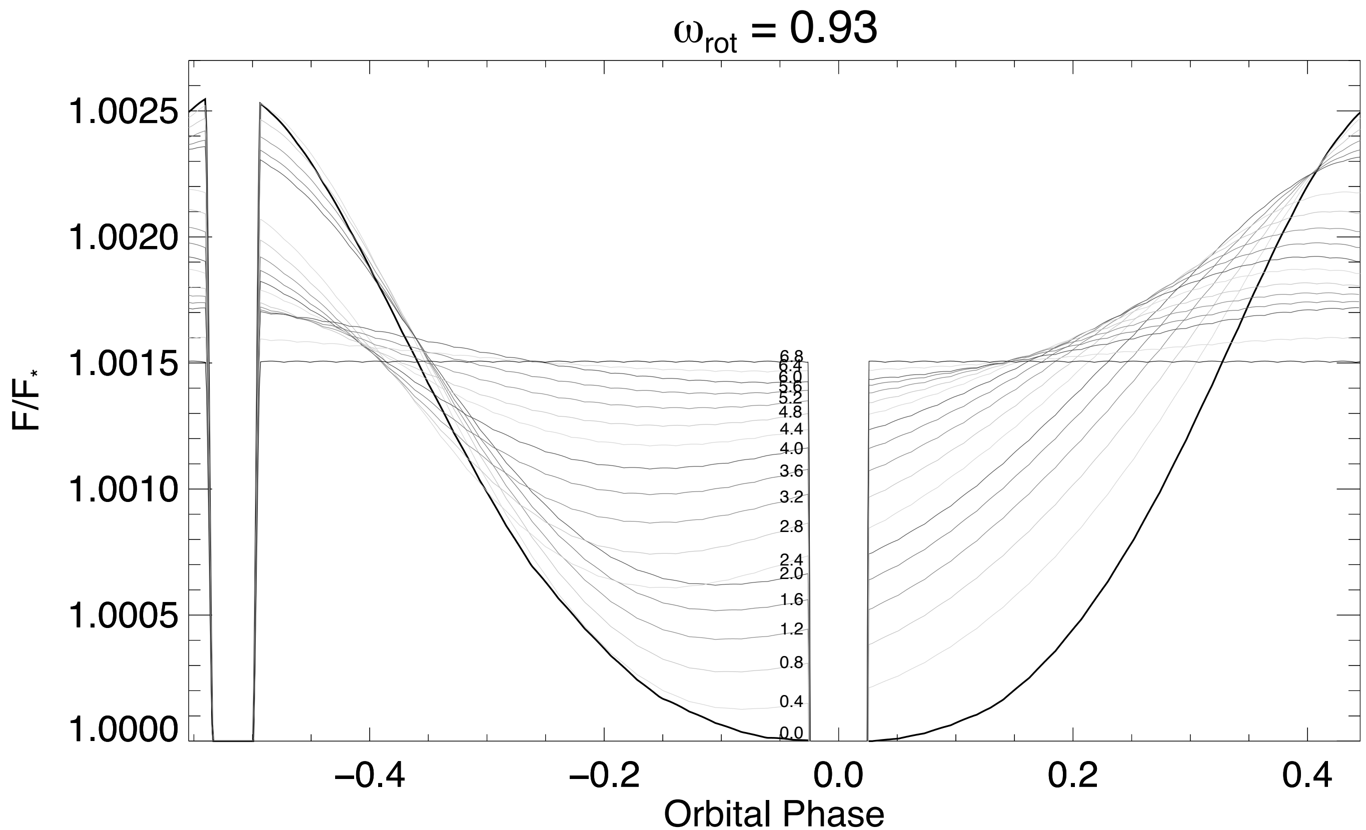
Exoplanet.org database real-time access

INPUTS

- Exoplanet system
- Wavelength of interest
- Atmosphere model parameters: Albedo, t_{rad}, rotational angular velocity. Cowan & Agol (2011) model of phase curve + Mandel & Agol (2002) transit and eclipse shapes
- Limb darkening parameters (nonlinear)
- Total duration (will split into 12 hr AORs, with 30min, each with re-pointing)
- Spitzer observation flags: Peakup, Subarray

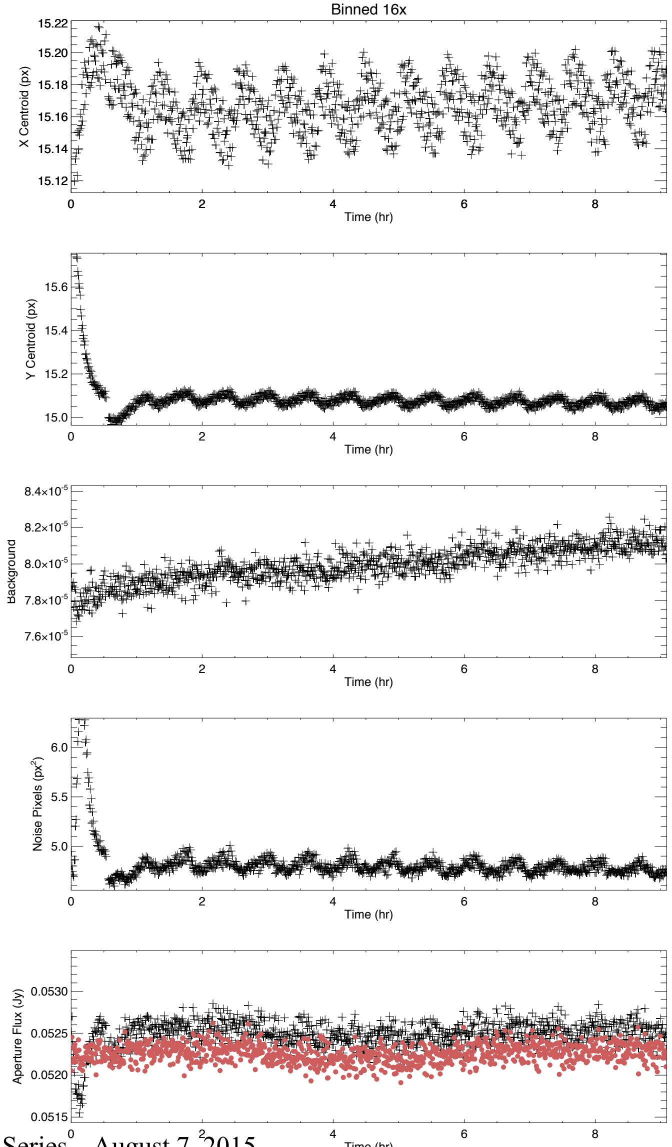
Model Phase Curves

- Wasp-14 b
- Approx. tidally locked (e=0.3)
- Changes with radiative timescale (dy) shown



