CrossMark

A Framework For Optimizing Exoplanet Target Selection For The *James Webb Space Telescope*

Charles D. Fortenbach[1] and Courtney D. Dressing[2]

[1] Department of Physics and Astronomy, San Francisco State University, San Francisco, CA 94132, USA; mailto:cfortenbach@att.net
[2] Astronomy Department, University of California, Berkeley, CA 94720, USA

Abstract

The *James Webb Space Telescope (JWST)* will devote significant observing time to the study of exoplanets. It will not be serviceable as was the *Hubble Space Telescope*, and therefore the spacecraft/instruments will have a relatively limited life. It is important to get as much science as possible out of this limited observing time. We provide an analysis framework (including publicly released computational tools) that can be used to optimize lists of exoplanet targets for atmospheric characterization. Our tools take catalogs of planet detections, either simulated, or actual; categorize the targets by planet radius and equilibrium temperature; estimate planet masses; generate model spectra and simulated instrument spectra; perform a statistical analysis to determine if the instrument spectra can confirm an atmospheric detection; and finally, rank the targets within each category by observation time required. For a catalog of simulated *Transiting Exoplanet Survey Satellite* planet detections, we determine an optimal target ranking for the observing time available. Our results are generally consistent with other recent studies of *JWST* exoplanet target optimization. We show that assumptions about target planet atmospheric metallicity, instrument performance (especially the noise floor), and statistical detection threshold, can have a significant effect on target ranking. Over its full 10 yr (fuel-limited) mission, *JWST* has the potential to increase the number of atmospheres characterized by transmission spectroscopy by an order of magnitude (from about 50 currently to between 400 and 500).

Key words: planets and satellites: atmospheres – space vehicles: instruments – surveys – techniques: spectroscopic

Online material: machine-readable table

1. Introduction

1.1. Background

The launch of the *James Webb Space Telescope (JWST)* is currently scheduled for early 2021.[3] *JWST* will usher in a new era of astronomical observation, studying everything from the history of the universe, to the formation of extrasolar planetary systems, to the evolution of our own solar system. It will complement and extend what has been achieved with the *Hubble Space Telescope (HST)* in a new wavelength regime and with much better sensitivity (Gardner et al. 2006; Kalirai 2018; Madhusudhan 2019).

One of the main mission goals of *JWST* is to "measure the physical and chemical properties of planetary systems, . . . and investigate the potential for life in those systems."[4] For planets outside our own solar system (exoplanets) one of the best ways to investigate the potential for life is to study the planetary atmospheres.

JWST will significantly advance studies of exoplanet atmospheres by providing access to a wide variety of molecular spectral features. For example, the main features of an Earth-like atmosphere in the 0.7–5.0 μm range are H_2O (near 2.8 μm, and 3.2 μm), and CO_2 (4.2 μm), and in the 4.0–20 μm range are CO_2 (15 μm), O_3 (9.6 μm), CH_4 (7.8 μm), H_2O (5.9 μm), and HNO_3 (11.2 μm) (Kaltenegger & Traub 2009). *JWST* has the capability to take high quality spectra over these wavelength ranges. The target exoplanets may well have atmospheres with significantly different composition, but many interesting molecular features are still expected to be observed.

1.1.1. The Mission

The spacecraft will orbit the Sun at the second Sun–Earth Lagrange point, L2. This is a meta-stable point about 1.5 million km away from the Earth along the Sun–Earth line outward (in the opposite direction) from the Sun. This orbital geometry allows the large sunshield to provide protective cooling for the telescope with minimal maneuvering, which would be very difficult, if not impossible, if the spacecraft were in a typical low-Earth orbit, as with *HST*.

---

[3] https://www.nasa.gov/press-release/nasa-completes-webb-telescope-review-commits-to-launch-in-early-2021
[4] https://jwst.nasa.gov/science.html

Due to this orbital location, *JWST*'s lifetime will be limited. There is currently no plan for on-orbit servicing as there was with *HST*. The spacecraft will have a nominal five-year mission, but will carry ten years worth of fuel to enable an extended mission.

We can estimate the time available for exoplanet observing over the spacecraft's ten-year fuel-limited lifetime. NASA expects *JWST* to be in routine science mode roughly six months after launch.[5] This allows for cooling down the spacecraft and doing various calibrations and maneuvering operations. This leaves approximately 83,220 hr available for the overall observing program over the 10 yr extended mission.

The question is: out of this 83,220 hr, how much time will be available for exoplanet studies and in particular for transmission spectroscopy?

The purpose of our work is to provide a tool to rank an exoplanet target catalog (simulated or actual, delivered by *Transiting Exoplanet Survey Satellite* (*TESS*) or other precursor mission) by atmospheric observability. We are not attempting to justify a particular time allocation from a bottom-up scientific perspective. We are only trying to provide a rough estimate of the amount of time that might be available for exoplanet transmission spectroscopy based on a top down view of high-level mission priorities.

Our ranking study is not in any way intended to be an observing proposal. Any strategic survey program related to our ranking work would need to go through the proper reviews and would be competitively chosen by the *JWST* Telescope Allocation Committee and/or the Director.

One of the early documents providing guidance on the allocation of *JWST* mission time to various scientific programs was the *JWST* Science Operations Design Reference Mission report (SODRM, released 2012 September 27).[6] The 2012 SODRM was an early estimate of how observing time might be allocated for the first year of *JWST* science operations. This document was built up from a diverse set of science and calibration programs.[7] It showed a notional allocation of 16.1% (13,363 hr) of the total mission time to exoplanet study (after removing the 6-month commissioning activity).

Since the release of the 2012 SODRM, interest in exoplanet studies and, in particular, in exoplanet atmospheric characterization has grown. This has led to a shift in emphasis of *JWST* observational priorities. A recent white paper (Greene et al. 2019) describes how guaranteed time observations (GTOs) and early release science (ERS) will "advance understanding of exoplanet atmospheres and provide a glimpse into what transiting exoplanet science will be done with *JWST* during

its first year of operations. . . . Approximately 3700 hr of GTO and an additional ∼500 hr of Director's Discretionary ERS observations have been accepted for *JWST* Cycle 1 (for general science, including exoplanet studies). This is ∼50% of the time available in the first year of science operations."

Greene et al. (2019) go on to say that "The transiting planet observations in the Cycle 1 GTO and ERS (Bean et al. 2018) programs will enable a large step forward in the characterization of exoplanet atmospheres . . . these programs (which include multi-wavelength transmission, emission, and phase curve observations of 27 transiting planets) sum up to 816 hr, 19% of the scheduled GTO + ERS observing time." If we extrapolate this to a full year program we might see as much as 1600 hr dedicated to transit spectroscopy; and if extrapolated to the full 10 yr (less commissioning) fuel-limited mission, we might see a total program of as much as 15,000 hr for exoplanet transit spectroscopy.

Additionally, a significant GTO + ERS time allocation will be made for coronagraphic imaging and direct spectroscopy of young planets (Beichman et al. 2019), but that is not our focus.

Greene et al. (2019) also mentioned that, "*JWST* may well characterize the atmospheres of over 50 transiting planets in its first year of science operations." Extrapolating, this suggests that *JWST* could possibly characterize the atmospheres of as many as 475 transiting planets over the course of its full mission. In some cases this would include multiple-transit observations and revisits of the same planet with multiple instruments for broader wavelength coverage.

It has been noted that transmission spectroscopy "is expected to be the prime mode for exoplanet atmospheric observations and provides the best sensitivity to a wide range of planets" (Kempton et al. 2018). It would seem likely that a somewhat greater emphasis will be placed on transmission spectroscopy than emission (eclipse) spectroscopy.

Of course all of the *JWST* instruments will be involved in the observing program; however, for our baseline ranking we will focus exclusively on transmission spectroscopy with the NIRSpec G395M, operating in Bright Object Time Series (BOTS) mode (the basis for selection of this instrument/mode is discussed further in Section 2.3.1). This should give us a reasonable ranking of the atmospheric observability of our target planet catalog at least for the NIRSpec G395M wavelength range and capabilities. Future work could rank the targets with NIRISS SOSS, or other instrument/modes. An optimal target ranking using a mix of instruments would be a complex problem; but one that could potentially be studied with our code. Such a study is beyond the scope of our current work.

It is important to emphasize that the available observing time does not affect the target rankings; it only comes into play in establishing the cut-off line for the list of ranked targets. In order to set this cut-off line for our study, we will assume an overall program of 8300 hr for <u>transmission</u> spectroscopy. This is a little over half (∼55%) of the total 15,000 hr transit

---

[5]   https://jwst.nasa.gov/faq.html
[6]   http://www.stsci.edu/files/live/sites/www/files/home/jwst/about/history/science-operations-design-reference-mission-sodrm/_documents/SODRM-Revision-C.pdf
[7]   http://www.stsci.edu/jwst/about/history/science-operations-design-reference-mission-sodrm

spectroscopy program extrapolated from the Cycle 1 GTO and ERS program allocation.

1.2. Aim of This Work

Given the limited lifetime of *JWST*, the scarce resource of observing time must be allocated as efficiently as possible. In particular, for the study of exoplanet atmospheres we want to prioritize targets where the exoplanet atmospheres have the best chance of successful detection and characterization.

The scientific community needs a tool to implement this target prioritization in a robust and efficient way. We provide a framework for analysis and an associated computer program to assist with ranking targets of interest for *JWST* detected by *TESS* and other precursor efforts.

In addition, our ranking tools will help to prioritize targets for high precision radial velocity (RV) follow-up to further constrain the masses of the target planets.

1.3. Related Work

A number of studies have addressed the range of issues associated with optimizing *JWST* targets for atmospheric characterization. These issues include: precursor missions, synthetic target surveys, exoplanet mass–radius relationships, model atmosphere spectra, simulated instrument spectra, and other target ranking schemes.

1.3.1. Synthetic Target Surveys

The earliest simulation of exoplanet yield from a space-based all-sky survey of bright stars was described in Deming et al. (2009). This was specifically aimed at simulating the *TESS* survey yield. It was a particularly notable effort given that the *Kepler* mission had not even been launched at the time of their work. This was an exercise in scaling, and extrapolating from the limited data available.

We have chosen to use the target catalog developed by Sullivan et al. (2015) (the "Sullivan" catalog) as the basis of our analysis and code development. This catalog is a simulation of expected *TESS* planet detection results based on existing occurrence rate information. The *TESS* exoplanet statistics have been studied recently by several groups, including Bouma et al. (2017), Muirhead et al. (2018), Barclay et al. (2018), and Ballard (2019). These studies considered *TESS* mission extension, yields of planets around M-dwarf hosts, and updates of planet yields using the actual *TESS* catalog of target stars.

These updated studies do show some significant differences with the Sullivan work. Specifically, by accounting for co-planarity of planets in multi-planet systems, Ballard (2019) found that *TESS* would detect up to 50% more M-dwarf planets than predicted by Sullivan. Bouma et al. (2017) estimated 30% fewer Earths and 22% fewer super-Earths for the primary 2×10^5 target stars that will be sampled at a 2 minute cadence,

and Barclay et al. (2018) predicted 36% fewer Earths and 60% fewer super-Earths for the primary target stars.

Our study was too advanced to incorporate the results of these new simulated catalogs into our baseline target ranking. The primary goal of our project was to develop a framework for ranking planets for study with *JWST*. With modest effort, our code can be adapted to work with any input planet sample including the actual *TESS* detections.

1.3.2. Target Ranking Studies

A number of studies have considered the issue of finding an optimal target set for atmospheric characterization by *JWST*. We have previously mentioned the Deming et al. (2009) study. In addition to modeling the performance of the Mid-Infrared Instrument (MIRI) and the Near-Infrared Spectrograph (NIR-Spec), they coupled their simulated *TESS* target yield to their sensitivity model. They found that *JWST* should be able to characterize dozens of *TESS* super-Earths with temperatures above the habitable range. They also found that *JWST* should be able to measure temperature and identify absorption features in one to four habitable Earth-like planets orbiting lower-main-sequence stars.

They asserted that, "although the number of habitable planets capable of being characterized by *JWST* will be small, large numbers of warm to hot super Earths and exo-Neptunes will be readily characterized by *JWST*, and their aggregate properties will shed considerable light on the nature of icy and rocky planets in the solar neighborhood."

More recently Crossfield & Kreidberg (2017) determined that the prominence of features in the transmission spectrum for a warm-Neptune exoplanet is related to its equilibrium temperature and its bulk H/He mass fraction. They were able to construct an analytical relation to estimate the overall observing time needed to distinguish a gas giant's transmission spectrum from a flat line. They suggested that the atmospheric trends they describe could result in a reduction in the number of *TESS* targets with atmospheres that could be detected by *JWST* by as much as a factor of eight.

Howe et al. (2017) explored the optimization of observations of transiting hot Jupiters with *JWST* to characterize their atmospheres. They constructed forward model sets for hot Jupiters, exploring parameters such as equilibrium temperature and metallicity, as well as considering host stars with a wide brightness range. They computed posterior distributions of model parameters for each planet with all of the available *JWST* instrument modes and various programs of combined observa-tions. From these simulations, trends emerged that provide guidelines for designing a *JWST* observing program.

Morley et al. (2017) offered a study of atmospheric detection for the seven TRAPPIST-1 planets, GJ 1132b, and LHS 1140b. These are some of the smallest planets discovered to date that might have atmospheres within the detection capability of

JWST. This was not strictly a target ranking study, but it did involve generating model atmospheres and simulating *JWST* instrument performance. They considered the observability of each planet's atmosphere in both transmission and emission. GJ 1132b and TRAPPIST-1b are excellent targets for emission spectroscopy with MIRI, requiring less than 10 eclipse observations. Seven of the nine planets are good candidates for transmission spectroscopy. Using estimated planet masses they determined that less than 20 transits would be required for a 5σ detection of a transmission spectrum.

Another recent study (Louie et al. 2018) was aimed at understanding the suitability of expected *TESS* planet discoveries for atmospheric characterization using *JWST*'s Near-Infrared Imager and Slitless Spectrograph (NIRISS) by employing a simulation tool to estimate the signal-to-noise ratio (S/N) achievable in transmission spectroscopy. The tool was applied to predictions of the *TESS* planet yield and then the S/N for anticipated *TESS* discoveries was compared to estimates of S/N for 18 known exoplanets. They analyzed the sensitivity of the results to planetary composition, cloud cover, and the presence of an observational noise floor. Several hundred anticipated *TESS* discoveries with radii $1.5R_{\oplus} < R_p \leqslant 2.5R_{\oplus}$ will produce S/N higher than currently known exoplanets in this radius regime. In the terrestrial planet regime, only a few expected *TESS* discoveries will result in higher S/N than currently known exoplanets.

A study, Kempton et al. (2018) (the "Kempton study" or "Kempton"), was published recently that is similar to our work in its general goal of finding an optimal target set for *JWST*. The authors use a set of two analytic metrics, quantifying S/N values in transmission and thermal emission spectroscopy to rank the target planets in the Sullivan catalog. They use the S/N predictions from the *JWST*/NIRISS simulation performed by Louie et al. (2018) as the basis for their transmission metric. They determine a sample of roughly 300 transmission spectroscopy targets that meet the threshold values of their metrics for observation.

We have organized the remainder of this paper as follows: In Section 2, we describe our analysis framework and the high level structure of our ranking code. We then validate the analysis and code in Section 3. Next, we discuss our results including a baseline target ranking run in Section 4. Finally, we present our conclusions and opportunities for further study in Section 5.

2. Analysis Framework and the JET Code

We have developed an analysis framework that takes a planet detection target catalog (simulated or actual) and processes it to result in a list of prioritized exoplanet targets for atmospheric characterization by *JWST*.

Our analysis of a target catalog proceeds in a straightforward manner. We first set target selection parameters and other values (e.g., *JWST* instrument/mode, target list limits, atmosphere model

equations of state, etc.). Next, we take the target catalog data, and for each of the targets determine various planetary system parameters: orbital semimajor axis, planet equilibrium temperature, planet mass, etc. We then make an estimate of the number of transits observable in a 10 yr mission, given target position on the sky and spacecraft pointing constraints, for each target. Then we divide the parameter space into seven demographic categories, by planet radius and equilibrium temperature, and categorize the targets.

We then generate model transmission spectra for each target. In this paper we focus exclusively on transmission spectroscopy because, as discussed in Kempton et al. (2018), this is expected to be the best approach for observations of exoplanet atmospheres and provides the best sensitivity to a wide range of planets. We consider two bounding cases for model atmospheres for each target: low metallicity with no clouds, and high metallicity with a mid-level cloud deck. This should span the range of atmospheres that we might encounter, and their potential detectability (see Section 2.2.6).

Next, we make sure that the host star is not so bright as to saturate the detector of the particular instrument under consideration. If not, then for each of the targets and for each model atmosphere, we run an instrument simulation which includes noise effects. We then determine the number of transits required for a high confidence detection.

Next, we check to see if the number of transits required for detection is available in a ten-year mission. If so, we can then determine the total observing time for the target, with all of the overheads included, for each atmosphere case. From this we can determine an average total observing time for that particular target.

Finally, we sort the targets by category and then rank them within each category by the average total observing time needed for detection. No attempt is made to prioritize one demographic category over another, but future users could modify the code according to their own priorities

We will describe each of these analysis steps in more detail in the following subsections. In addition, we have developed a suite of computer tools: the JWST Exoplanet Targeting code (JET) that embodies our analysis procedure.

2.1. Top Level Analysis Framework and JET Architecture

The analysis (and code) can be divided logically into five main parts: (1) General input, (2) Generating model spectra, (3) Detecting the spectra with *JWST*, (4) Sorting and ranking the target list, and (5) Controlling the program execution. Figure 1 presents a flowchart of this top level analysis framework and program architecture.

