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

Justin Rogers1,2,3, Dániel Apai3,4, Mercedes López-Morales2,5 & Elisabeth Adams6

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

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.

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).

Positively, 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-10b, although other detections, such as those of WASP-4b, HAT-P-26b, and HAT-P-16b 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 data above right, Rogers et al. 2010). For WASP-10b, 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-CO2 conditions, unexpected chemical abundances)? Or is it simply a product of systematics in the analytical methods used to detect and model the eclipses?

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).

Testing and Analysis

The red noise trends often correlate with other factors:
- Atmospheric effects – airmass, seeing, sky brightness
- Instrumental effects – e.g. position, shape of images on chip
- These can be removed manually or with a blind routine (e.g. Python), looking at only the out-of-eclipse baseline points, but a more robust method that uses all of the data points is to combine the systematic trends and the actual eclipse shape into a single model with a number of parameters to solve for:

\[ C_\epsilon(e) = 1 + m_\epsilon \theta \]

All the trends are then combined with the eclipse shape \( F_{\text{ecl}} \), consisting of adjustable parameters over baseline flux \( F_{\text{bas}} \), eclipse depth \( D \), and eclipse phase \( \phi \).

\[ F_{\text{ecl}} = F_{\text{bas}}(D) + m_\epsilon \theta \phi \]

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

Conclusions / Future Work

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, containing 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. a differential flux from comparison stars in the CoRoT-1b band, observed in its band with the NGSTPS 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 (0.11%), the point-to-point dispersion of the comp-comp noise was 0.06%; 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.

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