(Krick et al in prep.)

Example: XO-3 b (Data Challenge)



DATA CHALLENGE 2015

Jim Ingalls
Jessica Krick

What do we want to accomplish?

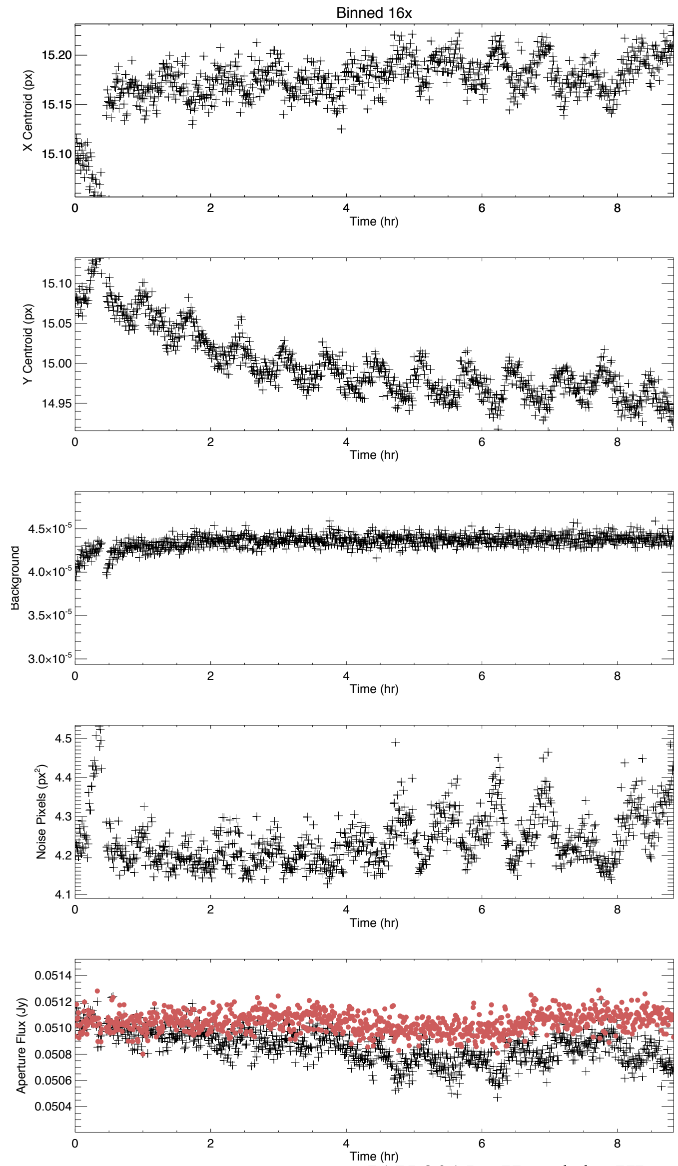
Instrumental Repeatability: are results consistent from one visit to the next (real)?

Methodological Consistency: Do different techniques give similar results (real and simulated)?

Technical Accuracy and Precision: How well do different techniques recover light curve features (eclipse depths) (simulated data)?