The complete source code, detailed installation instructions, and sample input files for JET are available via Github (https://github.com/cdfortenbach/JET). JET has been written
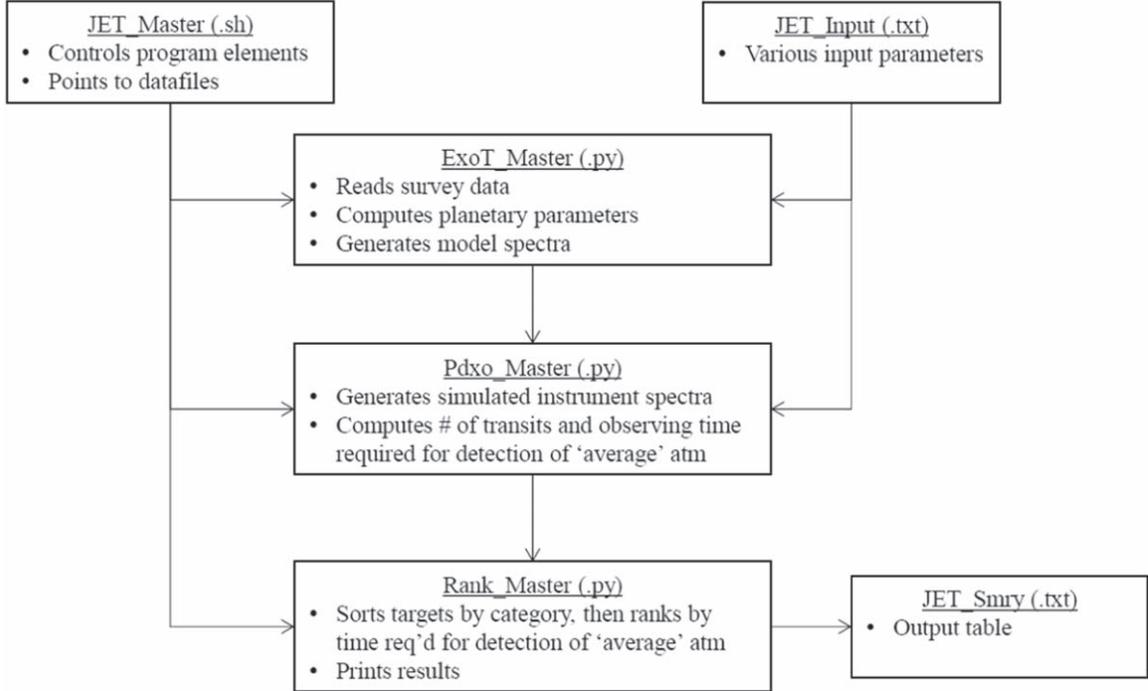
Figure 1. Top level `JET` program architecture.

in `Python`, but it incorporates a fully compiled executable of the `Exo-Transmit` atmospheric modeling code which was written in C (see Section 2.2.6). The `JET` code is designated as open source under the GNU General Public License.

2.2. Generating Model Transmission Spectra (ExoT_Master)

The first major step in the analysis process is to generate model transmission spectra for the various targets. This task is carried out by the `JET` program element: `ExoT_Master`, shown in flowchart form in Figure 2.

2.2.1. Reading the Survey Data

Our analysis begins by reading in a planet detection target catalog. As we discussed in Section 1.3.1, there have been several attempts at developing a synthetic detection catalog that would emulate a full *TESS* (or other mission) program catalog. As previously mentioned, we have adopted the format of the catalog described in Sullivan et al. (2015). Of course, when an actual *TESS* (or other mission) catalog is available, that would be used for a `JET` production/science run.

2.2.2. Estimating Planet Masses

Given that the precursor missions are transit surveys, they will not directly provide planet mass data. `JET` estimates planet

masses using an approach implemented in the `Forecaster` package described by Chen & Kipping (2017). To estimate the median values of planet mass, we use two power laws that can be derived from this work. We note that our power law derivation here is consistent with the analysis of Louie et al. (2018).

For $R_p < 1.23\ R_\oplus$:

$$M_p = 0.9718\ (R_p)^{3.58}. \tag{1}$$

For $1.23 R_\oplus \leqslant R_p \leqslant 10 R_\oplus$:

$$M_p = 1.436\ (R_p)^{1.70} \tag{2}$$

where M_p is the planet mass relative to the mass of the Earth ($M_\oplus \cong 5.9736 \times 10^{24}$ kg) and, R_p is the planet radius relative to the radius of the Earth ($R_\oplus \cong 6.37814 \times 10^6$ m).

We have set an upper limit for R_p of 10 $R_\oplus$. The mass–radius relationship described in the Chen & Kipping (2017) study is ambiguous for radii above this level. This is due to the very wide range of mass for planets with radii similar to that of Jupiter. While there are target planets in the Sullivan catalog (and there will be in actual catalog data) that have larger radii, we will flag them as outside the range of our analysis. This is equivalent to imposing a planet mass upper limit of roughly 72 $M_\oplus$ (roughly 3/4 the mass of Saturn) for our analysis.
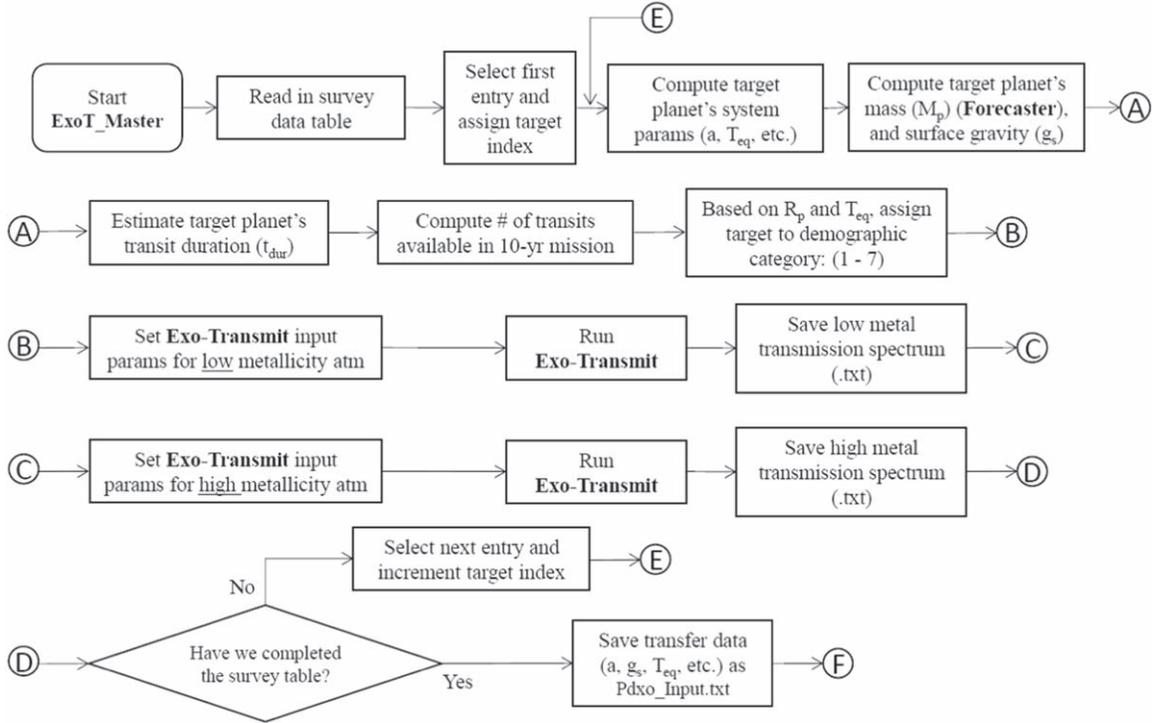
Figure 2. Analysis and code flow for `JET` program element `ExoT_Master`.

2.2.3. Deriving Other Planetary System Parameters

In addition to estimating planet mass, we need to determine certain other planetary system parameters for each of the target planets. These parameters include the semimajor axis of the planet's orbit, the planet's equilibrium temperature, the planet's surface gravity, and the transit duration. These parameters are derived directly from the catalog data, given certain assumptions.

Semimajor axis (a_{au}): The relative insolation of the planet, S, is given by Equation (13) of Sullivan et al. (2015). By rearranging this equation we can determine the semimajor axis of each of the target planets as follows:

$$a_{au} = \left(\frac{R_*}{\sqrt{S}}\right)\left(\frac{T_{eff}}{T_{eff,\odot}}\right)^2 \tag{3}$$

where S is the stellar insolation at the top of the planet's atmosphere in $S_\oplus$ units, T_{eff} is the effective temperature of the host star in K, while R_* is the radius of the host star in solar radii, $T_{eff,\odot}$ is the Sun's effective temperature $\cong 5777$ K, and a_{au} is the semimajor axis of the planet's orbit in astronomical units.

Surface gravity (g_s): The target planet's surface gravity is needed to calculate the pressure scale height, and is given by:

$$g_s = \frac{GM_p}{(R_p)^2} \tag{4}$$

where G is the constant of universal gravitation, 6.67428×10^{-11} N m^2 kg^{-2}; and g_s is the surface gravity in m s^{-2}.

Equilibrium temperature (T_{eq}): The planet's equilibrium temperature is also needed in the generation of spectra of model atmospheres. We are assuming circular obits, zero albedo, and full day-night heat distribution through the atmosphere. We can estimate equilibrium temperature as follows (see Equation (12) of Sullivan et al. 2015):

$$T_{eq} = T_{eff}\sqrt{\frac{R_*}{2a}} \tag{5}$$

where T_{eq} is the planet's equilibrium temperature in K, and a is the orbital semimajor axis in solar radii. This, of course, ignores any potential greenhouse effect.

Transit duration (t_{dur}): The transit duration for each target planet is a key element in our analysis. For our purposes, we define the transit duration to be the interval between the first (t_I) and fourth (t_{IV}) contacts as shown in Figure 2 of Winn (2010). This interval is otherwise known as t_{14}.

If we assume a circular orbit, with an inclination of 90 deg., we can estimate the total transit time by the method described in Section 2.4 of Winn (2010). The full transit duration is given by:

$$T_{tot} \cong \left(1 + \left(\frac{R_p}{R_*}\right)\left(\frac{R_\oplus}{R_\odot}\right)\right)\frac{R_*(P*24)}{\pi a} \tag{6}$$

Table 1
Target Planet Categories

Category	Description	T_{eq} Range	R_p Range
1	Cool Terrestrials	$T_{eq} < 400$ K	$R_p < 1.7 R_\oplus$
2	Warm Terrestrials	$400\ \text{K} \leqslant T_{eq} \leqslant 800$ K	$R_p < 1.7 R_\oplus$
3	Hot Terrestrials	$T_{eq} > 800$ K	$R_p < 1.7 R_\oplus$
4	Cool sub-Neptunes	$T_{eq} < 400$ K	$1.7 R_\oplus \leqslant R_p \leqslant 4.0 R_\oplus$
5	Warm sub-Neptunes	$400\ \text{K} \leqslant T_{eq} \leqslant 800$ K	$1.7 R_\oplus \leqslant R_p \leqslant 4.0 R_\oplus$
6	Hot sub-Neptunes	$T_{eq} > 800$ K	$1.7 R_\oplus \leqslant R_p \leqslant 4.0 R_\oplus$
7	sub-Jovians	⋯	$R_p > 4.0 R_\oplus$

Note. Category 7 captures all catalog targets of large radii. Targets with $R_p > 10.0 R_\oplus$ are flagged and no transit detection calculations are made on them. These targets drop to the bottom of the ranking for Category 7.

and,

$$t_{dur} = t_{14} = T_{tot} \qquad (7)$$

where P is the planet's orbital period in days, t_{dur} is the transit duration in hours, and again a is the orbital semimajor axis in solar radii.

2.2.4. Categorizing Targets

The next step in our analysis process is to categorize the target planets. We have chosen to divide the catalog into seven categories along planet radius and equilibrium temperature dimensions with breaks at planet radii of 1.7 and 4.0 $R_\oplus$, and at equilibrium temperatures of 400 and 800 K (as shown in Table 1). In this work, as mentioned in Section 2.2.2, we are focusing our attention on planets with generally smaller radii ($R_p < 10 R_\oplus$).

The category divisions are somewhat arbitrary; however, the break at an R_p of 1.7 $R_\oplus$ is based on recent exoplanet population studies (Fulton et al. 2017; Fulton & Petigura 2018) that indicate that a gap exists in the exoplanet population around this radius.

2.2.5. Determining JWST Observing Constraints

JWST has viewing constraints that are defined by its orbit and configuration. The spacecraft will orbit the Sun (in the ecliptic plane) at the L2 position with a one year period. At all times the large sunshield must be positioned between the Sun and the science instruments to maintain the very cold environment necessary for infrared observations. The sunshield geometry creates limitations on where the spacecraft can point at any particular time and for how long. *JWST* can point to solar elongations between 85 and 135 deg. It can also point to any location in the 360 deg. circle perpendicular to the Sun line. This defines a large annulus where *JWST* can observe at any given time (the field of regard). Targets at low ecliptic latitudes are particularly impacted by these constraints. This observing geometry is well described in Figure 1 of the Space Telescope Science Institute's (STScI) on-line User Documentation for *JWST* Target Viewing Constraints.[8]

For a target at a particular ecliptic latitude there are only a certain number of days per year that *JWST* can observe it. This means that depending on the position and transit period of the target system, not all transits may be observable. This could be a problem for low S/N targets (meaning many transit observations necessary) that have long periods, and that are at low ecliptic latitudes.

For each target we compute the number of transits that would be observable over the full 10 yr fuel life of the spacecraft. If the transit requirement for detection exceeds the number of transits available we raise a flag in the output.

Within ExoT_Master, our algorithm for computation of transits available over the 10 yr mission is only approximate. We would need to include the absolute transit timing and absolute *JWST* orbital position to more accurately determine this parameter for a given target. However, our estimate should be acceptable for our purposes.

We first compute the number of mission days available in the 10 yr (fuel limited) mission, less a six month commissioning period:

$$t_{mission} = (10 - 0.5) * 365\ \text{days}. \qquad (8)$$

Then we compute the number of transits available (but not necessarily observable) in a 10 yr mission:

$$nt_{10\ yr,\ available} = (1/P) * t_{mission} \qquad (9)$$

where P is the planet's orbital period in days.

Next, we convert the target's decimal coordinates into hours–minutes–seconds coordinates. This is done using the Python astropy SkyCoord routines.[9] Then we convert to Ecliptic coordinates using the Python PyEphem package.[10]

---

[8] https://jwst-docs.stsci.edu/jwst-observatory-hardware/jwst-target-viewing-constraints

[9] http://docs.astropy.org/en/stable/api/astropy.coordinates.SkyCoord.html

[10] https://rhodesmill.org/pyephem/index.html

We then take advantage of the STScI analysis that produced an estimate of observable days versus ecliptic latitude for *JWST* shown in Figure 2 of the previously referenced STScI on-line User Documentation for *JWST* Target Viewing Constraints. For the target's particular ecliptic latitude we can then estimate the number of days out of 365 that the target will be observable: the observable days factor ($f_{obsdays}$). Below 45 deg. this observability comes as two shorter time periods separated by six months. Above 45 deg. one longer viewing period is available, increasing until the continuous viewing zone is reached at approximately 85 deg. ecliptic latitude. We are not considering the impact of the six month split for observations below an ecliptic latitude of 45 deg. We only consider aggregate observable days available.

Finally, we can estimate the rough upper bound on the total number of transits that would be observable in the 10 yr mission as:

$$nt_{10\,yr} = nt_{10\,yr,\,available} * f_{obsdays}. \quad (10)$$

For further information on target visibility we recommend the *JWST* General Target Visibility Tool (GTVT). This is a command-line Python tool that provides quick-look assessments of target visibilities and position angles for all *JWST* instruments.[11]

2.2.6. Generating Model Transmission Spectra with Exo-Transmit

We have chosen to implement the Exo-Transmit code for our study. Kempton et al. (2017) presents a detailed description of Exo-Transmit, an open source code for generating model exoplanet transmission spectra. This is an extension of a super-Earth radiative transfer code originally described by Miller-Ricci et al. (2009), and Miller-Ricci & Fortney (2010).

Exo-Transmit is a flexible tool aimed at calculating transmission spectra for a wide range of exoplanet size, surface gravity, equilibrium temperature, and atmospheric composition. The code essentially solves the equation of radiative transfer for absorption of the host star's light as it travels through the planet's atmosphere during a transit. It was originally developed to work with low-mass exoplanets, but it can be used to model giant planet transmission spectra as well. The models can be set up with or without clouds, and can include the effects of Rayleigh scattering, and collision induced absorption.

In general we will not know the characteristics of the target planet's atmosphere. We might be able to make an educated guess, but our approach is to assume that we do not know.

Figure 5 of Kempton et al. (2017) showed that spectral line strength is related to atmospheric metallicity and cloud levels. High metallicity (related to high mean molecular weight) and the presence of higher altitude (lower pressure) level clouds reduce the strength of spectral features.

---

[11] https://jwst-docs.stsci.edu/jwst-other-tools/jwst-target-visibility-tools/jwst-general-target-visibility-tool-help

Table 2
JET Baseline Run—Input Parameters

Parameter	Value
Starting target from catalog for this run:	1
Ending target from catalog:	1984
JWST instrument:	NIRSpec G395M
Wavelength short limit (μm):	2.87
Wavelength long limit (μm):	5.18
J_{mag} limit (T_{eff} = 10,000 K):	6.2
J_{mag} limit (T_{eff} = 5000 K):	6.8
J_{mag} limit (T_{eff} = 2500 K):	7.4
Detector linear response limit (% FW):	80
Noise floor (nfloor) (ppm):	25
R value of sim (Res):	100
dBIC samples for each ntr grid pt.:	2000
Detection threshold (dBIC):	10
Free model spectrum BIC parameters:	5
Equation of State (lo metal atm):	eos_5Xsolar_cond
Cloud_lo (Pa):	0
Equation of State (hi metal atm):	eos_1000Xsolar_cond
Cloud_hi (Pa):	10,000
Out of transit factor (% tdur):	100
+ Out of transit "timing tax" (s):	3600
Slew duration avg. (s):	1108
SAMs: small angle maneuvers (s):	0
GS Acq: guide star acquisition(s) (s):	284
Targ Acq: target acquisition (s):	492
Exposure Ovhd: factor 1:	0.0393
Exposure Ovhd: factor 2 (s):	26
Mech: mechanism movements (s):	110
OSS: Onbd Script Sys. compilation (s):	65
MSA: NIRSpec MSA config. (s):	0
IRS2: NIRSpec IRS2 Detector setup (s):	0
Visit Ovhd: visit cleanup (s):	110
Obs Ovhd factor (%):	16
DS Ovhd (s):	0
RunExoT (Y/N):	Y
RunPdxo (Y/N):	Y
RunRank (Y/N):	Y

Note. The "EOS (Equation of State)" file nomenclature is defined in the Exo-Transmit user manual. The detection threshold (dBIC) is discussed in detail in Section 2.3.2.

