Rohan Subramani

Research Plan: TESS Updates to Plan JWST Observations for Atmospheric Characterization of Promising Exoplanets

1. Rationale

Pythagoras is generally remembered for his contributions to the field of mathematics, but he was a philosopher as well. In fact, his fascination with numbers was driven by his core philosophical principle: "Number is the ruler of forms and ideas." What he meant by this was that mathematical relations described everything in the universe, and that by pursuing these formulae, people could gain new insights to answer the most burning questions of philosophy (Buckingham et al., 2017). This pursuit branched into many fields, including physics and astronomy. At times, the great distances to objects of astronomical interest can lead people to believe that studies in the field are of minimal importance to humans and unnecessarily drain money and resources. However, the significance of astronomy research cannot be overstated: it provides us with glimpses into our distant future and the universe's distant past, and it gives us information to answer such deep questions as, "Where did we come from?" and "Are we alone in the universe?"

The field of exoplanet research in particular is shedding new light on these questions, and it has taken off since the discovery of the first known exoplanet, 51 Peg b, in 1995. Today, there are over 4,000 known exoplanets, and the recently launched Transiting Exoplanet Survey Satellite (TESS) is expected to add significantly to this total ("Exoplanet Archive," 2019). In fact, it will provide a thorough catalog of transiting planets around bright, nearby stars (Ricker et al., 2015). "Transiting" means that they pass directly between their host stars and the Earth, which is an important characteristic because the dip in light we receive from the star during a transit is how we detect most planets and describe their characteristics. These transiting planets around bright nearby stars will be the easiest to study further with future telescopes, so this mission is the first step towards a much greater understanding of extrasolar planets.

Scientists have a number of goals when studying exoplanets: finding planets where life may exist, discovering planets where humans could potentially live, learning how planets form, and, most relevant to this project, determining atmospheric characteristics of exoplanets in order to better understand how Earth's atmosphere developed and how it compares to those of other worlds. The atmosphere is an essential component of how life developed on Earth, and studying exoplanet atmospheres has the potential to answer some fundamental questions about both the origins and the future of our existence. No telescopes launched to date have the sensitivity to detect exoplanet atmospheres, but the James Webb Space Telescope (JWST), scheduled for launch in 2021, will change that (NASA, 2019). However, even JWST will not be able to characterize all exoplanet atmospheres: planets with thin or nonexistent atmospheres can be near impossible to study. For that reason, it is important to select planets that JWST would be able to provide significant information about for study. In order to quantify this "atmospheric characterizability," two metrics have been developed based on the methods that will be employed to identify and describe atmospheres: the Transmission Spectroscopy Metric (TSM) and the Emission Spectroscopy Metric (ESM) (Kempton et al., 2018). TSM represents how characterizable a planet's atmosphere will be based on the absorption spectrum received during a transit. This absorption spectrum is the spectrum of light wavelengths we receive, and certain wavelengths will not be received depending on what elements are in a planet's atmosphere. A high TSM for a planet means that the signal to noise (S/N) ratio of its absorption spectrum will be high, increasing the likelihood of successful atmospheric characterization. ESM represents how characterizable a planet's atmosphere will be based on the absorption spectrum received just prior to a secondary transit. A secondary transit is when the planet passes behind the star, which also causes a slight dip in light because planets also emit electromagnetic radiation due to their temperatures (though much less than stars). When a planet is very near a star from the point of view of Earth, this thermal emission will also be received in addition to the star's radiation. This thermal emission will also pass through the planet's atmosphere before reaching Earth, so once the stellar radiation is filtered out, an absorption spectrum would again be obtained. Again, this absorption spectrum indicates which wavelengths of light are absorbed by the elements in the atmosphere, and because each element has a unique set of absorption wavelengths, this information allows for atmospheric characterization. The calculations of TSM and ESM values are described in Kempton et al. (2018).