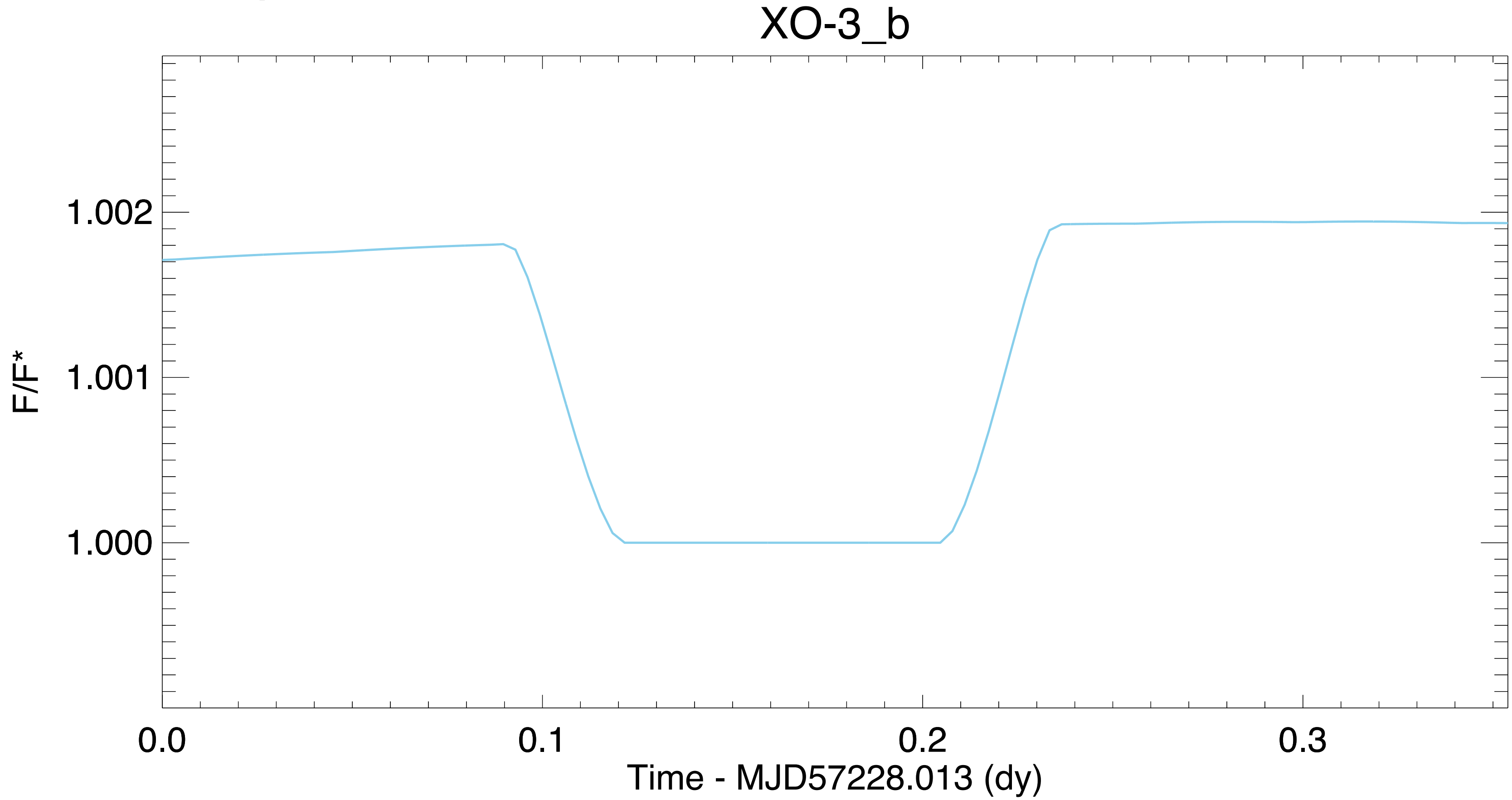
Target Selection: XO-3 b

- XO-3: V=9.8 mag, F5V
- XO-3 b: 11.8 M_J, P = 3.2 day, e = 0.26, i = 84.2 deg
- 10 **secondary eclipse** AORs on Ch2 Subarray "Sweet Spot"
- Analysis published by Wong et al (2014)

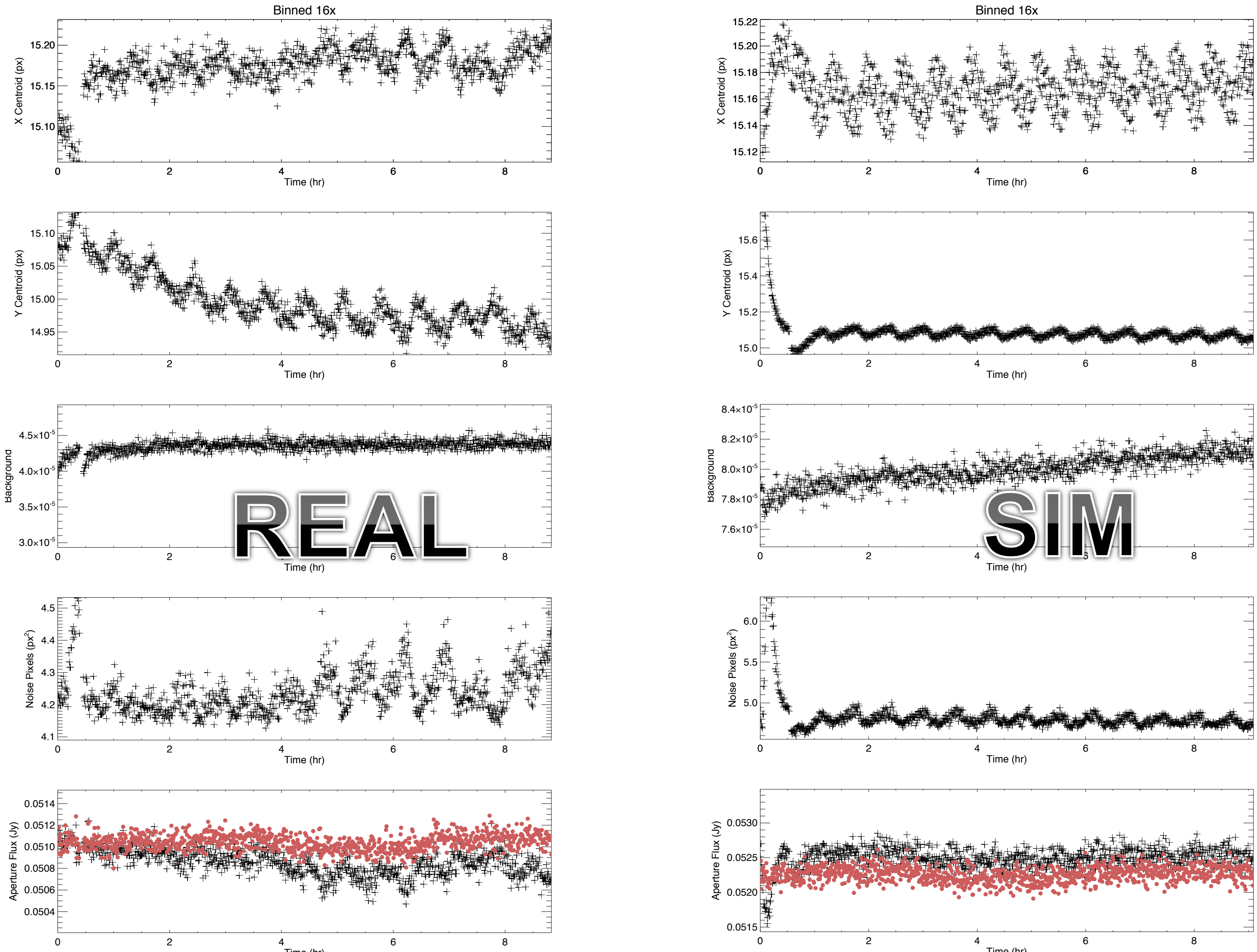


Model of XO-3 b

- Radiative timescale: 1 day, Rotational period: 1 orbit
- Each AOR starts at a slightly different phase
- Time-dependent BG