We set up two bounding cases of observability: (1) a relatively easy to detect low-metallicity atmosphere (5× Solar) with no clouds, and (2) a more difficult to detect high-metallicity (1000× Solar) atmosphere with a low (100 mbar) cloud deck. In each case the condensation and removal via rain-out of molecules (excluding graphite) from the gas phase is included. The specific Exo-Transmit Equation of State (EOS) input file names and other input parameters for the two atmosphere cases are given in Table 2.
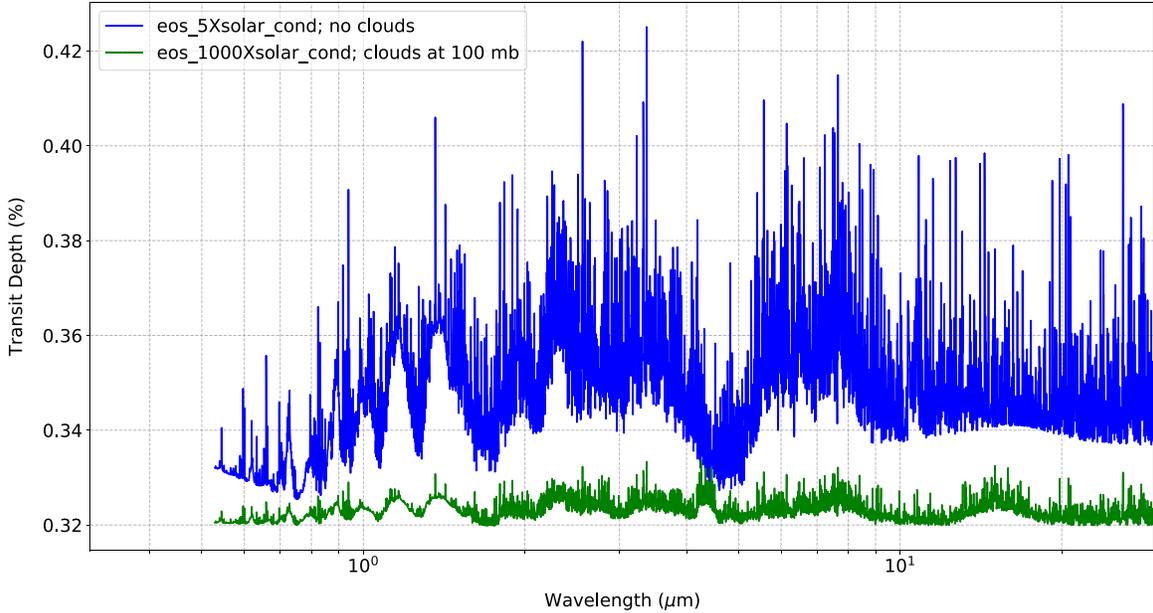
Figure 3. Model transmission spectra for Sullivan Target 1292. The spectra show the full resolution computed by `Exo-Transmit`. Note the substantial difference (on the order of 100–300 ppm) between the low and high metallicity atmospheric spectra.

`Exo-Transmit` provides the user with over 60 different EOS file choices, modeling various conditions of metallicity, condensation, etc. We have chosen files with solar constituents, and with high and low levels of metal abundance. These particular files also allow for modeling condensation and removal via rain-out of molecules from the gas phase—but they do not include condensation and rainout of graphite. This was an arbitrary choice. Alternative EOS files are available that would allow consideration of the rain-out of graphite, which would deplete the model atmosphere of all carbon-bearing species at low temperature.

As described in Kempton et al. (2017), `Exo-Transmit` "allows the user to incorporate aerosols into the transmission spectrum calculation following one of two ad-hoc procedures. The first is to insert a fully optically thick gray cloud deck at a user specified pressure. The second is to increase the nominal Rayleigh scattering by a user-specified factor." We are using the first procedure.

For a given target, and for each of the atmosphere cases, our `ExoT_Master` subprogram will write the appropriate set of planetary parameters to a transfer input file (`userInput.in`). The parameters written include the equilibrium temperature (we assume an isothermal atmosphere); the EOS (defines the metallicity and other atmospheric characteristics, e.g., rain-out of condensates); planet surface gravity; planet radius; host star radius; and in the case of clouds, the pressure at the cloud top. We have assumed the default Rayleigh scattering factor (1.0), and default collision induced absorption factor (1.0).

For each target/atmosphere case our `Python` program `ExoT_Master` calls `Exo-Transmit` to generate model spectra. For each case, `Exo-Transmit` generates a moderate resolution ($R \sim 1000$) spectrum across a wide wavelength range (0.3–30 μm). Figure 3 presents an example of the model spectra generated for Sullivan Target 1292, a Category 1 planet (Cool Terrestrial) with $R_p \sim 1.52 \, R_\oplus$, and $T_{eq} \sim 300$ K.

We have chosen Target 1292 arbitrarily, but it is a reasonable example case. The catalog data and `JET`-computed parameters for this target as well as its location (as simulated) on the sky are shown in Figure 4.

2.3. Detecting Transmission Spectra with JWST (Pdxo_Master)

The next major step in the analysis process is to determine if the *JWST* instruments can detect the target transmission spectra. Where "detect" indicates that the simulated spectrum fits the model spectrum better than a flat line to a statistically significant level. This task is carried out by the `JET` program element: `Pdxo_Master`. Figure 5 presents a flowchart for the `Pdxo_Master` program element.

2.3.1. Generating Simulated Transmission Spectra with PandExo

For each target/atmosphere case, `Pdxo_Master` first sets up a `Python` input dictionary for the instrument simulator, `PandExo`.

Batalha et al. (2017) presented an open-source `Python` package, `PandExo`, to provide community access to *JWST* instrument noise simulations. `PandExo` relies on the STScI's

| RA | Dec | R_p | P | S | R_* | T_eff | J_mag | | a | T_eq | g_s | M_p | t_{dur} | nt_{10yr} |
[deg]	[deg]	[$\mathrm{R}_\oplus$]	[days]	[$\mathrm{S}_\oplus$]	[$\mathrm{R}_\odot$]	[K]	[mag]		[$\mathrm{R}_\odot$]	[K]	[m s^{-2}]	[$\mathrm{M}_\oplus$]	[hrs]	
248.8	-29.47	1.52	14.27	1.4	0.25	3351	9.1		15.3	300	12.4	2.92	1.8	68
			survey data								*JET computed*			
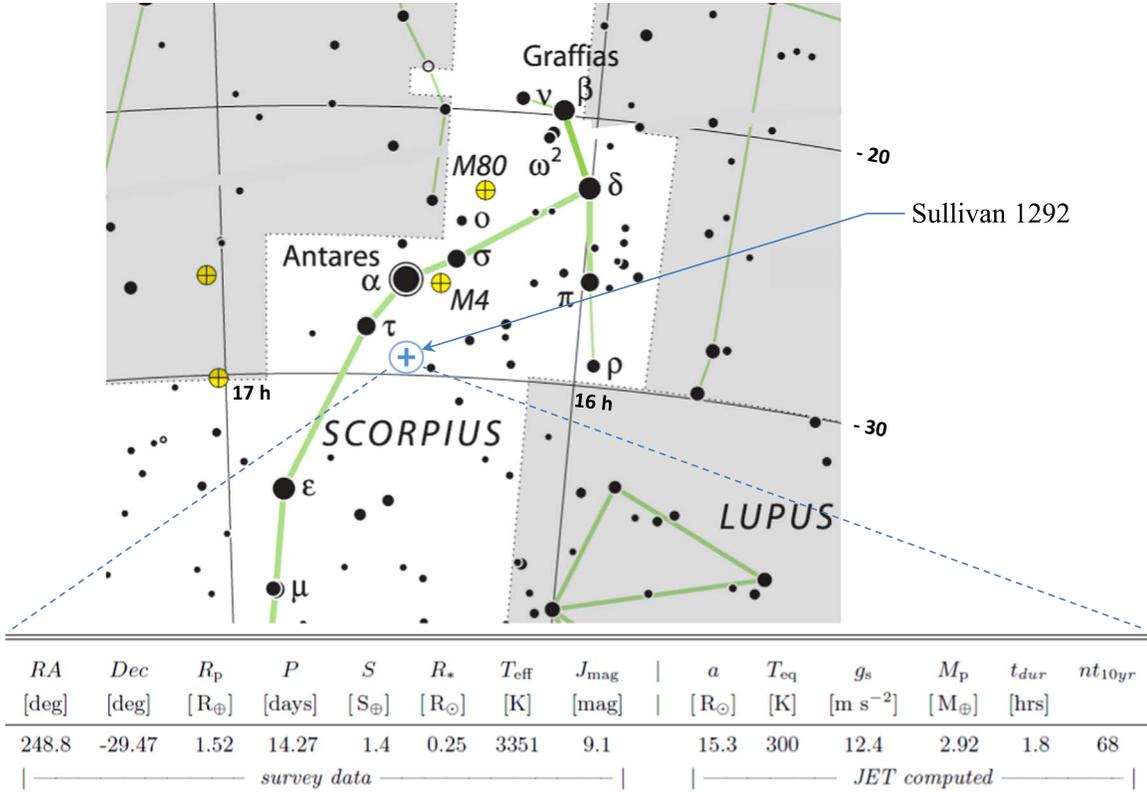
Figure 4. We have arbitrarily chosen Sullivan Target 1292 as a reasonable example case. We show the catalog data and the `JET` computed parameters for the target, as well as its location (as simulated) on the sky. The chart shows relative V_mag. For reference, Antares has $V_\mathrm{mag} = 0.9$, and Sullivan 1292 has $V_\mathrm{mag} = 12.7$. Image source: Mag 7 Star Atlas Project by Andrew L. Johnson, Cloudy Nights LLC (https://www.cloudynights.com/articles/cat/articles/observing-skills/free-mag-7-star-charts-r1021), licensed under a Creative Commons license: CC BY-NC-SA 2.0 (https://creativecommons.org/licenses/by-nc-sa/2.0/).

Exposure Time Calculator, `Pandeia` (Pickering et al. 2016). It can be used as both an online tool[12] and a `Python` package for generating instrument simulations of *JWST*'s NIRSpec, Near-Infrared Camera (NIRCam), NIRISS, MIRI, and *HST*'s WFC3. `PandExo` has been shown to be within 10% of the *JWST* instrument team's noise simulations and has become a trusted tool of the scientific community.

`PandExo` takes as input the target catalog data, the system parameter data, and the model spectrum for the particular target/atmosphere in question. It also takes input of the top level settings that define the particular instrument considered, its wavelength range, noise floor value, and brightness limits, among other things.

The brightness limits for NIRSpec are given in Table 3 of the STScI's on-line User Documentation for NIRSpec BOTS Spectroscopy.[13] The full-well *J*-magnitude brightness limits as a function of a target's T_eff are given for each mode/filter. The SUB2048 subarray is assumed in all cases except for

PRISM/CLEAR values that were determined using SUB512. These values are for gain = 2, and a full well depth of 65,000. The `JET` user manually inputs the magnitude limits for T_eff values of 2500, 5000, and 10,000 K for the chosen disperser. `JET` then uses an interpolation routine to determine the value of the full-well magnitude limit for any particular target T_eff over the full range of targets considered in the survey/catalog. For targets with T_eff greater than 10,000 K, or less than 2500 K, the `JET` code will extrapolate the full-well magnitude limit versus T_eff curve, holding a constant endpoint slope. The detector percent-full-well limit (e.g., 80%) is a user specified value. We use a simple correction formula to estimate the *J*-magnitude limit for a particular percent-full-well condition:

$$J_\mathrm{mag,\,limit}(\mathrm{X\%\ FW}) = J_\mathrm{mag,\,limit}(100\%\ \mathrm{FW}) \\ - 2.5 * \log(\mathrm{X\%\ FW}). \quad (11)$$

In order to provide a margin for linearity we are using 80% full-well in our baseline run with NIRSpec.

Similarly, the brightness limits for NIRISS SOSS are shown in Table 1 of the STScI's on-line User Documentation for
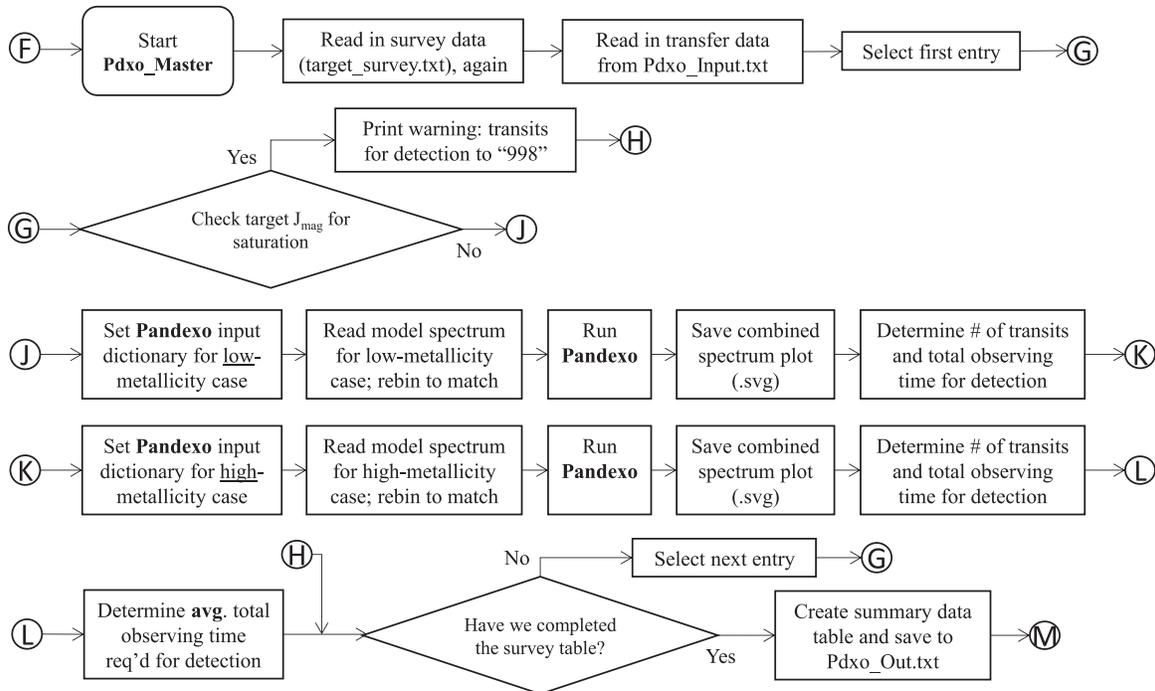

Figure 5. Analysis and code flow for `JET` program element `Pdxo_Master`.

NIRISS Bright Limits.[14] Currently `JET` can simulate the performance of the SUBSTRIP96 subarray, with Order 1. The full-well J-magnitude limit is listed as 7.5 for a G2V star. This bright limit varies by $+0.15$ mag for an A0V star to -0.10 mag for an M5V star. The `JET` user manually inputs the three data points (for $T_{eff} = 2500$ K, $J_{mag} \sim 7.4$; for 5000 K, $J_{mag} \sim 7.5$ and for 10,000 K, $J_{mag} \sim 7.65$) for the full-well limits. The same correction formula described in Equation (11) above is used to estimate the J-magnitude limit for a particular percent-full-well condition for NIRISS SOSS. Again, we are using an 80% full-well cap on the brightness limit in our variation studies with NIRISS.

`PandExo` is first run for a single transit. This yields a baseline simulated instrument spectrum with 1σ error bars for each data point (at n wavelength locations, depending on the instrument and spectral resolution chosen). Figure 6 presents simulated transmission spectra for Sullivan Target 1292 using NIRSpec G395M, with a low-metallicity atmosphere ($5\times$ Solar, no clouds), and a high metallicity atmosphere ($1000\times$ Solar, clouds at 100 mbar). The NIRSpec G395M disperser covers the wavelength range from 2.87 to 5.18 μm, with a noise floor estimated to be 25 ppm.

The noise floor assumption is an important one. Unfortunately, in the literature it is not always clear if a given noise floor value is associated with a single transit observation, or is a

residual value after multiple transits. We use the term in the latter sense. This is consistent with our detection algorithm and the way that `PandExo` is structured. In `PandExo` the propagated error generally falls with the square root of the observed transit time and number of transits. When the error reaches the value of the noise floor it will go no lower for additional observation time or number of transits.

First, we will address the NIRISS noise floor assumption. For the variation studies that we present in Section 4.2, we assume a noise floor of 20 ppm for NIRISS SOSS.

Beichman et al. (2014) presented "a transmission spectrum simulation of what we would expect with NIRISS for GJ 1214b, a super-Earth around a star of magnitude $J = 9.75$. The simulation assumes 12 hr of clock time spread over 4 transits, and a 20 ppm noise floor." Here the noise floor is given in a multi-transit residual sense.

Greene et al. (2016) suggested 20 ppm for the NIRISS SOSS noise floor. They pointed out that, "The best *HST* WFC3 G141 observations of transiting systems to date have noise of the order of 30 ppm (Kreidberg et al. 2014), ... We adopt reasonably optimistic systematic noise floor values of 20 ppm for NIRISS SOSS, ... These are less than or equal to the values estimated by Deming et al. (2009) for the *JWST* NIRSpec . . . . The excellent spatial sampling of the NIRISS GR700XD SOSS grism approaches that of the *HST* WFC3 G141 spatial scanning mode, and both instruments have reasonably similar HgCdTe detectors. We anticipate that decorrelation techniques will continue to

---

Table 3
Observation Time Elements

Time Element	NIRSpec BOTS (s)	NIRISS SOSS (s)
Science observing time:		
Transit observation (e.g., t_{dur}[a] = 2.4 hr)	8640	8640
Out of transit observation,[b] including "timing tax" [t_{dur} + 3600] (s)	12240	12240
Slew time[c,d] [(1800 + avg. slew time * 4)/5] (s):	1108	1108
Instrument Overheads:[e]		
SAMs: small angle maneuvers	0	0
GS Acq: guide star acquisition(s)	284	284
Targ Acq: target acquisition, if any	492	602
Exposure Ovhd: some instruments require initial reset		
NIRSpec: [0.0393 * (Science) + 26]	846	
NIRISS: [0.2523 * (Science) + 15]		5283
Mech: mechanism movements, including filter wheels	110	52
OSS: Onboard Script System compilation	65	30
MSA: NIRSpec MSA configuration	0	0
IRS2: NIRSpec IRS2 Detector Readout Mode setup	0	0
Visit Ovhd: visit cleanup activities	110	62
Observatory Overheads [16% * (Science + Slew + Instr. Ovhd)]:[f]	3823	4528
Direct Scheduling Overheads:[g]	0	0
Total charged time [Science + Slew + Instr. Ovhd + Observ. Ovhd + DS Ovhd]:	27718	32829

Notes.
[a] t_{dur} is the transit duration.
[b] To avoid systematic bias, the "Out of transit" observation is split into two parts, pre-ingress, and post-egress.
[c] Avg. slew time per target, based on program of one 53 deg. slew (of 1800 s) followed by four 13 deg. slews (each 935 s) repeating 95 times over course of full 10 yr fuel limited mission. This assumes a program of approx. 8300 hr of exoplanet transmission spectroscopy with one dedicated instrument over the full mission. If we assume roughly 7.7 hr for each transit observed, and ~2.3 transits for each target we could visit roughly 475 targets/transits over the course of the full mission. Raw slew duration versus distance (angle) are given in the *JWST* User Documentation: (https://jwst-docs.stsci.edu/jppom/visit-overheads-timing-model/slew-times).
[d] Avg. slew angle comes from the fact that there are 41,253 sq. deg. on the sky, and if our 475 targets/transits are evenly distributed, then each target is in a box of 87 sq deg (or about 9 deg. on a side). The diagonal of this box and the avg. angular separation of the targets is ~13 deg.
[e] Instrument overhead times are based on a set of trial observing programs using the APT.
[f] This charged time supports calibration, station keeping, and momentum management activities.
[g] For very tight timing constraints or rapid turnaround targets of opportunity.

improve, so we assign a 20 ppm noise floor value to NIRISS even though *HST* has not yet (as of early 2016) done quite this well."