Not all the characteristics required to determine TSM and ESM values are known for every exoplanet, but in all cases where it was possible, these values have been calculated. My research will focus on updating the characteristics of the top 10 ESM targets using TESS data, especially epoch and period. In doing so, future transits can be predicted far more precisely, providing brief, specific windows during which JWST can be aimed towards these planets' host stars for atmospheric characterization. Minimizing observation time is essential, as JWST will have a busy observation schedule and should be observing targets only at the times when useful data can be gathered.

2. Research Questions

a. What are the planetary characteristics of the ten known exoplanets with the highest ESM values?

b. When should JWST be directed towards each of their stellar hosts?

3. Hypotheses

a. If planets have high ESM values, then they are likely large, hot planets with low masses because these characteristics make atmospheric detection more likely (Kempton et al., 2018). No general hypotheses can be formulated for the exact planetary characteristics because they will vary from planet to planet.
b. If we utilize data from TESS, then the ideal times for JWST observation of the top ten ESM planets will be obtained with high precision and accuracy because TESS data combined with archival information provides a clear picture of planets' orbital history over several years, making it easier to predict future secondary transits (Kempton et al., 2018).

4. Expected Outcome

It is expected that most, if not all, of the top 10 ESM targets will be Hot Jupiters, and that the procedures of this study will provide an update to the known characteristics of these planets. The parameters that this project will inspect include planet vs. star radius ratio, semimajor axis, epoch (which is a reference point for the beginning of one transit, such that all other transits start an integer multiple of the period after the epoch), the period, and more. The accuracy and precision of period and epoch calculations are expected to be greatly enhanced because TESS allows for a check of period and epoch values from years prior. Even small margins of error for these values can lead to large discrepancies between prediction and observation after several years, so knowing modern transit timings greatly decreases the margin of error for previous observations. This will prove especially useful when selecting observation times for JWST, as it will allow for minimal time requirements of a widely requested telescope.

5. Procedure

<u>Role of Mentor and Role of Student</u>

My mentor will provide me with guidance regarding how to acquire data from TESS and previous studies of the top ten ESM targets. He will teach me how to prepare data files such that a software package he prepared with a colleague can read them and conduct fits of models to the data. He will teach me the concepts behind the model that allow for these fits. I will go through databases and published journal articles to collect photometric and radial velocity data for each of the planets of interest, gather this data in a format that the software can use, and create the other files required for the fits to run. My mentor will provide assistance at any time when I am unable to determine the cause of a problem, resolve any high-level technical difficulties that arise regarding the code of the software, and create Python programs to help with formatting the data when it is otherwise difficult. Once the data is properly formatted and all supplementary files are prepared, I will inform my mentor and have him run the fit either on his personal computer or a more powerful computer that is available at the research site (the software was not easy to install on my computer, and it likely would not have had the computing power regardless). Any issues that arise with the fit will be addressed jointly. My mentor will teach me how to analyze the output files as we go through the planets, allowing me to contribute more to the data analysis after gaining some experience.

Steps

- For each exoplanet, a series of steps will be followed to ensure that the final fit for all data (TESS and archival) accounts for baseline fluctuations of the star and the noise that is a part of each instrument's detection.
- The first step will be to conduct a fit using only the TESS out-of-transit data.
  - The TESS data will be obtained by downloading a zip file from the MAST archive and copying the desired data file into a folder for data preparation.

- My mentor will write a Python program that takes the original data and creates a comma separated values (csv) file, which is the format that the software package (allesfitter) can read when doing a fit.
- My mentor will also write a program to mask the light dips associated with transits, leaving a csv file with data that is reasonably smooth with some gaps where the transits occurred.
- I will execute the programs written by mentor for the original TESS data, converting it into the masked data csv file and then inserting this file into the first working directory that will be modelled.
- I will fill out the other two csv files required for an allesfit, which detail the settings and initial guesses for parameters.
- In this first fit, no astrophysical parameters will be fitted because the goal is to train a Gaussian Process (GP) to the masked data so that when it is run on the full data set, it properly recognizes the transits as deviations from the baseline fluctuations. The GP is basically an approximation for the light curve without the transits, providing this baseline for comparison with the full data set. A GP initial guess will be made based simply on the way the original data looks: a rough length scale and amplitude for baseline fluctuations can be easily seen in preliminary plots.
- The outputs of this first fit will reveal with more precision the GP hyperparameters (the aforementioned length scale and amplitude), which will be used in the next step.

