Exploring Hot Jupiter Atmospheres via Ground-based Secondary Eclipse Detections: Biases, Limitations, and Lessons Learned

Justin Rogers[1,2,3], Dániel Apai[3,4], Mercedes López-Morales[2,5] & Elisabeth Adams[6]

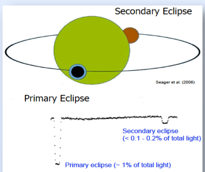
[1]Johns Hopkins University; [2]Department of Terrestrial Magnetism, Carnegie Institution of Washington; [3]Space Telescope Science Institute; [4]Steward Observatory, University of Arizona; [5]Institut de Ciencias de l'Espai; [6]Massachusetts Institute of Technology

Work funded by:

Missions.22

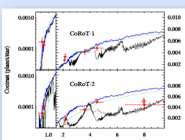
Introduction

Key to our understanding of transiting exoplanets is the wealth of information we can obtain from detection of secondary eclipses, in which the planet passes behind its host star, as viewed from Earth. Eclipses have been detected from space (*Spitzer, Hubble, CoRoT, Kepler*) and from ground-based observatories.

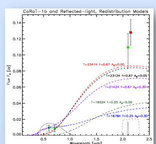


Real data for HD 189733b (Knutson et al. 2007)

By separating the light from the planet from that of the star, we directly measure the planet's combined thermal and reflected emission in a given passband. From this we can constrain the planet's day-side temperature, albedo, and atmospheric energy circulation, and multiple detections at different wavelengths can describe the atmosphere's chemical composition and pressure-temperature profile, including the presence or absence of a thermal inversion (e.g. Burrows et al. 2008, Fortney et al. 2008)



CoRoT, ground-based, and Spitzer detections of exoplanets CoRoT-1b and Corot-2b from the optical to mid-infrared (Deming et al. 2011)

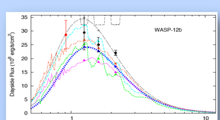


Detail of the optical and near-IR detections for CoRoT-1b, with blackbody models (Rogers et al. 2010, Gillon et al. 2009, Alonso et al. 2009, Snellen et al. 2009)

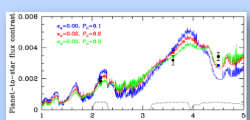
Puzzlingly, many of the ground-based detections suggest that planets are hotter than predicted by standard models, particularly around 2 µm. This is seen in both CoRoT-1b and WASP-12b, as well as WASP-19b, although other detections, such as those of WASP-4b, TrES-2b, and TrES-3b do not exhibit this trend.

For CoRoT-1b, the best-fitting model is a blackbody hotter than the maximum expected equilibrium temperature of the day-side of the planet (see detail above right, Rogers et al. 2010).
For WASP-19b, the measured brightness temperature of the NB2090 detection is 2540 +/- 180 K, compared to a maximum equilibrium temperature of ~2400 K.

Is this a real physical effect, suggesting alternative atmospheric models (e.g. non-LTE conditions, unexpected chemical abundances)? Or is it simply a product of systematics in the analytical methods used to detect and model the eclipses?



Ground-based detections of WASP-12b in z', J, H, K (Lopez-Morales et al. 2010, Croll et al. 2011)



Ground-based and Spitzer detections of WASP-4b (Caceres et al. 2011, Beerer et al. 2011)

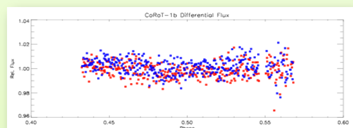
Exploring Strange New Worlds, Flagstaff, AZ, May 1-6, 2011

Observations

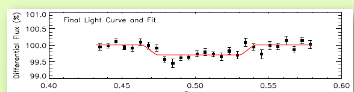
Eclipse detections are primarily determined by:
- Eclipse Depth
- Point-to-point uncertainty
- Magnitude and structure of the red noise
- Stable comparison star(s)

Secondary considerations include:
- Length of Baseline
- Sampling rate
- Fixed position on detector vs. dithering
- Sky subtraction
- Airmass, focusing

Even with optimal noise-minimization techniques, a significant amount of correlated, or "red" noise will remain, as in this raw differential light curve (red and blue represent separate offset positions):



These trends must be removed in order to accurately detect the eclipse signal, as shown here (points in 12-minute bins; Rogers et al. 2009).