More recently, *HST* WFC3 has demonstrated even better performance. Line et al. (2016) reported a precision for *HST* WFC3 (at 1.1–1.6 μm) of ~17 ppm for multiple secondary eclipse observations of HD209458b. In spite of this we choose to remain conservative with a noise floor estimate of 20 ppm for NIRISS SOSS. There are still many unknowns for *JWST*.

Next, we consider our noise floor assumption for NIRSpec. Of course the actual performance of NIRSpec will not be known until *JWST* is on orbit and fully commissioned, but there are some known issues that could make it difficult for NIRSpec to improve upon the *HST* WFC3's very low noise qualities. Specifically the NIRSpec image sampling is not quite as good as *HST* WFC3, the detector electronics (in particular the System for Image Digitalization, Enhancement, Control, and Retrieval (SIDECAR) ASICs, used for detector readout and control)[15] are not well characterized at this point, and the measured wavefront-error is acceptable, but not excellent.

NIRSpec may have more systematic noise than NIRISS because it is not spatially sampled well at wavelengths less than about 3 μm. The PSF of *JWST* is on the order of ~70 mas FWHM at 2 μm,[16] but the NIRSpec detectors have 100 mas pixels.[17] This under-sampling condition will tend to degrade resolution and add systematic noise that can generally not be completely eliminated by co-adding transits.

"Since the NIRSpec PSF is under-sampled at most wavelengths, dithering is required to achieve nominal spectral and spatial resolution."[18]

This under-sampling issue has come up with other space telescope instruments. Discussing *Spitzer*/IRAC performance, Ingalls et al. (2012) commented that "due to the under-sampled nature of the PSF, the warm IRAC arrays show variations of as much as 8% in sensitivity as the center of the PSF moves across a pixel due to normal spacecraft pointing wobble and drift. These intra-pixel gain variations are the largest source of correlated noise in IRAC photometry."

In addition, Grillmair et al. (2012) identified a number of *Spitzer*/IRAC systematic issues "that limit the (telescope's) attainable precision, particularly for long duration observations (e.g., transits). These include initial pointing inaccuracies, pointing wobble, initial target drift, long-term pointing drifts, and low and high frequency jitter. Coupled with small scale, intra-pixel sensitivity variations, all of these pointing issues have the potential to produce significant, correlated

---

[15] https://www.cosmos.esa.int/web/jwst-nirspec/nirspec-s-design
[16] https://jwst.nasa.gov/content/forScientists/faqScientists.html#howbig
[17] https://jwst-docs.stsci.edu/near-infrared-spectrograph#NearInfrared Spectrograph-Opticalelementsanddetectors
[18] https://jwst-docs.stsci.edu/near-infrared-spectrograph/nirspec-observing-modes/nirspec-fixed-slits-spectroscopy
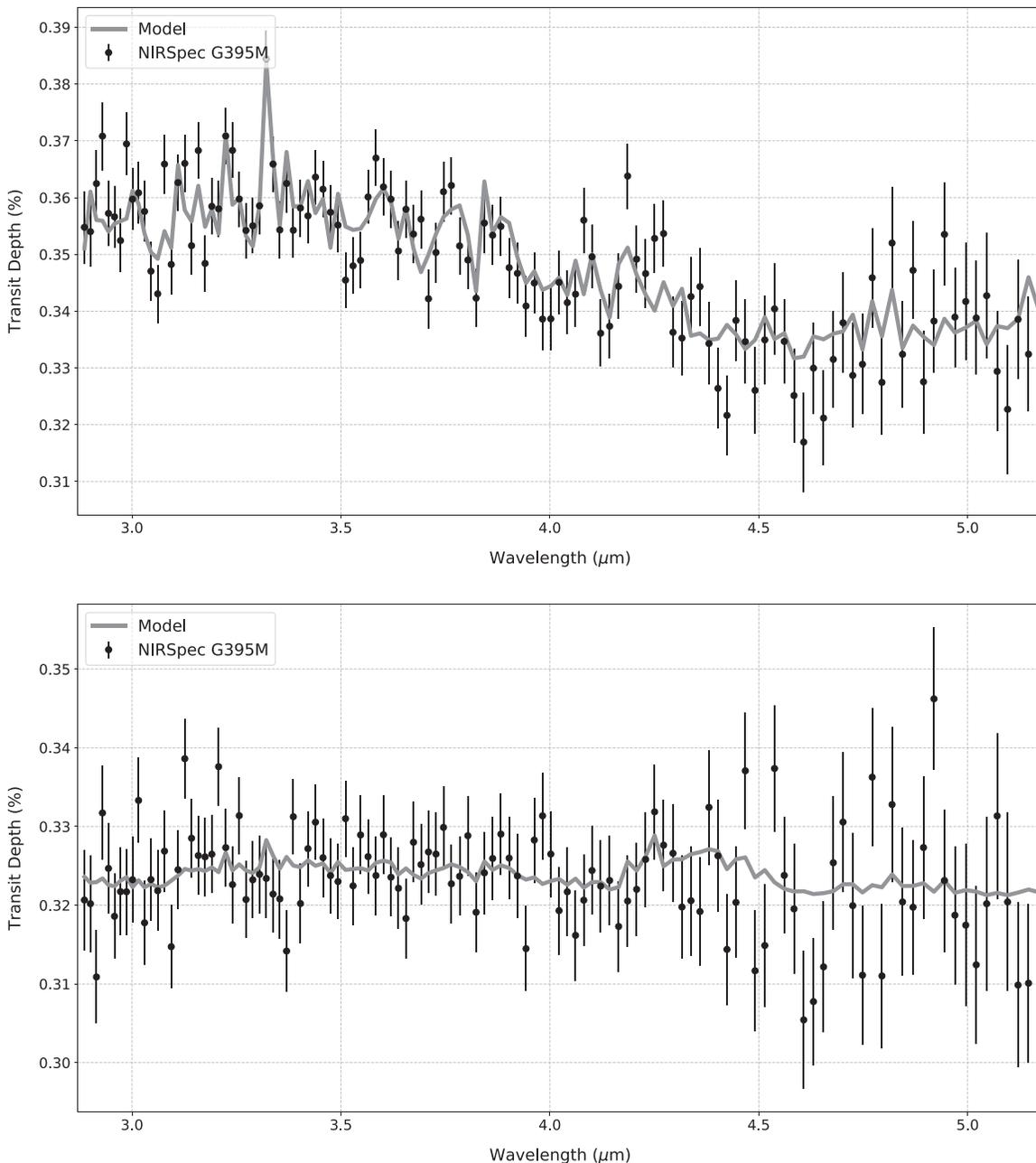
Figure 6. Simulated spectra for a single transit of Sullivan Target 1292 using NIRSpec G395M. The model spectra shown as gray background lines have been binned down to a resolution consistent with the simulated data ($R \sim 100$). Top: a low-metallicity ($5\times$ Solar, no clouds) atmosphere. For the lo-metal case the mean dBIC (less $1\,\sigma$) is >10 for a single transit. Bottom: a high-metallicity ($1000\times$ Solar, clouds at 100 mbar) atmosphere. The mean dBIC (less $1\,\sigma$) for a single transit of the Sullivan 1292 hi-metal case is ~ -18. It takes 17 transits to reach the dBIC detection threshold of 10.

photometric noise." *JWST* will be affected by these issues to a degree, to be determined.

The NIRSpec wavefront-error performance is acceptable, but it is not equal to the excellent performance of *HST* WFC3. Aronstein et al. (2016) noted that ". . . there are currently violations in the requirements for the uncertainties of 3rd order aberrations in NIRSpec."

Louie et al. (2018) argue that systematic noise can be removed by co-adding multiple transits. In our view this may be true for some systematic noise sources, but not for all (e.g., thermal shock after slew).

Greene et al. (2016) also point out that, "Astrophysical noise (e.g., Barstow et al. 2015) and/or instrumental noise (e.g., decorrelation residuals) produce systematic noise floors that are

not lowered when summing more data. . . . We do not know how well co-adding observations of multiple transits or secondary eclipses will improve our results at this time. The simulated single-transit and single-eclipse observations of our selected systems typically have total noise values only 10%–50% larger than our adopted noise floors . . . so systematic noise assumptions have already significantly influenced the precision of our simulated data . . . . Given this, co-adding more data would not substantially improve the results for these very observationally favored systems with bright host stars (generally the case with *TESS* targets)."

It should be mentioned that the noise floor generally has more of an impact on the more difficult detection situations, where co-adding many transits is needed for a detection. These targets will tend to be further down the ranking list and may well fall below a time allocation cut-off line. Reducing the noise floor would likely only have a marginal effect on the cut-off-limited ranking.

While it could be argued that our noise floor values are somewhat conservative, it would seem that for the purpose of establishing preliminary target rankings a conservative approach is appropriate. In the end, the noise floor for any particular instrument configuration is an input value and can be changed at the user's discretion.

For the purpose of our full catalog baseline run with JET, we have chosen the NIRSpec instrument with the G395M disperser and the F290LP filter operating in BOTS mode. There were a number of factors that influenced this decision. First, according to Batalha & Line (2017), "An observation with both NIRISS [SOSS] and NIRSpec G395M/H always yields the highest information content spectra with the tightest constraints, regardless of temperature, C/O, [M/H], cloud effects or precision." In addition, Bean et al. (2018) included the NIRSpec G395H as one of the consensus high priority modes for the ERS Program for *JWST*. Also, PandExo run-times for NIRSpec are significantly shorter than for NIRISS. We have made small variation-set runs with NIRISS, but as a practical consideration, for the full catalog baseline run we felt that NIRSpec would be a better choice. Finally, the NIRSpec G395M wavelength coverage gives us (simulated) access to the 4.5 μm region and effects of the important atmospheric constituents CO and CO_2, which would not be true for NIRISS.

The NIRSpec simulations are performed using the S1600A1 aperture with a fixed $1\!\!''6 \times 1\!\!''6$ field of view.[19] Exposures use the SUB2048 subarray (2048×32 pixels)[20] to record the full spectrum using the NRSRAPID read mode (Ferruit et al. 2014).

Given the long PandExo run times for our baseline instrument mode (NIRSpec G395M), and given the purpose

of our study (a simple detection, not a full constituent retrieval) we felt that the M disperser mode was an appropriate choice over the H mode which would have even longer run times.

It should be kept in mind that the instrument/mode chosen for the baseline run was primarily for the purpose of demonstrating the JET code. The user can study other modes (e.g., NIRSpec G140M/BOTS, and G235M/BOTS, as well as NIRISS SOSS, with the GR700XD disperser) that have been implemented and tested in JET.

In the figures, we show the model spectrum in the background, and have re-binned the data to $R \sim 100$ to match the resolution set in the PandExo simulation. The lower (than native instrument) resolution is used to reduce overall computation time. In the description of the analysis that follows, the re-binned model spectral values are designated y_{rebin}. The re-binning process is accomplished using the Python package SpectRes as described in Carnall (2017).

2.3.2. Detecting the Transmission Spectra

For a particular target/atmosphere, the questions that we need to answer are: (1) Can the spectrum be detected given the instrument's noise characteristics?, and (2) If the observation of one transit is insufficient to detect the spectrum, will the observation of additional transits pull a detectable signal out of the noise?

Specifically we want to determine, in a formal way, whether our simulated spectrum is a better fit to the model spectrum or to a flat-line spectrum. We use a Bayesian Information Criterion (BIC) approach to guide this selection (Kass & Raftery 1995; Wall & Jenkins 2012). This technique will allow us to determine if the spectrum is detectable at all, and if so, how many transit observations will be needed for a strong detection.

Fortunately, we do not need to re-run the full PandExo simulation over and over to determine the results of observations of multiple transits. We can compute the improvements in signal to noise, and the effects of the instrument noise floor analytically.

Our one-transit PandExo run, described in the previous section, generates a simulated spectrum with n spectral data elements of wavelengths (x), spectral values (y), and error values (e). In addition, we can extract another set of data from PandExo that does not include random noise. We will call these the wavelengths (x_{1trans}), the spectral values (y_{1trans}), and the noise values (e_{1trans}).

We then compute the integrated multi-transit noise (noise$_{\mathrm{multi\text{-}transit}}$):

$$\mathrm{noise}_{\mathrm{multi\text{-}transit}} = \frac{e_{\mathrm{1trans}}}{\sqrt{ntr}} \qquad (12)$$

where *ntr* is the number of transit observations co-added.

If this value is less than the noise floor, then we reset it to the noise floor lower bound. Clearly, this multi-transit noise term

---

will tend to smaller and smaller values as the number of transits increases (bounded by the noise floor).

Next, we compute a random noise value:

$$\text{noise}_{\text{random}} = \text{noise}_{\text{multi-transit}} * f_{\text{random}} \qquad (13)$$

where for each statistical sample, the f_{random} term is drawn from a standard normal distribution with mean 0 and variance 1.

Now we can recast the spectrum with random noise:

$$x_1 = x_{1\text{trans}} \qquad (14)$$

$$\text{spectrum}_{\text{sim}} = y_{1\text{trans}} + \text{noise}_{\text{random}}. \qquad (15)$$

Adjusting for transit depth % units:

$$y_{100} = \text{spectrum}_{\text{sim}} * 100 \qquad (16)$$

$$e_{100} = \text{noise}_{\text{multi-transit}} * 100. \qquad (17)$$

Next, we compute the BIC value for the model spectrum case. We first determine the chi-squared statistic (χ^2_{MS}), and the reduced chi-squared statistic ($\chi^2_{\text{MS,r}}$), for the single transit spectral data:

$$\chi^2_{\text{MS}} = \sum \left(\frac{y_{100} - y_{\text{rebin}}}{e_{100}} \right)^2 \qquad (18)$$

$$\chi^2_{\text{MS, r}} = \frac{\chi^2_{\text{MS}}}{(n - n_{\text{params}})}. \qquad (19)$$

For the model spectrum case the number of free parameters (n_{params}) used in the calculation of the reduced chi-squared ($\chi^2_{\text{MS,r}}$) is five.

As discussed in Section 2.2.6, JET uses the transmission spectroscopy code Exo-Transmit to generate model spectra of the target planet atmospheres. In the JET implementation, Exo-Transmit reads in survey/catalog data for each target planet and generates a moderate resolution transmission spectrum of the planet's atmosphere for two atmosphere cases: (1) low metallicity solar with no clouds, and (2) high metallicity solar with low altitude clouds.

The JET code then generates a simulated observational spectrum. Our observational simulator, PandExo (described in Section 2.3.1), uses the primary model as the starting point for the simulation. It introduces noise based on the target characteristics, and the performance of the JWST instrument under study.

The goal of our target ranking exercise is to determine the targets for which JWST observations could potentially be the most informative. Given that we do not know the atmospheric composition a priori, simulating transmission spectra for two drastically different atmospheres allows us to investigate which targets would be the most interesting for future atmospheric characterization via retrieval studies.

The question is: for our calculation of the reduced chi-squared, and the model spectrum BIC, how many free parameters should we employ?

A typical JWST retrieval, assuming high S/N and high spectral resolution, may employ ten or even more free parameters (Madhusudhan 2018). For our initial ranking and selection of targets, we are more interested in assessing the overall level of information content within each spectrum rather than determining detailed atmospheric abundances for individual elements. We therefore adopt a relatively low spectral resolution ($R = 100$) and consider a broad set of potential targets, some of which may be of somewhat lower S/N. For the model spectrum BIC analysis, we therefore consider a lower number of free parameters to correspond with the lower resolution and lower S/N of our simulated spectra. (Benneke 2013).

We have chosen five representative free parameters for our model spectrum BIC calculation for each target/atmosphere case. The parameters include: (1) planet equilibrium temperature, (2) bulk metallicity/atmospheric composition ($1\times$ solar, $10\times$ solar, etc.), (3) Rayleigh scattering (on/off), (4) cloud height, and (5) whether to include rainout of molecules out of the gas phase (i.e., condensation/gas models). Of course other "fixed" catalog values (e.g., R_p, R_*, g_s) are needed to generate the model spectra for each target, but these are not included as "free" parameters because the planets and their host stars will have known radii, masses, and surface gravities.

To provide flexibility, the number of free parameters for the model spectrum BIC calculation is an input value and can be changed by the user.

Next, we compute an error factor to drive $\chi^2_{\text{MS,r}}$ to 1. JET runs PandExo for a single transit and adds a random noise value with raw error bars (based on the single transit S/N level) to the underlying model spectrum. Since the model spectrum is the correct model for the simulation by construction, the reduced χ^2 should indicate a good fit. In principle, a value of $\chi^2_{\text{MS,r}} = 1$ indicates that the extent of the match between the simulated observation and the model is in accord with the error variance. The error factor (f_{error}) is used to adjust the simulation error bars to result in $\chi^2_{\text{MS,r}} = 1$. These adjusted error bars are then used in the χ^2 (and BIC) calculations for the model spectrum and the associated flat-line spectrum.