- The second step will be to model the entire TESS data set. This fit will use a constrained GP based on the first fit as the baseline and astrophysical parameters determined from previous studies of the given planet as initial guesses.
    - The outputs of the second fit will provide the astrophysical parameters that we will use as initial guesses in the following step.
- The third step will involve numerous fits: one for every instrument used to study the planet prior to TESS. This will include both photometric instruments (which, like TESS, detect light curves and provide transit information) and radial velocity (RV) instruments (which detect shifts in wavelengths of light from a star which are caused by changes in the speed of the star relative to Earth as a result of the gravitational effects from a planet).
    - The data from each prior instrument will be gathered from previous journal articles and databases and formatted into csv files.
    - A separate allesfit working directory will be created for each instrument, such that each has its own settings.csv, params.csv, and (instrument name).csv files. Each of these working directories will be run in order to determine the GP baseline and noise for each instrument.
- The final step will be to make one working directory with every data file from every instrument, and one settings.csv and one params.csv file, each of which tell the final fit how to account for the previously found astrophysical parameters, GP baselines, and noise.
    - The outputs of this final fit will be analyzed.

Note: Some steps may need to be repeated due to issues such as initial guesses that are too far away from what the data suggests for the fit to properly work.

6. Data Analysis

The majority of the statistics for this project is done within the allesfitter software. It uses a Monte Carlo simulation to take the data it has been fed and create a probability map of the parameter space, and then Markov Chains are followed by "walkers." This is known as an MCMC fit, and here is a more basic description of the process: based on the data, there is a certain probability that the planet being observed has each of a number of characteristics. Starting from the initial guess that is entered, artificial walkers take on step at a time towards a different value for the parameter. The direction of the step is based on the probabilities that are calculated. Thousands of steps are taken by thousands of walkers, with the idea being that eventually, all of the walkers will be near the most likely value for the parameter. These walkers simultaneously walk for every parameter, and hopefully converge around one value for each at the end of the fit. Here is an analogy that makes it easier to understand: each walker begins at some point on a hill. It is easier to walk downhill than uphill, so any given step is more likely to be downhill than uphill. The bottom of the hill represents the most likely value based on the data, and the starting point represents the initial guess. An informed guess may be near the bottom of the hill to begin with, requiring fewer steps for all walkers to converge at the bottom, while an uninformed guess may be near the top of the hill. Even in such a case, enough steps would lead most of the walkers to the bottom of the hill because each step is more likely to be down than up.

This process is plotted automatically within allesfitter, and shows the chains that each of the walkers follow. Convergence is desired at the end of the chain for each parameter. Corner plots show the distribution of each walker's final value for each parameter, and shows a scatter plot for every combination of two parameters. Plots that show parameters pushed against the edges of the boundaries that were initially set indicate a need to reconfigure the initial parameters, and such changes will be made when necessary and refitted. Once any issues are resolved, the runs will be saved and the parameter values noted for future reference and for calculation of JWST observation time possibilities.

7. Bibliography

(2019). Retrieved July 14, 2019, from https://exoplanetarchive.ipac.caltech.edu/index.html

Buckingham, W., Burnham, D., Hill, C. E., King, P. J., Marenbon, J., & Weeks, M. (2017). *The Philosophy Book*. London: DK.

Kempton, -R., E. M., L., J., Louie, R., D., Drake, . . . Barclay. (2018, September 14). A Framework for Prioritizing the TESS Planetary Candidates Most Amenable to Atmospheric Characterization. Retrieved from https://arxiv.org/abs/1805.03671

NASA. (2019). James Webb Space Telescope. Retrieved July 14, 2019, from https://www.jwst.nasa.gov/

Ricker, R., G., Winn, N., J., W., D., Bakos, . . . Ge. (2014, October 28). The Transiting Exoplanet Survey Satellite. Retrieved from https://arxiv.org/abs/1406.0151

NO ADDENDUMS EXIST