XO-3 b Real vs. Simulated



Summary of De-trending Methods

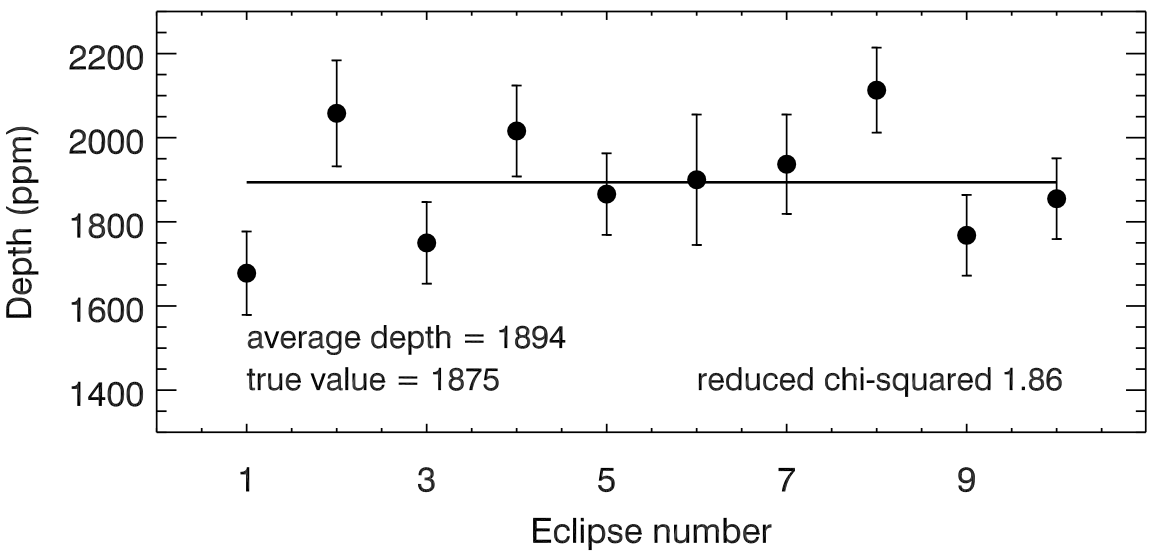
Deming: Pixel-level decorrelation
Evans: Gaussian processes
Krick: Kernel regression with pixel map
Morello: Independent component analysis
Stevenson & Diamond-Lowe: BLISS mapping
Wong et al. (paper): Kernel regression with data

Pixel-level decorrelation (PLD), Drake Deming (UMD)

A limited effort – only Gaussian centroiding and fixed-radius aperture photometry (no variable aperture radii or center-of-light) and eclipse depth errors using only the binned-sigma slopes (no MCMC)

Broad-bandwidth PLD fits as per Sec. 3.1-3.3 of ApJ 805, id.132.

Quadratic temporal baseline to account for the phase effect near eclipse

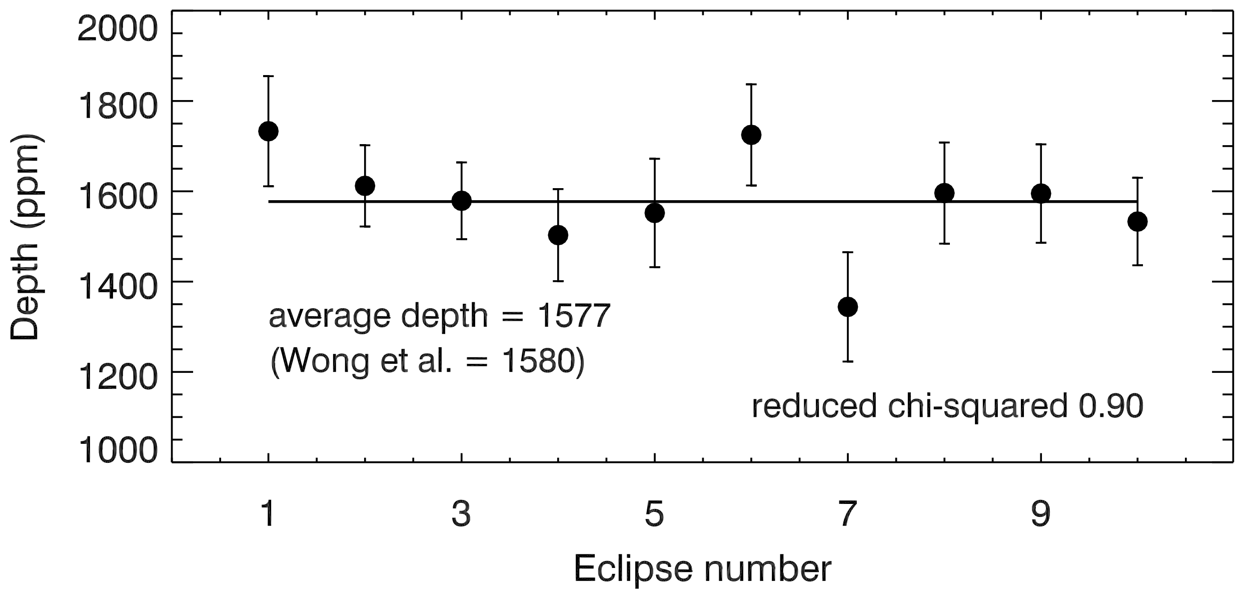


simulated data:
the scatter from eclipse to eclipse exceeds my tentative error estimates, but average depth very close to true value

no red noise; average slope of my binned sigma relations is -0.50

average eclipse depths for the real data scatter consistently with the error estimates

small amount of red noise: average binned sigma slope = -0.44



Average eclipse depth in excellent agreement with Wong et al.

but: Wong et al used a linear temporal baseline and my average depth for that case differs from Wong et al. by 120 ppm
(the temporal baseline is a major factor at the ~ 100 ppm level)

IRAC data challenge: Gaussian process (GP) fits by Tom Evans

Same method used as described in Evans et al (2015).