The error factor based on the raw spectrum data is:

$$f_{\text{error}} = \sqrt{\chi^2_{\text{MS,r}}}. \qquad (20)$$

From this we can determine the rescaled χ^2 for the model atmosphere case:

$$\sigma_{\text{combined}} = \sqrt{(f_{\text{error}} * e_{100})^2} \qquad (21)$$

and,

$$\chi^2_{\text{MS}} = \sum \left(\frac{y_{100} - y_{\text{rebin}}}{\sigma_{\text{combined}}} \right)^2 \qquad (22)$$

and now,

$$\chi^2_{\text{MS,r}} = \frac{\chi^2_{\text{MS}}}{(n - n_{\text{params}})} = 1. \tag{23}$$

Finally, we can compute the BIC value for the model spectrum case (BIC_{MS}). The BIC is formally defined, using the traditional nomenclature, as:

$$\text{BIC} = \ln(n)k - 2\ln(\hat{L}) \tag{24}$$

where $\hat{L}$ is the maximized value of the likelihood function of the model M. $\hat{L} = p(x|\hat{\theta}, M)$, and $\hat{\theta}$ are the parameter values that maximize the likelihood function. x is the observed data, in our case the spectral transit depth values; n is the number of data points in x (in our case the wavelength points), essentially the sample size; and k is the number of parameters estimated by the model (in our case the number of "free" parameters, n_{params}).

As discussed in Kass & Raftery (1995), under the assumption that the errors are independent and follow a normal distribution, the BIC can be rewritten in terms of the χ^2 as:

$$\text{BIC} = \chi^2 + k\ln(n). \tag{25}$$

For the model spectrum case we have:

$$\text{BIC}_{\text{MS}} = \chi^2_{\text{MS}} + (n_{\text{params}}) * \ln(n) \tag{26}$$

where we have set $n_{\text{params}} = 5$.

Next, we repeat the BIC calculation for the case of the flat-line spectrum. In this case only one number, the y-intercept of the flat line, needs to be specified. This number can be determined by taking the median value of the simulated spectrum values. This is consistent with the conventional approach where the model parameters are determined from the observational data.

We can now determine the BIC value for the flat-line spectrum case (BIC_{FL}):

$$\text{BIC}_{\text{FL}} = \chi^2_{\text{FL}} + (1) * \ln(n). \tag{27}$$

For model selection we are interested in the difference in BIC. When picking from multiple models, the one with the lower BIC is preferred.

We can define the parameter:

$$\text{dBIC} = \Delta\text{BIC} = \text{BIC}_{\text{FL}} - \text{BIC}_{\text{MS}}. \tag{28}$$

For a very strong detection we need dBIC > 10 (Kass & Raftery 1995). This loosely corresponds to a 3.6σ confidence level of detection (Gordon & Trotta 2007) for the mean value of dBIC.

Now we can loop on this analysis sequence, holding the number of transits constant and just considering new random samples of noise. With enough samples (we use 2000 samples for each number-of-transits grid point) we build up a distribution of dBIC. The distribution tends toward Gaussian from which we can determine the mean and standard deviation.

We then move to the next higher number of transits on the grid and build up another distribution of dBIC, repeating the process outlined in Equations (12) through (28). We use 15 transit grid points, from 1 transit to 50 transits (with spacing increasing as the number of transits increase). The result is a data set of mean dBIC (and standard deviation) versus number of transits.

For a particular number-of-transits grid point, if we use the mean value of dBIC as our critical parameter, then we can say that we have a 50% confidence level of a very strong detection if dBIC > 10. Assuming a Gaussian distribution, if we reduce our $\text{dBIC}_{\text{mean}}$ parameter by 1σ, then for this $\text{dBIC}_{\text{mean}}$ less 1σ we have an ~84% confidence level of a very strong detection.

For each target/atmosphere combination we first check the $\text{dBIC}_{\text{mean}}$ less 1σ for one transit. If it is >10 we record that detection and move on. If not, then we check to see if the $\text{dBIC}_{\text{mean}}$ less 1σ for 50 transits is >10. If not, then we flag this target/atmosphere as a non-detection and move on.

If neither of these conditions are met, then we need to determine the number of transits to reach the detection threshold. We use the `Python` routine `UnivariateSpline` from the `scipy` package to interpolate the 15 element $\text{dBIC}_{\text{mean}}$ less 1σ data set. The smoothing parameter (s) is dynamic. We set it to 2 (soft spline) if the $\text{dBIC}_{\text{mean}}$ at 50 transits is less than 175; otherwise we set it to zero (spline goes through all points). Then we use the `scipy` routine `brentq` to find the crossing point of number of transits for a $\text{dBIC}_{\text{mean}}$ less 1σ of 10.

Figure 7 presents a plot of dBIC versus number of transits for Sullivan Target 1292 with a high metallicity atmosphere.

2.3.3. Determining the Observing Cycle Time Needed for Detection

The critical resource is total observing time. For a single transit observation there are a number of actions that take time. These include: (1) slewing the spacecraft to the proper target; (2) acquiring the target; (3) allowing the detectors to settle; (4) observing the transit; (5) observing the host star out-of-transit to establish the baseline flux level; and (6) various overheads added based on experience with other space telescopes.[21] We have captured these activities and overhead factors for NIRSpec-BOTS mode and NIRISS-SOSS for an example case with a transit duration of 2.4 hr (the median transit time of the Sullivan survey) in Table 3.

The Instrument overheads shown in the table have been generated using the *JWST* Astronomer's Proposal Tool (APT).[22] The Slew time and Instrument Overheads shown do not change for the particular instrument configuration considered over a reasonable science-time range.

---

[21] https://jwst-docs.stsci.edu/jwst-observatory-functionality/jwst-observing-overheads-and-time-accounting-overview
[22] https://jwst-docs.stsci.edu/jwst-astronomers-proposal-tool-overview
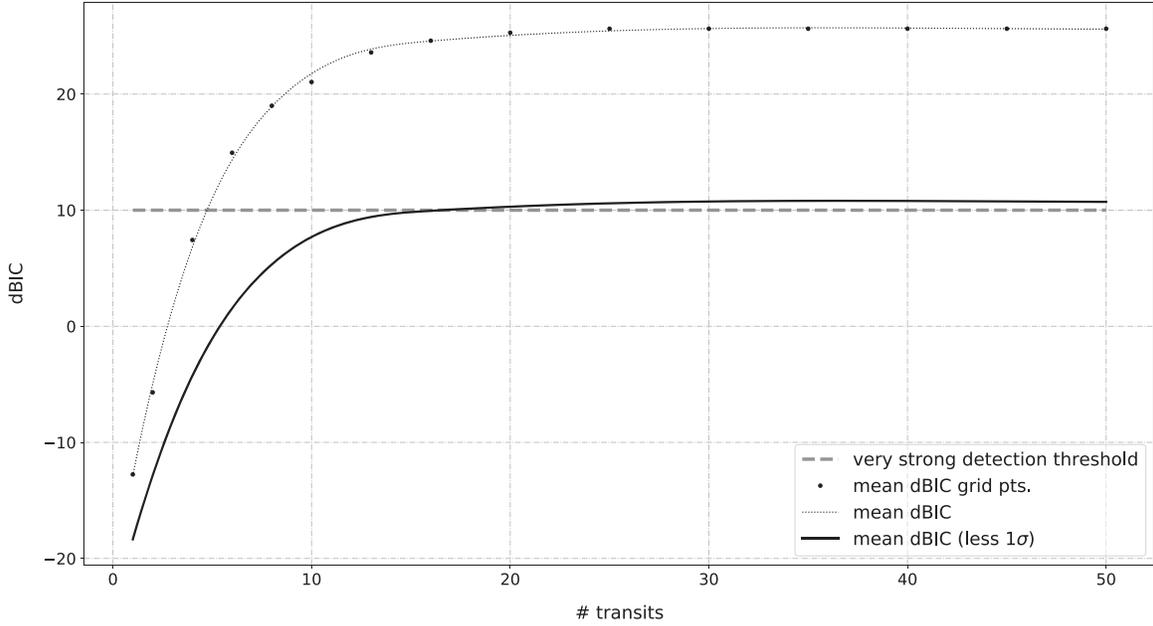
Figure 7. dBIC vs. number of transits for Sullivan Target 1292 with a high metallicity (1000× Solar, clouds at 100 mbar) atmosphere. The mean dBIC grid points are shown and the univariate spline fit with soft smoothing is shown as well for this mean value. The mean less 1σ value is also shown and is the critical line for determining detection with very high confidence. The plot shows a very strong detection for 17 transits.
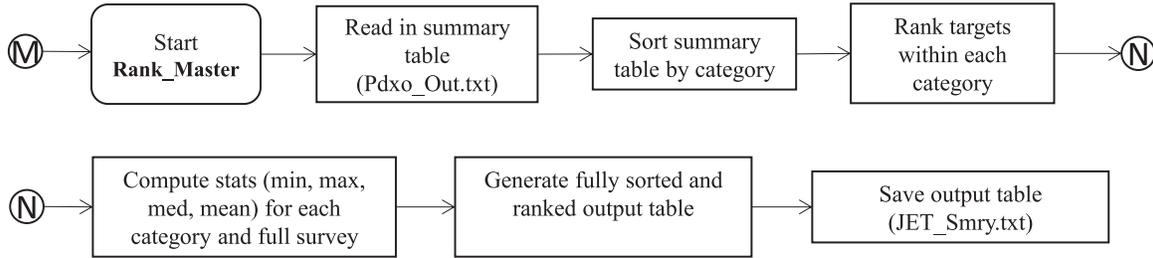

Figure 8. Analysis and code flow for `JET` program element `Rank_Master`.

So, for a given number of transits needed for detection, we can determine the total observation time needed:

$$t_{tot} = ntr * (\text{Total charged time}) \qquad (29)$$

where t_{tot} is the total observation time in hours, ntr is the number of transits required for detection of the transmission spectrum.

We determine the total observation time for both atmosphere cases of each target. Then, given the broad range of plausible atmospheric compositions we simply take an average of the two to determine our figure of merit (tT_{avg}):

$$tT_{avg} = \frac{(t_{tot,lo} + t_{tot,hi})}{2} \qquad (30)$$

where tT_{avg} is the average total observation time in hours.

Finally, summary data for each target is gathered by `Pdxo_Master`. Upon completion of the analysis of the target list, this summary data is written to the file: `Pdxo_Out.txt` in the working directory. The data here is in unranked target order.

2.4. Sorting and Ranking Targets (Rank_Master)

Once we have completed the analysis of all of the targets and computed the average total observation time required for detection of the atmosphere it is time to sort and rank the targets and provide summary statistics. Using a two level sorting approach, we first sort the targets by category, then rank the targets within each category on the average total observing time parameter (tT_{avg}). Targets that have hit warning flags (e.g., detector saturation, $R_p > 10R_\oplus$, etc.) are listed at the end of the viable ranking for each category. The sorting and ranking process is carried out by the program element `Rank_Master`, shown in flowchart form in Figure 8.

Table 4
Basic Survey Data for Validation

Target	R_p ($R_\oplus$)	P (days)	S ($S_\oplus$)	R_* ($R_\odot$)	T_{eff} (K)	J_{mag} (mag)
1	3.31	9.1	361.7	1.41	6531	7.6
2	2.19	14.2	2.1	0.32	3426	11.6
3	1.74	5.0	235.0	0.95	5546	8.8
4	1.48	2.2	1240.0	1.13	5984	7.0
Earth	1.00	365.2	1.0	1.00	5777	⋯
Mercury	0.38	87.6	6.7	1.00	5777	⋯
Mars	0.53	686.2	0.4	1.00	5777	⋯
GJ 1214b	2.67	1.6	17.6	0.22	3026	9.8

Note. Ref Data Sources: Targets 1 through 4, Sullivan catalog; solar system planets, NASA Planetary Fact Sheet (https://nssdc.gsfc.nasa.gov/planetary/factsheet/planet_table_ratio.html); GJ 1214b, NASA Exoplanet Archive (https://exoplanetarchive.ipac.caltech.edu/); Sun, Carroll & Ostlie (2006).

3. Validating the JET Code

In order to validate that the JET code is performing its calculations correctly, we have considered the various parameters that are determined by JET and compared those to manual calculations using the methods described in Section 2, and to reference data if available.

For the model spectra and simulated instrument spectra generated by Exo-Transmit and PandExo respectively, we demonstrate that our implementation of these programs is working correctly within the JET framework.

3.1. Planet Masses

We first consider the planet mass estimates. Our method for estimating planet mass based on planet radius is described in Section 2.2.2 above. Table 4 presents basic data for the first four targets of the Sullivan catalog, for several solar system planets, and for the well known exoplanet GJ 1214b. The planet mass (M_p) estimated by JET is compared to a manual calculation based on the previously described methods. The JET-to-Manual residuals for mass are well under 1% for the example cases considered (see Table 5).

For the solar system planets and for GJ 1214b, we compared the JET mass estimates to reference masses. For the solar system planets the JET mass estimates are within 7% of reference, except for Mercury, which with its very small radius, has a mass that is not well approximated by the methods of Chen & Kipping (2017). Mercury has a radius of roughly 0.38 $R_\oplus$, whereas the smallest planet radius in the Sullivan catalog is roughly 0.70 $R_\oplus$. The mass estimate for GJ 1214b is off by about 16%. This is still within the acceptable bounds of uncertainty of our estimation approach.

3.2. System Parameters, and Transits Observable

In a similar manner, we validate the JET calculations for the other planetary system parameters (a, T_{eq}, g_s, t_{dur}, and $ntr_{10\,yr}$).

Table 5
Validation of JET Calculation of Planetary Parameters

Target	a ($R_\odot$)	T_{eq} (K)	g_s (m s^{-2})	M_p ($M_\oplus$)	t_{dur} (hr)	$nt_{10\,yr}$
JET computed values:						
1	20.4	1215	9.8	11.0	4.9	163
2	17.0	333	11.1	5.4	2.2	137
3	12.3	1091	11.9	3.7	3.0	280
4	7.4	1653	12.5	2.8	2.5	514
Earth	215.1	279	9.5	1.0	13.1	⋯
Mercury	83.3	448	2.1	0.0	8.1	⋯
Mars	328.0	226	3.5	0.1	16.1	⋯
GJ 1214b	3.0	570	10.5	7.6	1.0	⋯
Manual calculation (by methods of Section 2):						
1	20.5	1213	9.8	11.0	4.9	163
2	17.0	333	11.1	5.4	2.2	136
3	12.4	1089	11.9	3.7	3.0	282
4	7.4	1651	12.5	2.8	2.6	510
Earth	215.5	278	9.5	1.0	13.1	⋯
Mercury	83.4	447	2.1	0.0	8.1	⋯
Mars	328.7	225	3.5	0.1	16.1	⋯
GJ 1214b	3.0	570	10.5	7.6	1.0	⋯
JET to Manual residual (%):						
1	0.2	−0.1	−0.2	0.0	0.3	0.0
2	0.2	−0.1	−0.2	0.0	0.3	−0.7
3	0.2	−0.1	−0.2	0.0	0.3	0.7
4	0.2	−0.1	−0.2	0.0	0.3	−0.7
Earth	0.2	−0.1	−0.2	0.0	0.3	⋯
Mercury	0.2	−0.1	−0.2	0.0	0.3	⋯
Mars	0.2	−0.1	−0.2	0.0	0.3	⋯
GJ 1214b	0.2	−0.1	0.1	0.3	0.3	⋯
Reference parameters:						
Earth	215.5	279	9.8	1.0	13.0	⋯
Mercury	83.4	449	3.7	0.1	8.1	⋯
Mars	327.6	218	3.7	0.1	16.0	⋯
GJ 1214b	3.0	604	8.8	6.6	0.9	⋯
JET to Reference residual (%):						
Earth	−0.2	−0.2	−2.9	−2.9	0.6	⋯
Mercury	−0.1	−0.3	−44.4	−44.9	−0.5	⋯
Mars	0.1	3.5	−6.1	−6.7	0.4	⋯
GJ 1214b	0.3	−5.5	18.8	16.1	9.9	⋯

Note. Ref Data Sources: Targets 1 through 4, Sullivan catalog; solar system planets, NASA Planetary Fact Sheet (https://nssdc.gsfc.nasa.gov/planetary/factsheet/planet_table_ratio.html); NASA About Transits (https://www.nasa.gov/kepler/overview/abouttransits); GJ 1214b, NASA Exoplanet Archive (https://exoplanetarchive.ipac.caltech.edu/); Sun, (Carroll & Ostlie 2006).

Again, Table 5 presents the results of this validation study. For all of the targets the JET-to-Manual calculation residuals are well under 1%. The very small errors are due to minor roundoff and constant differences between the manual calculations and those carried out by JET.

For the JET-to-Reference comparison we see some bigger differences. The simple JET estimate of T_{eq} is slightly off for Mars and GJ 1214b, probably due to our assumption of zero

albedo and other small differences. The reference T_{eq} for Earth is shown for zero albedo. Finally, the JET-to-Reference residual for the GJ 1214b transit duration is off by roughly 10%. This could be due to many factors, including differences in the assumptions for eccentricity (value <0.14), inclination (values between 87.63 and 90 deg.), host star radius (values between 0.204 and 0.228 $R_{\odot}$), or planet radius (values between 2.19 and 3.05$R_{\oplus}$).[23]

3.3. Model Spectra

To validate the calculation of model transmission spectra we have taken a benchmark spectrum generated by Exo-Transmit for the exoplanet GJ 1214b (obtained by private communication with the Exo-Transmit code's lead author, E. Kempton, and described in Section 2.2 of Kempton et al. 2017) and compared it with a model spectrum produced by the JET ExoT_Master subprogram. The benchmark Exo-Transmit spectral data is based on planetary system parameters from the exoplanet.eu database.[24] For the comparison JET spectrum we have used exoplanet.eu for the basic catalog-like input data and relied on the JET computed values for the other system parameters.

The two spectra shown in the upper panel of Figure 9 are a very close match. The small differences seen are likely due to differences in the re-binning and smoothing algorithms used, and to minor differences in input parameters for the planet-host star system. The center panel shows the JET-to-Exo-Transmit residuals at the same scale. The lower panel shows the distribution of the JET-to-Exo-Transmit residuals. The mean of the residuals is approximately -4 ppm, with a standard deviation of approximately 55 ppm.

As shown in Figure 9, JET has properly implemented the underlying Exo-Transmit code and is producing consistent transmission spectra.