Modeling and Analysis

The red noise trends often correlate with other factors:
- Atmospheric effects – airmass, seeing, sky brightness
- Instrumental effects – x- and y-position, shape of images on chip

These can be removed manually or with a blind routine (e.g. SysRem), looking at only the out-of-eclipse baseline points, but a more robust method that uses all of the datapoints is to combine the systematic trends and the actual eclipse shape into a single model with a number of parameters to solve for.

Our model allows a linear correlation between differential flux and a number of parameters: orbital phase (i.e. time), airmass, FWHM of the images, x- and y-displacement from the average position, and sky brightness around each star. For each variable q, the trend is modeled by simply:

$$C_q(q) = 1 + m_q q$$

All the trends are then combined with the eclipse shape F_{obs}, consisting of adjustable parameters overall baseline flux F_b, eclipse depth D_e, and mid-eclipse phase ϕ_{me}:

$$F_{obs} = F_{ecl}(F_b, D_e, \phi_{me}) \prod_q C_q(q)$$

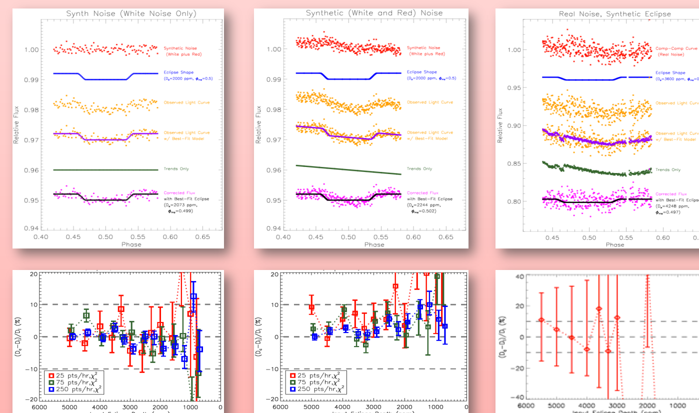
To find the best-fit set of parameters – these three plus all the m_q's – we run a series of Monte Carlo Markov Chains (MCMC), commonly used in eclipse modeling.

Tests and Results

To determine the reliability of the modeling and analysis routines, we designed a series of eclipse light curves on which to test the process. We began with completely synthetic data, combining an eclipse shape of known depth with a noise model, either white noise only or a combination of white and red noise. Three different red noise patterns were generated by combining sinusoids of different periods. Then the eclipse-plus-noise models were produced with a given sampling rate and baseline length, and treated as an observed differential flux from real photometry.

In the white-noise-only tests (example in left column below), the main determinant of recovering the input eclipse signal was the ratio of the eclipse depth to the white noise (amplitude fixed at 0.1%), although the sampling rate and baseline did have an effect as well. The input depths were recovered within 10% for depths down to 0.8 times the white noise for the shortest baseline and slowest cadence; for longer baselines they were recovered down to 0.3-0.5 of the white noise. At faster sampling rates, the dispersion of the errors between input and recovered depth decreased. With red noise added, the depths were recovered less accurately, and we sometimes saw a bias to the results, as in the example below (middle column), in which the red noise structure leads to systematically deeper eclipses recovered than the input.

After this, we tried using photometric noise (e.g. differential flux from comparison stars in the CoRoT-1 field, observed in Ks-band with the NICFPS instrument at Apache Point Observatory), combined with synthetic eclipse curves. This allowed correlation of the flux with each of the atmospheric and instrumental effects included in the models. While previously we used fixed amplitudes for the white noise (0.1%) and red noise (<0.1%), the point-to-point dispersion of the comp-comp noise was 0.6%; note the different scales in the examples (right column). The routines appeared to do a good job of removing the systematics, but the recovered depths still became inaccurate at higher depths than in the synthetic tests, and had much greater uncertainties on the measurements.



Conclusions / Future Work

Correlated, or "red" noise can play a significant part in the detected depths of exoplanet eclipses. We have seen that even with robust analysis methods, a red noise structure can cause systematic and significant biases to detected eclipse signals. With our current understanding of exoplanet atmospheres based largely on a small number of thermal emission detections, it is critical to know how accurate these measurements are.

To help characterize the accuracy and limitations of each dataset and analysis routine, we propose to create a consistent standard of simple benchmark tests. We have put together a series of sample light curves with known eclipse inputs that will be available for investigators to use to test their analysis routines and evaluate their accuracy and biases before applying them to new detections.

We welcome feedback or questions, please reach me by email at rogers@pha.jhu.edu.