Only the second, longer AOR was analysed for each lightcurve.

Flux, x and y were binned with respect to time (bins of about 30sec), resulting in lightcurves of ~1000 points.

Maximum likelihood fits were run on the binned lightcurves. Free parameters were the eclipse depth and mid-time, the covariance parameters (i.e. correlation amplitudes and length scales for the systematics), and the white noise.

Uncertainties for the eclipse depths and mid-times were obtained using MCMC with Metropolis-Hastings sampling in the region of maximum likelihood. For this step, the binned lightcurves were again used and the covariance parameters and white noise were fixed to allow fast evaluation of the GP likelihood.

Giuseppe Morello: Independent Component Analysis

Lightcurve detrended as follows (see Morello et al. 2014, ApJ, 786, 22 for details of the method):

- Wavelet transform of 25 (5x5) pixel time series
- ICA in the wavelet domain
- Transformation of the independent components to time domain
- Wavelet denoising of the independent components and raw light-curve
- Fit of the non-eclipse components + astrophysical model over the whole lightcurve (fitting parameters: components' coefficients, phase-shift of the eclipse center, eclipse depth at the eclipse center, out-of-transit slope, normalizing factor).
- Detrended lightcurve is raw minus components.

Stevenson & Diamond-Lowe

- Photometry pipeline:
  - POET, 2D Gaussian centering, 5x interpolated aperture photometry
- Intra-pixel model:
  - BLISS mapping
- Time-dependent models considered:
  - None, linear, quadratic, sinusoidal
- Used BIC to select best models
  - Typically None or Linear ramp
- Fit common eclipse depth, duration, and ingress/egress over all visits.
- POET uses larger aperture sizes (4.25 vs 2.5 pixels) and produces more precise light curves than CSV Tables, but eclipse depths and uncertainties are comparable.
- Including phase curve variation (through quadratic or sinusoid) finds more accurate eclipse depth with simulated data, but those models are not favored by BIC (ΔBIC > 7).

JK for the IRAC team

Data Reduction

- Box_centroider = center of light
- Aperture photometry [2.25, 3-7]
- Intrapixel gain correction = pmap_correct = nearest neighbors kernel regression
  - 50 nearest neighbors in X, Y, NP within .0025pixels
  - Advantage is not using the science data to correct itself
  - Must land on the sweet spot
- Minor outlier rejection
- Error bar determined by XXX