3.4. Simulated Spectra

Our approach to validation of our calculation of simulated instrument spectra is essentially the same as that of the previous section. We have taken a benchmark spectrum generated by PandExo for the exoplanet GJ 1214b (obtained by private communication with the PandExo code's lead author, Natasha Batalha, and generally described in Batalha et al. 2017) and compared it with a simulated spectrum produced by the JET Pdxo_Master subprogram.

The two spectra (in this case the data points representing the instrument spectra for a single transit with random noise) shown in the upper panel of Figure 10 are a close match. The small differences seen are likely due to the fact that each case is a different random noise instance and would only agree in

overall distribution. The center panel shows the JET-to-model residuals and the PandExo-to-model residuals. The lower panel shows the distributions of the two residual plots (JET-to-model, and PandExo-to-model) overlaid. They are a close match. The difference in the mean of the residuals is approximately 20 ppm, with the difference in standard deviation of the residuals approximately 5 ppm.

Figure 10 demonstrates that JET has properly implemented the underlying PandExo code and is producing consistent simulated instrument spectra.

4. Results/Discussion

Now that we have described our analysis framework and validated the calculations in the JET code, we present the results of our full baseline run. In addition, we present limited case studies of the effects of variations of some of the driving parameters of our analysis, including the noise floor, the atmospheric EOS, the detection threshold, and the instrument/mode selected.

4.1. Baseline (Full Survey)

The input parameters for the baseline run of the 1984 target Sullivan catalog were shown previously in Table 2. Excerpts of the summary output table for the baseline run are included in Table 6. A machine-readable version of the full output table is available, however it does not include the summary statistic sections.

We now have a list of (simulated TESS catalog) targets that is categorized and ranked within each category on the average total observation time to detect an atmosphere (tT_{avg}).

The overall statistics for the baseline run are presented in Table 7. Out of the total 1984 targets in the catalog, we show a full unambiguous detection (a detection with less than 50 transits observed for both low and high metallicity atmosphere cases) of 1070 targets. Some of these detections may well be unrealistic, given the large number of transit observations (and very long observation times) required. The Table shows how certain factors reduce the overall detection numbers. For example, 36 targets are eliminated because their host star is too bright, leading to a detector saturation condition for this instrument/mode. Of course the most significant screening factor is the lack of a strong detection (dBIC < 10) of the high metallicity atmospheres for 828 targets. A number of these targets may show a detection of the low metallicity atmosphere, but given that the high metallicity atmospheres are undetectable, we consider these to be ambiguous detections and eliminate them from further consideration.

We can see that there are 49 targets with $R_p > 10R_{\oplus}$, properly falling into Category 7 (sub-Jovians), but that are outside the range of our analysis. Also, we note (somewhat surprisingly) that there is only one target where the number of
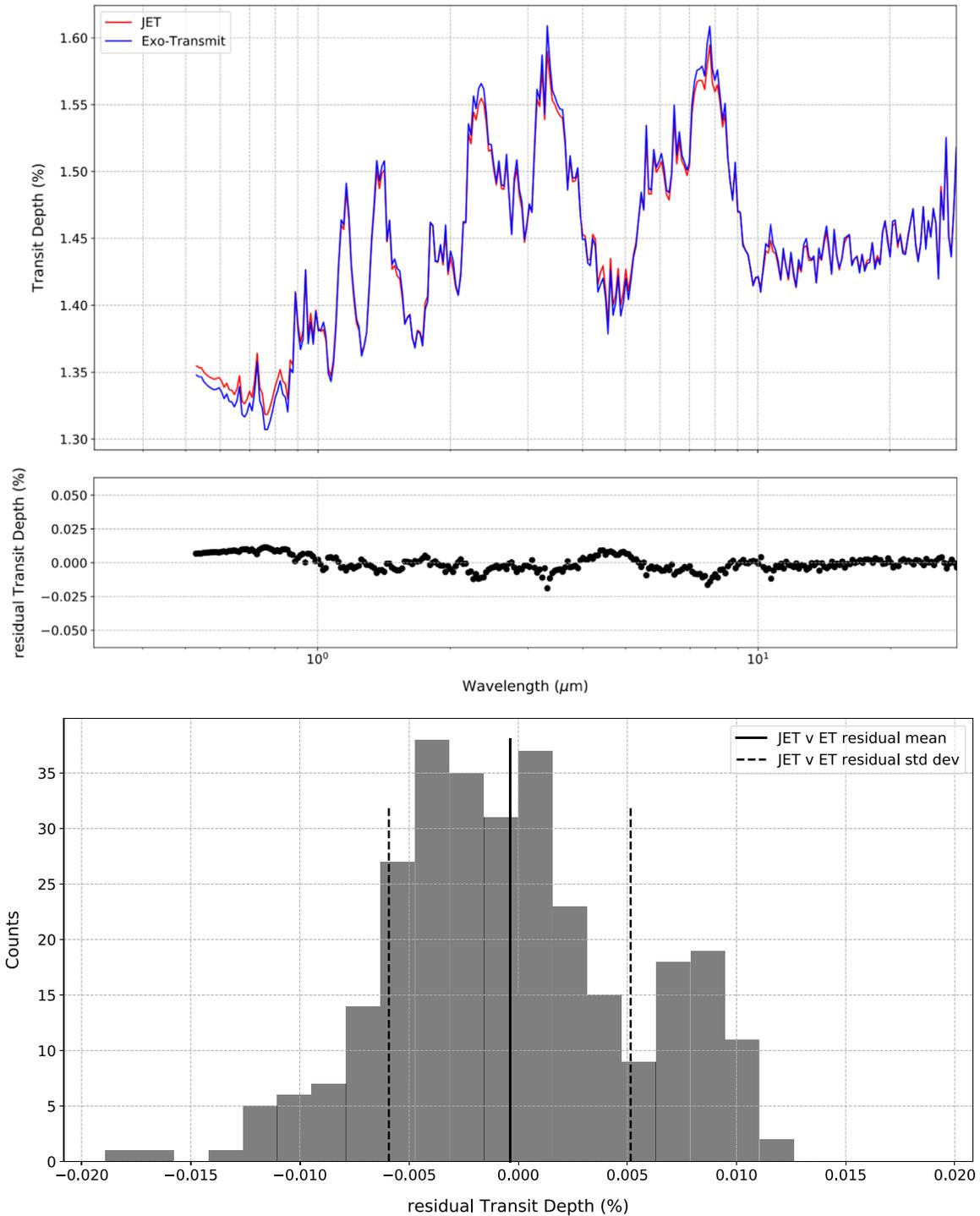

Figure 9. Top: comparison of a test transmission spectrum for GJ 1214b generated by our `JET` code with a benchmark spectrum generated by `Exo-Transmit` (E. M.-R. Kempton 2018, personal communication) shows excellent agreement across a broad spectral range. Center: residuals are shown at the same scale. Bottom: the `JET`-to-`Exo-Transmit` benchmark residuals are small and show no systematic bias. The mean of the residuals is approximately −4 ppm, with a standard deviation of approximately 55 ppm.
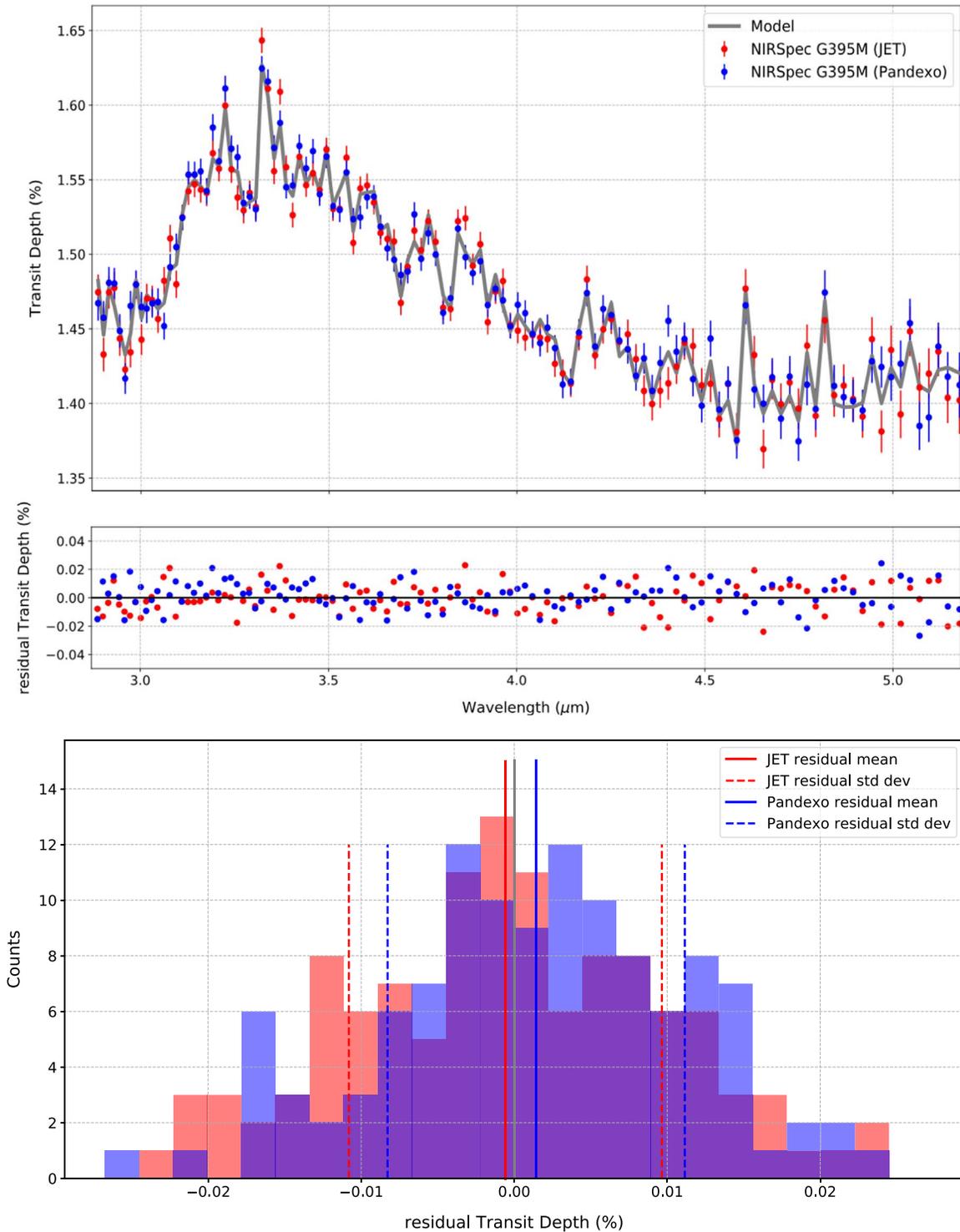
Figure 10. Top: comparison of a simulated spectrum for GJ 1214b generated by our `JET` code with a benchmark spectrum generated by `PandExo` (N. Batalha, 2018 personal communication) shows good agreement across the bandpass. Center: residuals are shown at the same scale. Bottom: here we overlay the `JET`-to-model residual histogram on the `PandExo`-to-model residual histogram. There is good agreement. The difference in the mean of the residuals is approximately 20 ppm, with the difference in standard deviation of the residuals approximately 5 ppm.

Table 6
Excerpt of Baseline Run (top 10 Targets for Each Category Shown)

Cat	Rank	Targ	R_P ($R_\oplus$)	P (days)	S ($S_\oplus$)	R_* ($R_\odot$)	T_{eff} (K)	J_{mag} (mag)	a ($R_\odot$)	T_{eq} (K)	g_s (m s^{-2})	M_P ($M_\oplus$)	t_{dur} (hr)	$nt_{10\,yr}$	nt_{lo}	nt_{hi}	tT_{avg} (hr)
1	1	922	1.40	4.39	1.17	0.12	2729	11.3	5.1	290	12.72	2.53	0.8	300	1	4	9.8
1	2	98	1.64	5.61	4.14	0.23	3325	8.6	8.0	397	12.11	3.33	1.3	173	1	3	10.1
1	3	105	1.68	7.58	1.12	0.14	3130	11.3	8.5	287	12.02	3.48	1.1	256	1	4	11.3
1	4	139	1.60	8.11	2.45	0.21	3170	8.1	8.8	348	12.20	3.20	1.6	145	1	3	11.5
1	5	1248	1.31	7.07	3.31	0.24	3351	7.9	9.5	376	12.98	2.26	1.4	275	1	4	13.4
1	6	204	0.77	14.32	1.27	0.24	3345	7.9	15.1	295	6.35	0.39	1.8	73	1	4	15.4
1	7	1247	1.41	4.24	3.59	0.16	3224	11.3	5.7	383	12.68	2.57	1.0	458	1	7	17.2
1	8	1919	1.40	9.41	0.51	0.12	2838	12.2	8.7	235	12.71	2.54	1.1	173	1	7	18.2
1	9	1878	1.08	9.96	0.78	0.14	3128	11.4	10.2	262	10.74	1.28	1.1	97	1	9	23.3
1	10	45	1.05	15.96	0.63	0.16	3228	10.9	14.0	248	10.27	1.15	1.5	70	1	8	25.1
2	1	1622	0.71	0.77	38.12	0.17	3231	11.4	1.9	692	5.48	0.28	0.6	2531	1	2	4.9
2	2	1856	1.55	0.55	32.21	0.13	3055	13.2	1.3	664	12.31	3.03	0.4	2077	1	3	6.0
2	3	17	1.11	0.79	53.39	0.22	3304	11.1	2.2	753	11.26	1.42	0.7	1446	1	3	7.0
2	4	857	1.50	2.27	7.37	0.16	3141	11.3	3.8	459	12.44	2.86	0.8	886	1	4	9.6
2	5	1255	1.22	0.59	26.65	0.14	2871	12.0	1.4	633	13.11	2.00	0.5	3223	1	6	10.7
2	6	882	1.06	1.61	15.80	0.19	3242	11.1	3.2	555	10.36	1.18	0.8	1187	1	5	11.2
2	7	893	1.07	2.56	7.17	0.17	3155	11.0	4.1	456	10.55	1.23	0.9	773	1	5	12.0
2	8	805	1.67	1.82	14.63	0.13	3758	12.5	3.1	545	12.06	3.41	0.7	571	1	6	12.2
2	9	1533	1.30	0.56	61.35	0.18	3231	12.9	1.6	780	13.00	2.24	0.5	3935	2	6	12.8
2	10	1013	1.59	0.88	19.35	0.14	3006	13.0	1.9	584	12.22	3.16	0.6	2176	1	7	13.0
3	1	1745	1.09	0.51	110.20	0.24	3337	10.4	1.7	902	10.88	1.31	0.6	3781	1	2	5.0
3	2	281	0.70	0.85	126.50	0.39	3557	10.1	2.8	934	5.39	0.27	0.9	2697	1	2	6.2
3	3	1421	1.40	0.63	92.11	0.27	3359	10.9	2.1	863	12.70	2.55	0.7	3226	1	3	7.1
3	4	1057	1.04	0.70	108.80	0.31	3425	10.2	2.3	900	10.07	1.10	0.8	1385	1	3	7.5
3	5	378	1.04	0.59	184.10	0.37	3511	10.1	2.1	1026	10.13	1.12	0.8	3290	1	3	7.6
3	6	146	0.89	0.76	113.70	0.34	3476	10.3	2.5	910	7.89	0.64	0.8	2550	1	3	7.8
3	7	419	1.09	0.64	174.90	0.38	3532	9.6	2.3	1013	10.94	1.33	0.8	2971	1	3	7.8
3	8	13	1.56	0.62	71.16	0.22	3283	12.1	1.8	809	12.29	3.06	0.6	3127	1	4	8.5
3	9	1855	1.35	1.12	108.00	0.42	3640	8.1	3.5	898	12.86	2.38	1.1	1148	1	3	9.0
3	10	917	1.65	1.35	89.71	0.43	3670	9.9	4.0	857	12.09	3.36	1.2	720	1	3	9.4
4	1	8	2.81	5.62	2.56	0.17	3228	11.1	7.0	352	10.30	8.28	1.2	173	1	1	4.8
4	2	921	2.68	9.98	0.39	0.12	2729	11.3	8.9	220	10.45	7.65	1.2	132	1	1	4.8
4	3	1810	2.83	3.84	3.46	0.16	2844	9.8	4.6	380	10.28	8.38	1.2	253	1	1	4.8
4	4	1444	2.89	6.90	2.75	0.21	3284	10.3	8.8	359	10.21	8.71	1.4	297	1	1	5.3
4	5	1864	2.56	4.82	0.71	0.11	2529	12.5	5.3	255	10.60	7.06	0.9	201	1	2	6.2
4	6	1804	1.94	17.79	0.45	0.16	2844	9.8	12.7	228	11.52	4.41	1.9	55	1	1	6.6
4	7	1296	1.92	11.56	1.00	0.17	3236	10.0	11.4	278	11.55	4.35	1.4	168	1	2	8.1
4	8	822	2.93	9.05	2.07	0.22	3300	10.6	10.8	334	10.17	8.89	1.6	134	1	2	8.6
4	9	783	2.29	6.20	2.42	0.19	3225	10.1	8.3	347	10.96	5.84	1.2	319	1	3	9.7
4	10	113	3.02	4.62	3.41	0.18	3006	12.5	5.8	379	10.07	9.37	1.3	413	1	3	10.0
5	1	182	2.27	0.57	31.90	0.14	2871	12.8	1.3	662	10.99	5.76	0.5	1778	1	1	3.2
5	2	1930	3.46	1.81	5.66	0.13	2877	11.6	3.0	430	9.67	11.79	0.8	613	1	1	3.8
5	3	934	3.64	2.16	11.39	0.19	3253	11.2	3.9	512	9.52	12.85	1.0	883	1	1	4.2
5	4	1260	3.91	1.91	19.91	0.25	3342	10.7	4.0	588	9.32	14.55	1.0	1015	1	1	4.4
5	5	901	3.80	2.81	5.40	0.16	2793	12.3	3.6	425	9.40	13.83	1.2	346	1	1	4.8
5	6	1130	2.01	0.66	30.50	0.15	2999	12.6	1.6	655	11.40	4.68	0.5	1747	1	2	4.8
5	7	60	3.19	2.46	34.13	0.41	3587	10.0	5.8	673	9.91	10.30	1.4	677	1	1	5.3
5	8	183	1.94	0.82	66.23	0.26	3363	11.2	2.4	795	11.52	4.40	0.7	1179	1	2	5.6
5	9	1050	3.91	1.94	8.92	0.15	3206	13.5	3.3	481	9.32	14.52	0.8	999	1	2	5.9
5	10	33	3.85	1.85	11.79	0.17	3225	13.0	3.4	516	9.36	14.18	0.9	1048	1	2	6.0
6	1	1956	2.38	0.60	95.05	0.25	3336	11.6	1.9	870	10.83	6.24	0.7	3258	1	1	3.5
6	2	993	1.91	0.52	98.50	0.23	3300	11.9	1.6	877	11.57	4.30	0.6	1885	1	2	5.0
6	3	1903	2.03	0.52	150.20	0.30	3407	11.7	1.8	975	11.36	4.78	0.7	1871	1	2	5.4
6	4	945	3.87	3.22	307.70	0.82	5383	8.3	8.8	1167	9.34	14.30	2.4	539	1	1	7.7