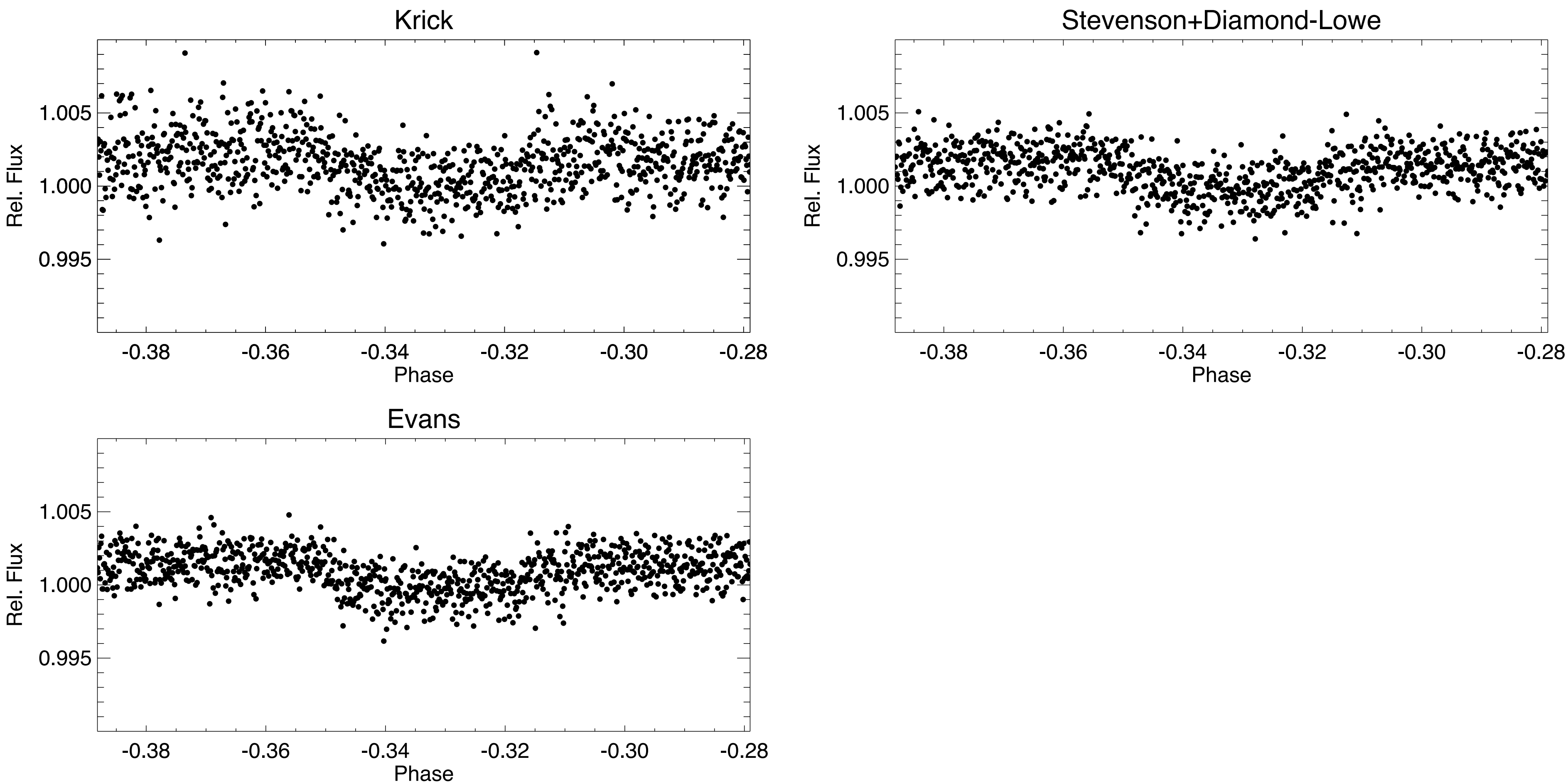
Eclipse Depth Fitting

- Cheated the phases to make them look like transits
- Applied Mandel & Agol 2002
- Used measured “transit depth” as an eclipse depth

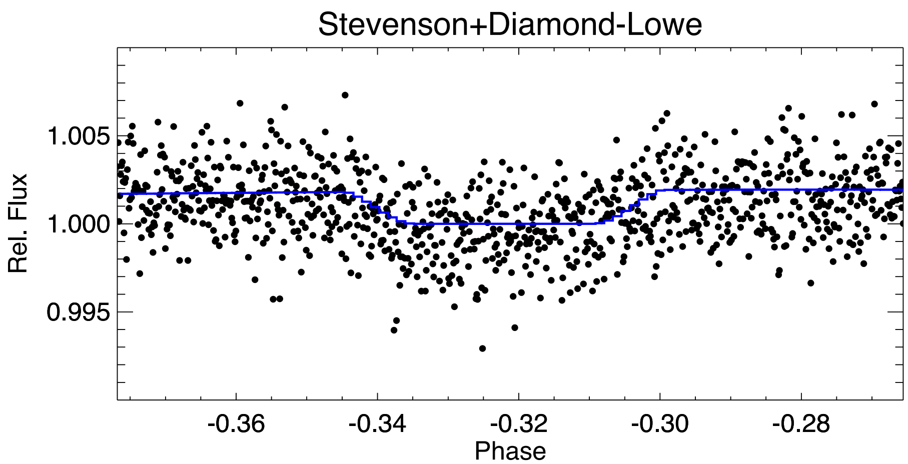
Note

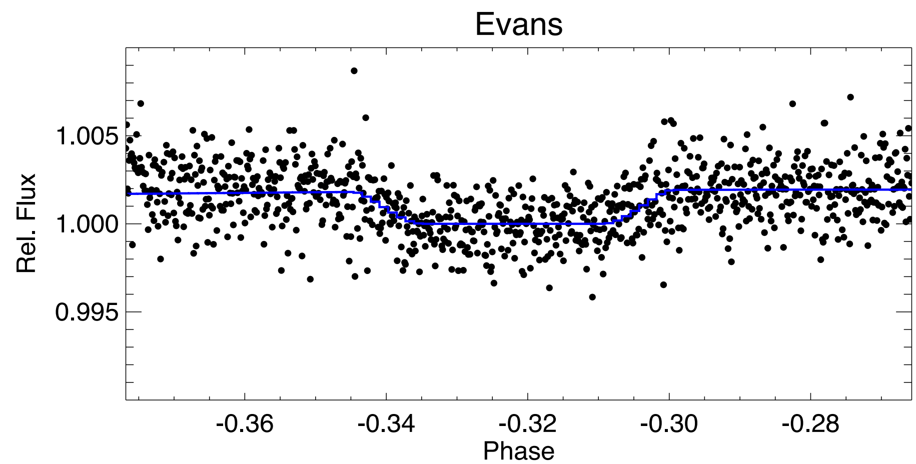
- Simulated data only used kernel regression for nearest neighbors in X, Y and *not* NP

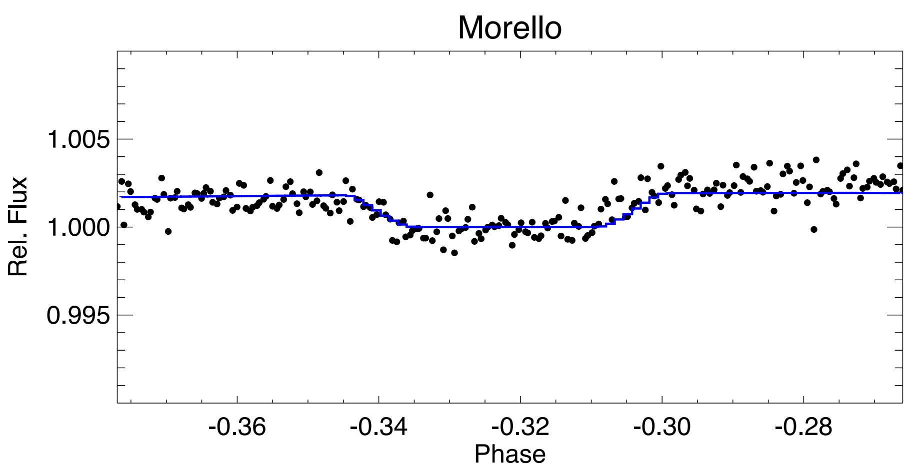
Data Challenge Results: Real



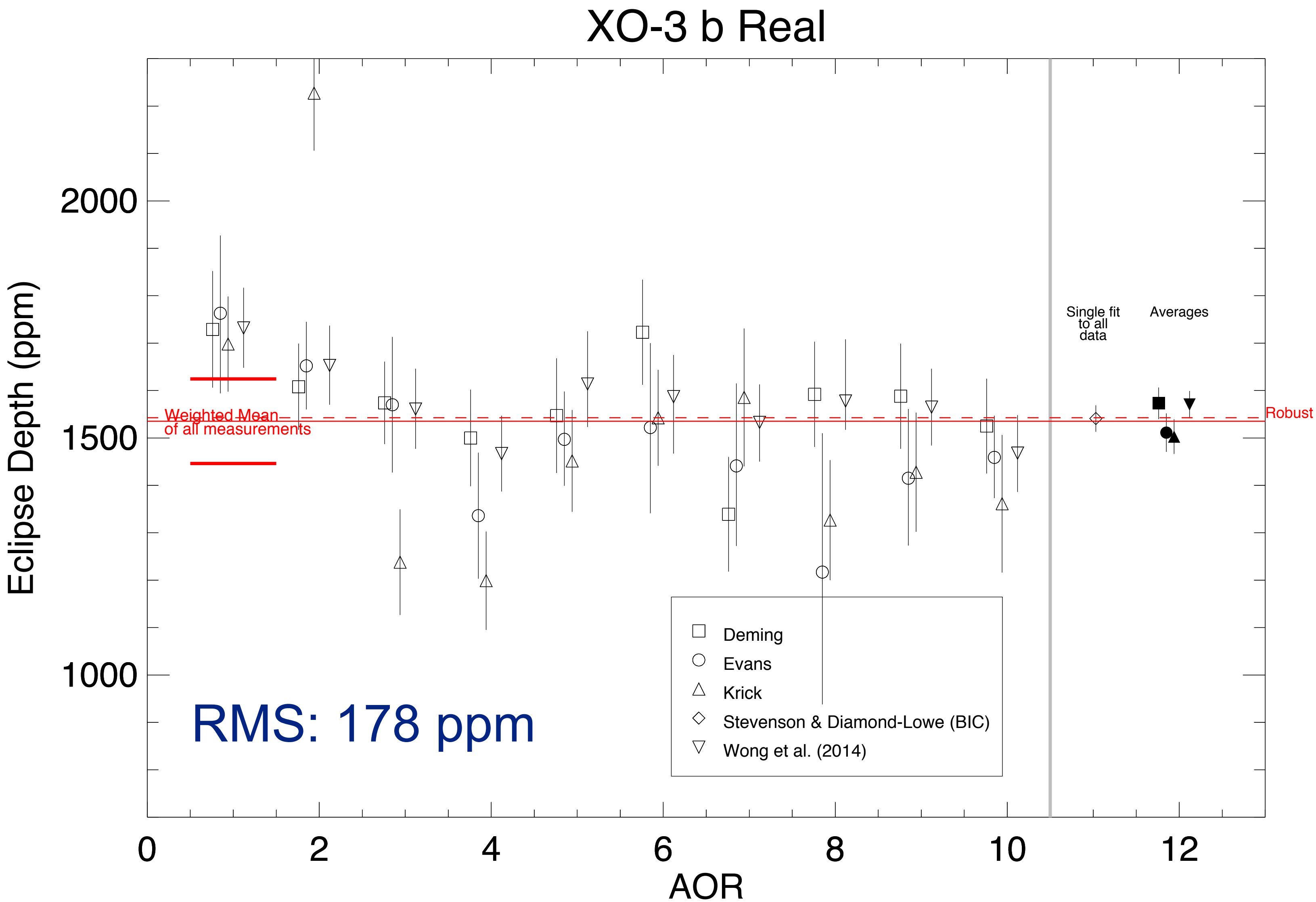
Data Challenge Results: Simulation



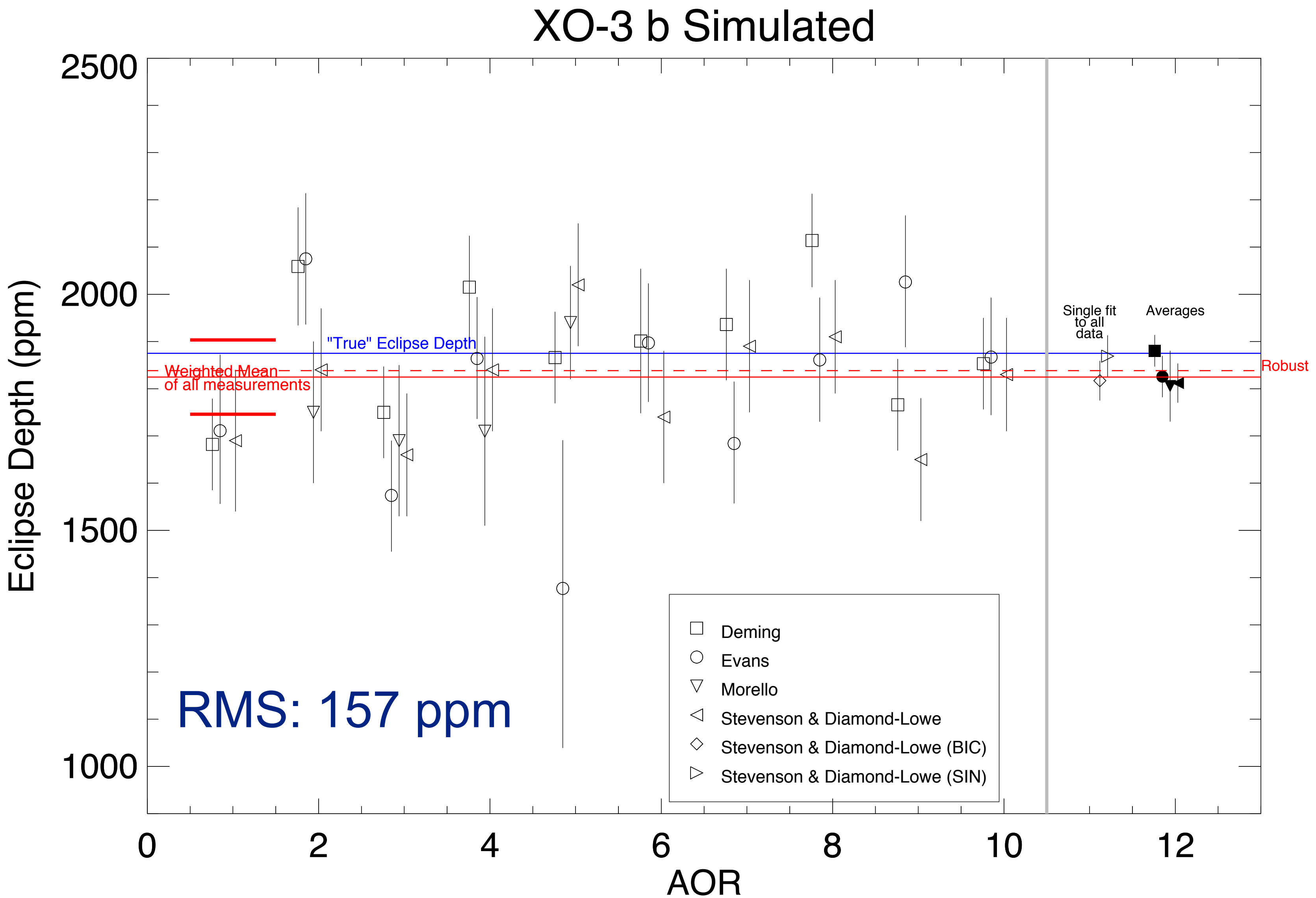




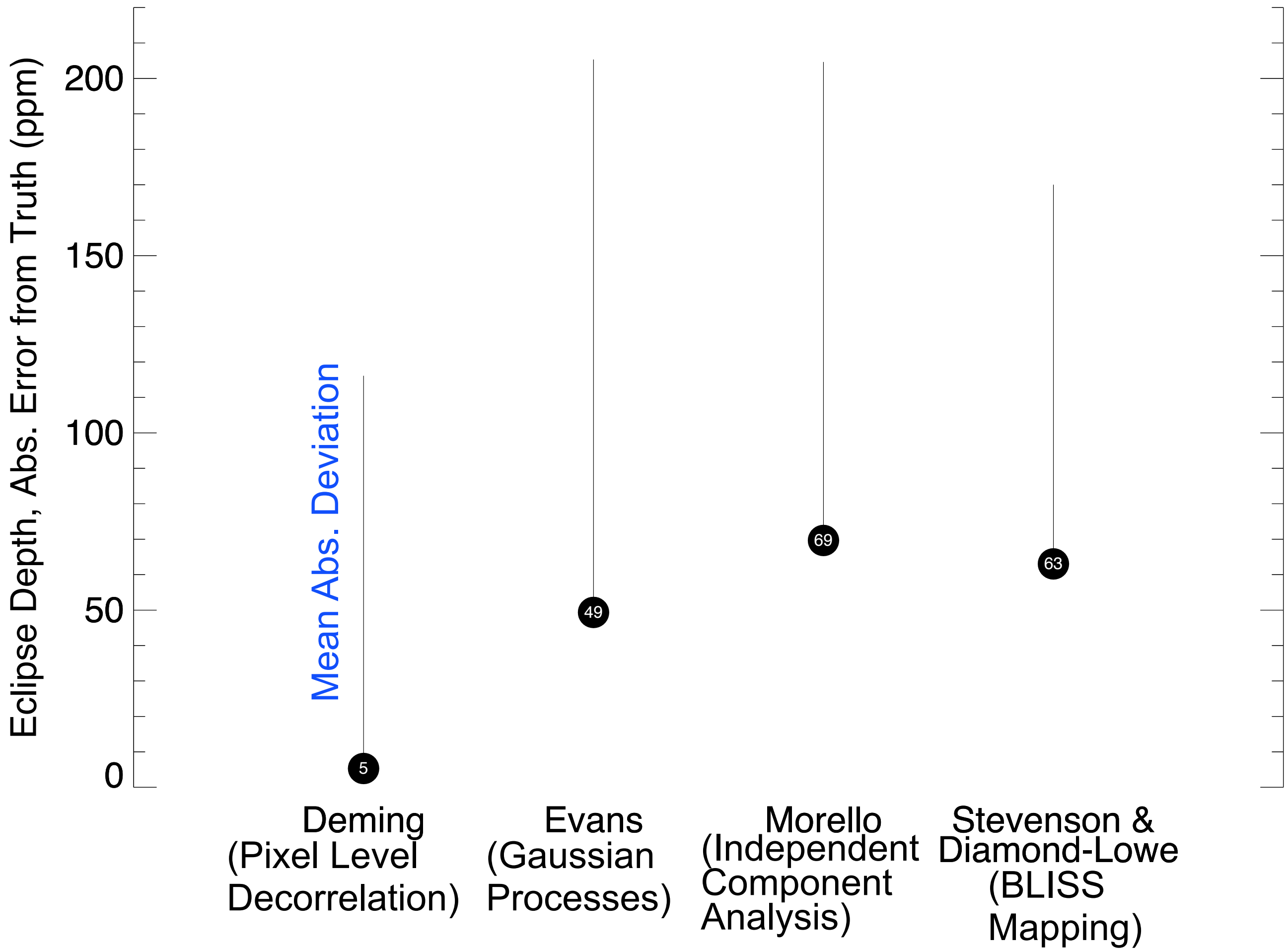
Eclipse Depth Measurements: Real



Eclipse Depth Measurements: Simulated



Eclipse Depth Accuracy: Simulated



Data Challenge: Conclusions

- **Instrumental Repeatability**: Scatter in eclipse depths in real data is comparable to that of simulated data (RMS 178 vs 157 ppm).
- **Methodological Consistency**: Techniques give mean results that are consistent within the uncertainties.
- **Technical Accuracy**: Mean eclipse depths for simulated data are correct to about 5-63 ppm.
- **Technical Precision**: Uncertainty in mean is about 50 ppm.