Table 6
(Continued)

Cat	Rank	Targ	R_p ($R_\oplus$)	P (days)	S ($S_\odot$)	R_* ($R_\odot$)	T_{eff} (K)	J_{mag} (mag)	a ($R_\odot$)	T_{eq} (K)	g_s (m s^{-2})	M_p ($M_\oplus$)	t_{dur} (hr)	$nt_{10\,yr}$	nt_{lo}	nt_{hi}	tT_{avg} (hr)
6	5	115	1.76	1.19	431.20	0.64	4583	7.5	4.2	1269	11.87	3.74	1.4	813	1	2	8.1
6	6	231	2.88	1.87	99.23	0.53	3918	10.2	5.3	879	10.21	8.67	1.5	612	1	2	8.3
6	7	1682	3.98	4.27	230.10	0.86	5470	8.3	10.9	1085	9.27	14.98	2.7	447	1	1	8.4
6	8	866	3.59	2.13	72.91	0.51	3824	10.4	5.6	814	9.56	12.54	1.6	521	1	2	8.5
6	9	1516	1.77	0.67	307.80	0.49	3804	9.8	2.6	1167	11.83	3.79	1.0	5152	1	3	8.6
6	10	207	1.79	0.64	83.65	0.25	3330	12.2	1.9	842	11.80	3.85	0.7	2965	1	4	8.8
7	1	79	7.51	2.07	129.10	0.60	4147	11.1	5.8	939	7.65	43.99	1.8	955	1	1	6.3
7	2	104	5.07	3.86	52.77	0.58	4122	10.7	8.8	751	8.61	22.58	2.1	252	1	1	7.0
7	3	860	7.70	1.66	868.70	0.86	5470	8.4	5.6	1512	7.59	45.94	2.1	629	1	1	7.0
7	4	905	6.61	2.82	264.50	0.74	5152	9.7	7.8	1123	7.95	35.45	2.2	626	1	1	7.3
7	5	352	8.93	3.56	116.10	0.67	4673	10.8	8.8	914	7.26	59.02	2.3	546	1	1	7.5
7	6	1898	6.77	4.63	42.24	0.59	4134	10.5	9.9	710	7.89	36.87	2.3	412	1	1	7.5
7	7	1784	6.88	4.46	101.40	0.69	4835	10.0	10.3	884	7.86	37.91	2.5	226	1	1	7.9
7	8	168	9.26	6.53	26.21	0.58	4144	10.3	12.5	630	7.18	62.77	2.6	149	1	1	8.3
7	9	608	9.71	4.15	153.20	0.79	5000	10.9	10.2	980	7.08	68.04	2.7	485	1	1	8.4
7	10	1213	8.34	2.69	710.60	1.00	5834	8.7	8.3	1438	7.41	52.55	2.7	839	1	1	8.4

(This table is available in its entirety in machine-readable form.)

Table 7
Baseline Run Statistics 1

					Non-detection			
Cat	Description	Targets	$J_{mag} < $ lim Targets	$R_p > 10R_\oplus$ Targets	dBIC< 10 for $nt_{lo} < 50$ Targets	dBIC< 10 for $nt_{hi} < 50$ Targets	nt req'd $> nt_{10\,yr}$ Targets	Full Detection Targets
1	Cool Terrestrials	41	1	⋯	⋯	21	⋯	19
2	Warm Terrestrials	134	2	⋯	2	53	⋯	79
3	Hot Terrestrials	119	6	⋯	5	43	⋯	70
4	Cool Neptunes	371	1	⋯	⋯	128	1	241
5	Warm Neptunes	768	7	⋯	17	343	⋯	418
6	Hot Neptunes	400	12	⋯	27	214	⋯	174
7	sub-Jovians	151	7	49	1	26	⋯	69
Total		1984	36	49	52	828	1	1070

Note. Some overlap in non-detection conditions ("Total" line will not sum across row).

transits needed for detection is greater than the number observable within the 10 yr fuel life of the spacecraft.

In Table 8 we slice the data a different way. Columns three and four show numbers of targets that meet certain cutoff values for nt_{hi}, and tT_{avg}. Column five shows accumulated observing hours by each category. We see that for $nt_{hi} < 10$ and $tT_{avg} < 35$ hr, we detect roughly 600 targets with approximately 11,000 hr of observing time. Likewise, the next three columns show that for $nt_{hi} < 6$ and $tT_{avg} < 20$ hr, we detect roughly 300 targets with approximately 4000 hr of observing time.

From Table 8, we can also see, for example, that for Category 4 (Cool Neptunes) there are 37 targets where a high metallicity atmosphere (the difficult case) can be detected with less than 6 transit observations, and 45 targets where an average atmosphere can be detected in under 20 hr of observation for each target.

In Table 9 we show the best targets for an 8300 hr (10% of the anticipated 10 yr mission total observing time) program by category for (1) the full Sullivan catalog, (2) the JET baseline ranking with all viable targets for Cat 1 (Cool Terrestrials), and the remaining hours applied evenly for other categories, (3) the JET baseline ranking with all viable targets for Cat 1, and a fixed number of targets for other categories, and (4) planets with actual atmospheric characterization by transmission spectroscopy to date, based on the NASA Exoplanet Archive (and associated references).

Table 8
Baseline Run Statistics 2

Cat	Description	$nt_{hi} < 10$ Targets	$tT_{avg} < 35$ hr Targets	accum hr by Cat[a]	for $nt_{hi} < 6$ Targets	for $tT_{avg} < 20$ hr Targets	accum hr by Cat[b]
1	Cool Terrestrials	10	9	167	4	4	50
2	Warm Terrestrials	29	37	768	4	20	295
3	Hot Terrestrials	50	51	898	21	35	468
4	Cool Neptunes	107	101	2165	37	45	611
5	Warm Neptunes	211	207	4077	96	108	1443
6	Hot Neptunes	131	108	2165	75	57	809
7	sub-Jovians	59	52	791	51	39	458
Total		597	565	11,032	288	308	4133

Notes. Columns 3, 4, 6, and 7 are numbers of targets; columns 5, and 8 are accumulated observing hr.
[a] Based on $tT_{avg} < 35$ hr column.
[b] Based on $tT_{avg} < 20$ hr column.

Table 9
Best Targets for 8300 hr *JWST* Transmission Spectroscopy Program

Cat	Description	Sullivan (full)[a]	JET targets Even hr[b]	JET targets Even Targets[c]	Actual atms Characterized to Date[d]
1	Cool Terrestrials	41	19	19	4
2	Warm Terrestrials	134	52	67	1
3	Hot Terrestrials	119	62	67	...
4	Cool Neptunes	371	74	67	1
5	Warm Neptunes	768	105	67	2
6	Hot Neptunes	400	83	67	1
7	sub-Jovians	151	67	67	39
Total		1984	462	421	48

Notes.
[a] From the simulated Sullivan et al. (2015) *TESS* Catalog.
[b] JET baseline ranking with all viable targets for Cat 1 (Cool Terrestrials), and remaining hours applied evenly for other categories (8300 hr program total).
[c] JET baseline ranking with all viable targets for Cat 1, and a fixed number of targets for other categories (8300 hr program total).
[d] Planets with actual atmospheric characterization by transmission spectroscopy to date, based on the NASA Exoplanet Archive.

We show the target planets plotted by category on a radius (R_p) versus equilibrium temperature (T_{eq}) grid in Figure 11. In this figure we combine a plot of the full Sullivan catalog (1984 targets), with the best 8300 hr JET target list from the Sullivan catalog (462 targets; as shown in Col 4 of Table 9), and with planets with actual atmospheric characterization (by transmission spectroscopy) to date (48 targets).[25] There is some ambiguity regarding which planetary atmospheres have actually been "characterized." Many of the spectroscopic observations obtained so far are of very low resolution and in some cases only multi-band photometry has been employed.

As we mentioned in Section 1.3.2, a recent paper by Kempton et al. (2018) ranked the Sullivan targets for transmission spectroscopy using an analytical metric (the

transmission spectroscopy metric, TSM). In Table 10 we show the top habitable zone transmission spectroscopy targets taken from Table 2 of their paper, with the JET target ranking side-by-side. There is general agreement, but differences are evident.

The TSM is an algebraic function of the planetary radius, mass, equilibrium temperature, and host star radius, that is proportional to the expected S/N of the observation. In our analysis we use a total observation time figure of merit (tT_{avg}) for ranking that is based on a statistical analysis of the fit of simulated spectra to model spectra for each target. Our ranking figure of merit is also related to the expected S/N of the observation, but in a more indirect and complex way. The imperfect agreement between the two rankings is reasonable given the significantly different analytical approaches taken.

We compare our ranking metric, tT_{avg}, to the TSM in Figure 12. This is for their "statistical" sample. We see a good

---

[25] https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=transitspec
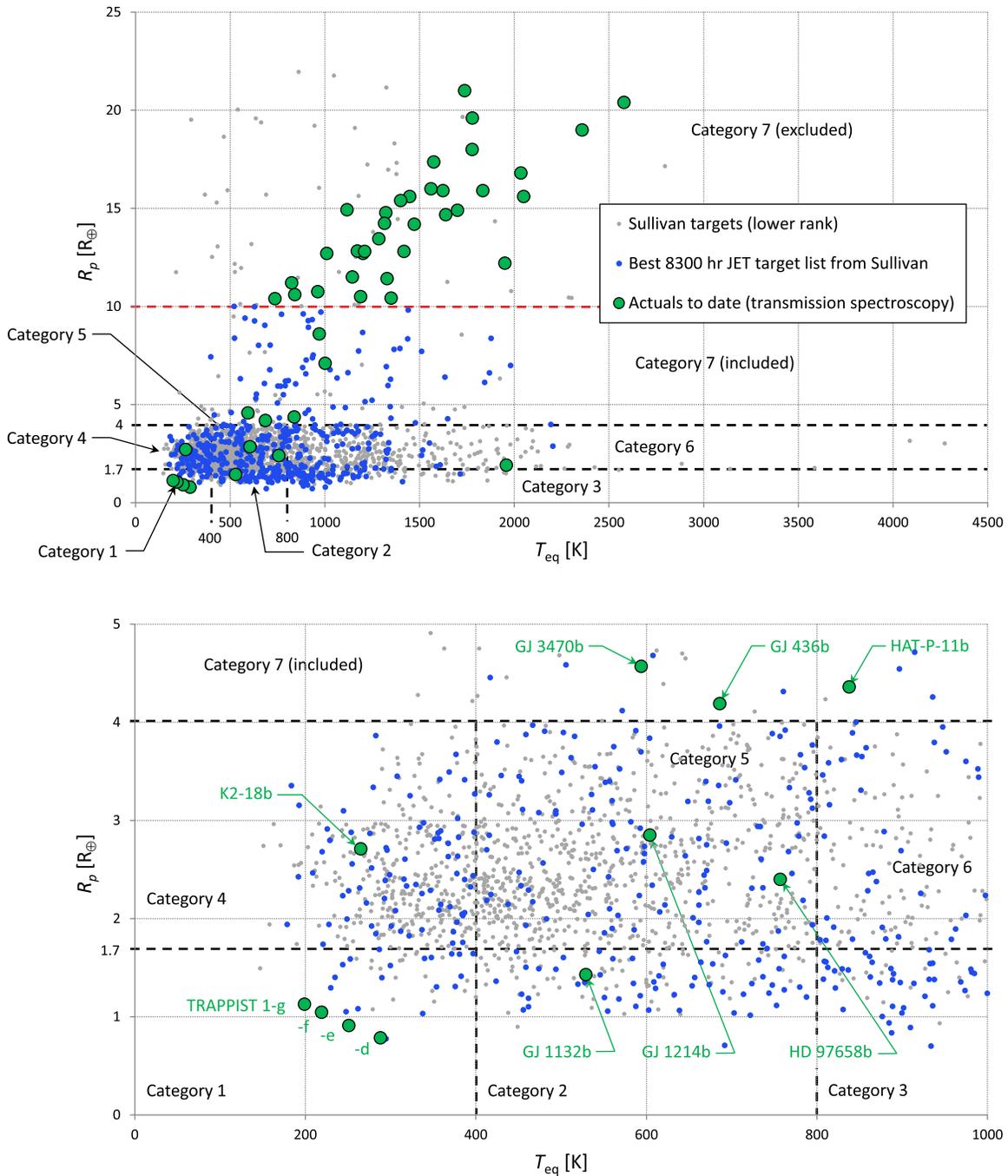
Figure 11. Target planets plotted by category on a radius (R_p) vs. equilibrium temperature (T_{eq}) grid. Top: full range of Sullivan catalog (1984 targets), the best 8300 hr JET target list from Sullivan (462 targets), and actual transmission spectroscopy targets to date (48 targets). Bottom: same data focused on the densest region of the plot, with planet names labeled.

general correspondence between the two approaches, but also some dispersion between these two ranking schemes.

In Figure 13 we show the comparison for their "small temperate" target sample. Again, we see good correspondence with only a small amount of dispersion.

4.2. Noise Floor Variation

Variation of the instrument noise floor has a powerful effect on the number of transits needed for detection for the more difficult targets (those with weak spectral features and bright host stars). Figure 14 presents the results of our noise floor

Table 10
Top Habitable Zone Transmission Spectroscopy Targets Ranked by Kempton et al. (2018) and by JET

		Kempton et al. (2018) Ranking						JET Ranking		
Targ[a]	TSM[b]	M_p	R_p	M_*	R_*	S	J_{mag}	Cat	Rank	tT_{avg}
		$(M_\oplus)$	$(R_\oplus)$	$(M_\odot)$	$(R_\odot)$	$(S_\oplus)$	(mag)		(in Cat)	(hr)
204	27.9	0.39	0.77	0.2	0.24	1.27	7.91	1	6	15.4
1296	26.8	4.35	1.92	0.12	0.17	1	10	4	7	8.1
1804	26.5	4.42	1.94	0.07	0.16	0.45	9.78	4	6	6.6
1308	23.2	2.83	1.49	0.12	0.25	1.15	7.97	1	no detection[c]	⋯
922	21.6	2.53	1.39	0.06	0.12	1.17	11.26	1	1	9.8
405	19.4	3.11	1.58	0.26	0.38	1.61	6.85	1	saturated[d]	⋯
105	17.9	3.48	1.68	0.1	0.14	1.12	11.27	1	3	11.3
48	17.3	4.64	1.99	0.12	0.16	0.64	11.1	4	12	10.8
1244	16.8	3.99	1.82	0.11	0.16	1.79	11.34	4	42	16.8
991	15.8	3.67	1.74	0.1	0.16	0.39	10.53	4	19	11.9

Notes.
[a] Target planet numbers from the simulated Sullivan et al. (2015) *TESS* Catalog.
[b] Transmission Spectroscopy Metric, for Scale factor = 0.167, calculated for small temperate sample.
[c] No detection, dBIC < 10 for 50 transits with hi-metal atm.
[d] Host star magnitude exceeds brightness limit of detector.

variation study for Sullivan Target 1292. As discussed in Section 2.3.1, the noise floor for NIRSpec G395M is estimated to be approximately 25 ppm. Of course, the true value of the noise floor will not be known until *JWST* is launched and has completed detailed commissioning procedures.

Our baseline run for this target with a noise floor of 25 ppm shows that we need 17 transits to detect the high-metallicity atmosphere. This falls to 7 transits for a noise floor of 20 ppm. Remarkably, if the noise floor turns out to be only 1 ppm higher than the baseline (or 26 ppm) there will be no detection of the atmosphere no matter how many transits are observed. Clearly the target ranking can be affected significantly by minor variations in the estimate of the instrument noise floor.

It should be pointed out that in the figure the error bars do not indicate uncertainty (which could be significant), but are meant to indicate that the number of transits has been rounded up to the next highest whole transit.

4.3. Atmospheric EOS Variation

Variation of the atmospheric metallicity has a very significant effect on our results. As we can see from Figure 5 of Kempton et al. (2017), spectral line strength is strongly driven by metallicity. Again, for Sullivan Target 1292, in Figure 15 we can see that the number of transits needed for detection falls almost by an order of magnitude between the 1000× Solar and the 50× Solar metallicity levels. Unfortunately, the only intermediate level that we could examine was a 100× Solar metallicity. The EOS files available in the Exo-Transmit installation do not include files in the regime between 100× and 1000× Solar. It was beyond the scope of this project to construct a new EOS file to fill in the gap. It is unlikely, however, that any intermediate data points would

significantly change our conclusions about the number of transits needed for detection here.

We did not directly examine the variation in our results due to changes in cloud levels. Again, Figure 5 of Kempton et al. (2017) guides what we would expect to see. Since we were trying to bound the atmospheric cases, our low metallicity (5× Solar) case with no clouds would seem to be a good choice for a best-on-best case. The high metallicity (1000× Solar) case with clouds at 100 mbar would seem to be a nearly worst-on-worst case. As the figure indicates, the effect of clouds (at any altitude) on the strength of spectral features at this high metallicity is not very strong.

4.4. Detection Threshold Variation

The number of transits needed for detection is significantly effected by our choice of dBIC detection threshold. Kass & Raftery (1995) suggest that a dBIC of 6 can be considered a strong detection, while a dBIC of 10 is very strong. Not surprisingly, our Figure 16 shows that for Sullivan Target 1292 with a high metallicity atmosphere, as we lower the detection threshold from 10 to 5, we see a drop in the number of transits required for detection from 17 to 8.

Given that we are using the dBIC mean less 1σ values for detection (the bold line shown in Figure 7 in Section 2.3.2), the threshold of 10 gives us a confidence level for a very strong detection of approximately 84%.

4.5. Instrument Variation

JWST has four main instruments, each with multiple observing modes. All four (NIRCam, NIRSpec, NIRISS, and MIRI) will to varying degrees be used for exoplanet
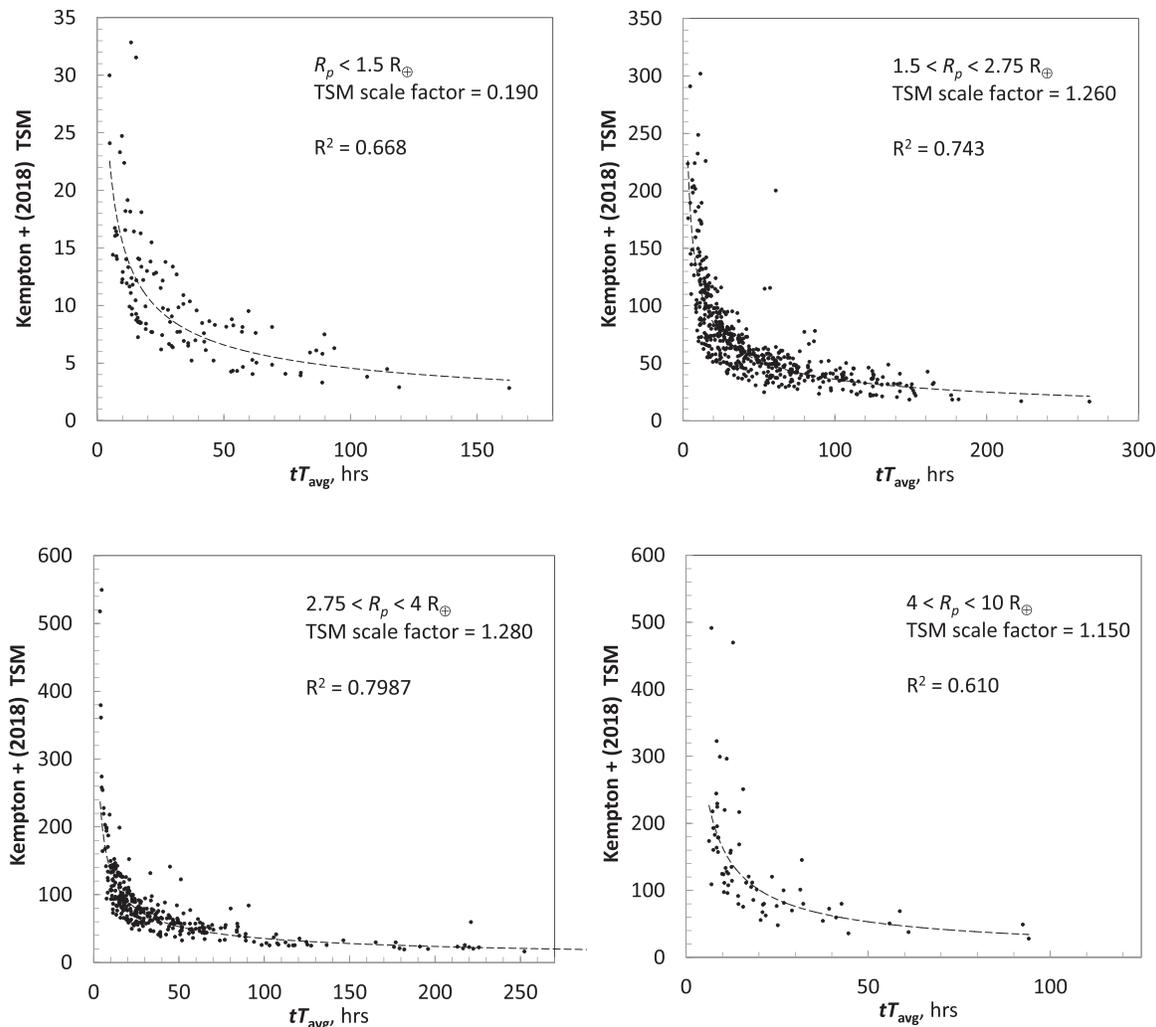
Figure 12. Comparison of the JET ranking metric, tT_{avg}, to the transmission spectroscopy metric, TSM, from Equation (1) of Kempton et al. (2018) for the full baseline JET run (1070 targets with unambiguous detection). Here we are using the Kempton demographic categories and scale factors for their "statistical" sample. The black dashed lines mark a power law least squares fit to the data, with R^2 values shown in each case. The Kempton ranking tracks the JET ranking reasonably well in all cases, and particularly well for the $2.75 < R_p < 4 R_\oplus$ demographic.

transmission spectroscopy. We have been able to implement four of these instrument/modes with JET. We previously presented single-transit spectra for NIRSpec G395M in Figure 6. We now present additional single-transit spectra for each of the other instrument/modes in Figures 17 through 19.

We have summarized results of variation of the *JWST* instrument/mode on the detection of a high metallicity atmosphere for Sullivan Target 1292 in Table 11.

We see that only two transits are needed to detect the target with NIRISS SOSS, while 17 transits are needed using NIRSpec G395M. This does not necessarily imply that the former instrument is better than the latter. The wavelength ranges are different, and the capabilities of the instruments are complementary rather than competing. It has been suggested (Batalha & Line 2017) that these two instruments be used in

tandem, since there is little overlap in their wavelength coverage, and both have relatively high precision.

5. Conclusions/Further Study

5.1. Conclusions

We have developed an analytical framework and associated computer code that can assist the community in determining the best exoplanet targets for atmospheric characterization by *JWST*. The tools can also be used to prioritize targets that would be worthy of follow up with RV observations to better determine the planet masses.

We have demonstrated that a target catalog can be categorized and ranked for minimum observation time to
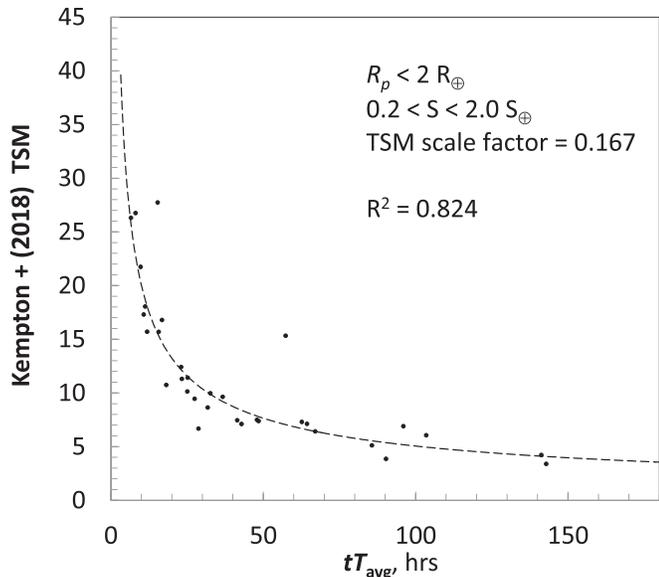
Figure 13. Same as the previous figure, but now showing only the "small temperate" target sample. The Kempton ranking tracks the JET ranking very well for this demographic.


Figure 15. Effect of atmospheric metallicity on number of transits needed for detection for Sullivan Target 1292, a detection threshold (dBIC) of 10, and using NIRSpec G395M with a noise floor of 25 ppm, and no clouds. This behavior is consistent with Figure 5 of Kempton et al. (2017). We were limited in the choice of metallicity levels by the equation of state files available in the Exo-Transmit installation.

Figure 14. Effect of noise floor variation on number of transits needed for detection, for Sullivan Target 1292, with a high metallicity atmosphere (1000× Solar), clouds at 100 mbar, a detection threshold (dBIC) of 10, and using NIRSpec G395M. There is no detection for a noise floor above approximately 25 ppm. The number of transits are rounded to the next highest whole number of transits.
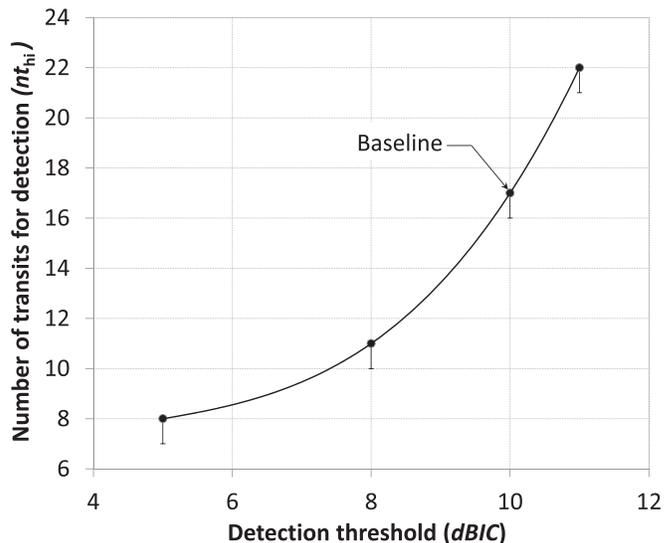
Figure 16. Effect of lowering the detection threshold (dBIC) on the number of transits needed for detection for Sullivan Target 1292, a high metallicity atmosphere (1000× Solar), with clouds at 100 mbar, and using NIRSpec G395M with a noise floor at 25 ppm. A dBIC of 10 is considered a very strong detection, while a dBIC of 6 is still considered a strong detection.

detect an atmosphere, even though we have no direct knowledge of the atmospheric properties of the target planets.

Our use of the Bayesian Information Criterion for model selection in our target atmosphere detection algorithm is one of the strengths of this work. In addition, any target prioritization tool must take into consideration all of the observing time

overheads for a transit observation. Even a very short transit can be expensive in terms of the overall observing time requirement.

In the Sullivan (simulated *TESS* detection) catalog we only saw one target where the number of transits observable during
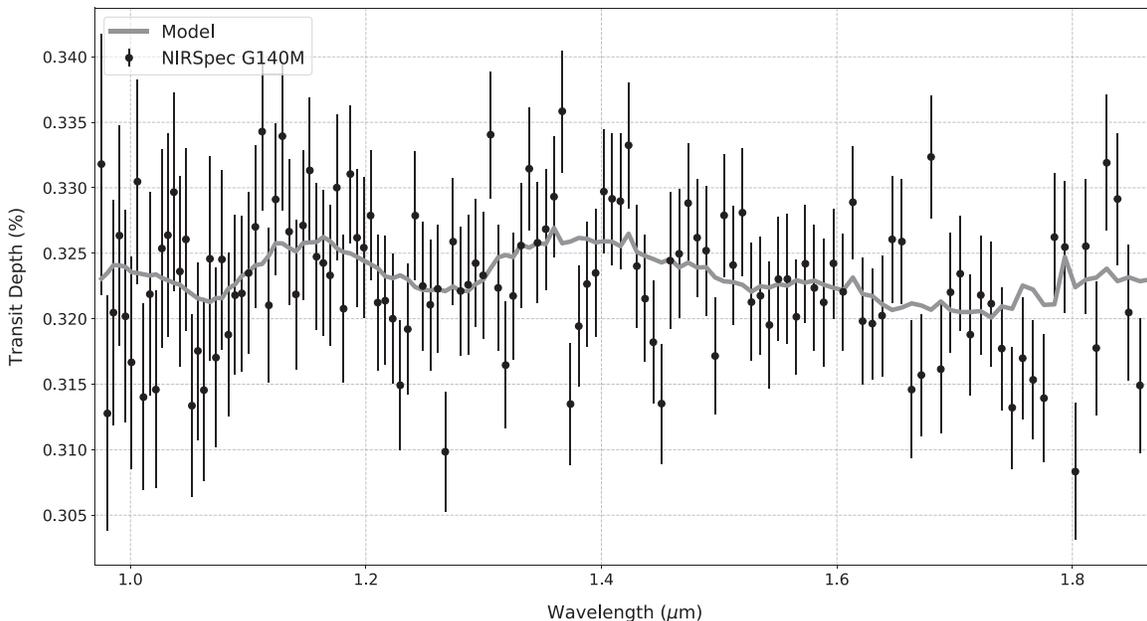
Figure 17. Simulated spectrum for a single transit of Sullivan Target 1292 with a high-metallicity atmosphere and clouds at 100 mbar, using NIRSpec G140M with a wavelength range of 0.97–1.87 μm. The noise floor for this instrument is approximately 25 ppm. The model spectrum shown as the gray background line has been binned down to a resolution ($R \sim 100$) consistent with the simulated data. A dBIC = 10 detection takes observation of four transits.
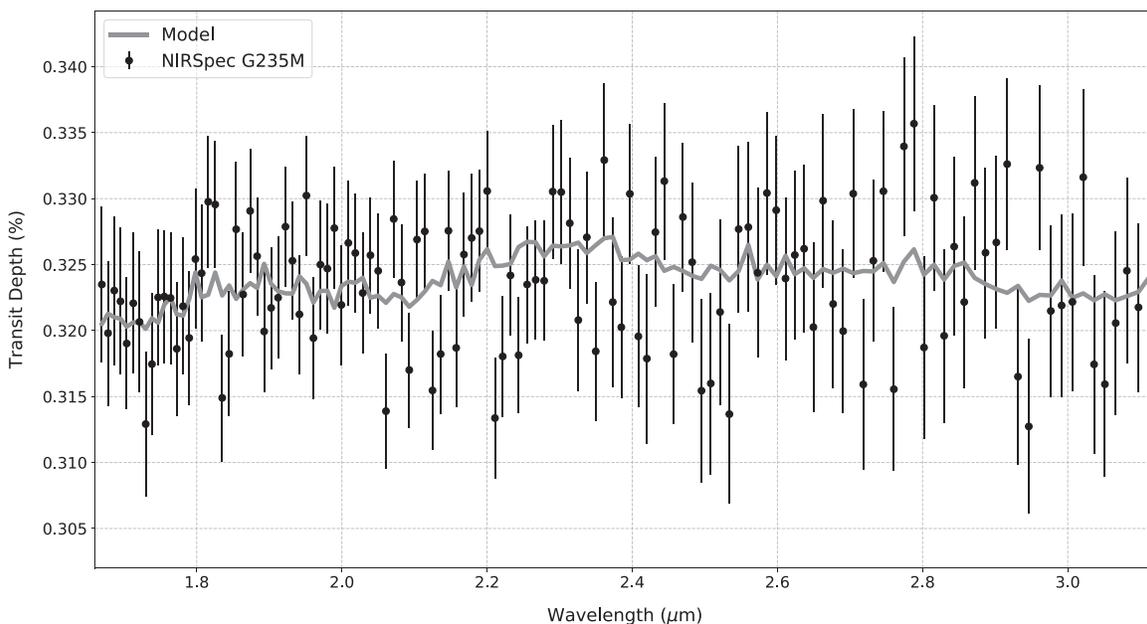

Figure 18. Similar to Figure 17, but for the G235M disperser and a wavelength range of 1.66–3.12 μm. A dBIC = 10 detection takes observation of five transits.

the fuel life of the spacecraft was less than the number needed for an atmospheric detection. The observing constraints imposed by the spacecraft's orbital location and pointing limitations do not appear to be significant, at least for the shorter period planets that are the focus of surveys like those coming from *TESS*.

The instrument noise floor is a critical parameter. Particularly for difficult targets (small transit depth and reduced feature prominence), a very small change in the assumption for the noise floor can change the number of transits needed for detection significantly. This, of course, could substantially change the overall target rankings. Once the instruments have
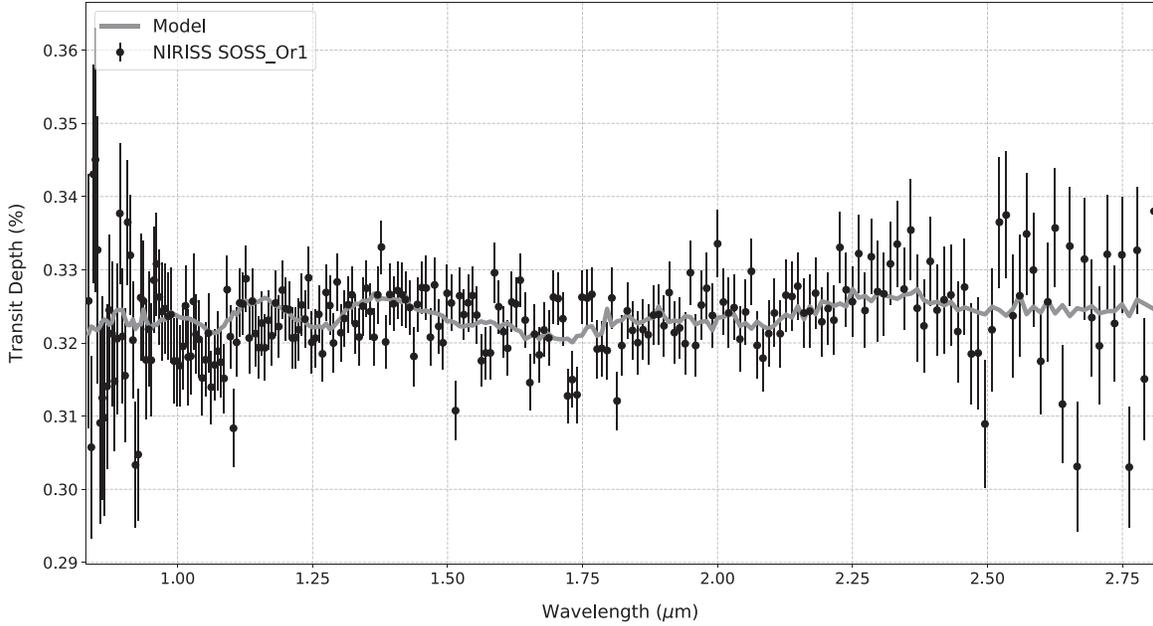
Figure 19. Simulated spectrum for a single transit of Sullivan Target 1292 with a high-metallicity atmosphere using NIRISS SOSS (Order 1) with a wavelength range of 0.83–2.82 μm. The noise floor for this instrument/mode is approximately 20 ppm. The model spectrum shown as the gray background line has been binned down to a resolution ($R \sim 100$) consistent with the simulated data. A dBIC = 10 detection takes observation of two transits.

Table 11
Effect of Instrument Choice on Detection (Sullivan Target 1292)

Instrument/Mode	R	nfloor (ppm)	λ Range (μm)	nt_{hi}
NIRSpec G395M	100	25	2.87–5.18	17
NIRSpec G235M	100	25	1.66–3.12	5
NIRSpec G140M	100	25	0.97–1.87	4
NIRISS SOSS_Or1	100	20	0.83–2.81	2

Note. Noise floor values (nfloor) are pre-launch estimates only. The R values shown are not the native resolving power of the instrument, but are the re-binned values for consistency between the model spectra and simulated spectra in our analysis.

been better characterized on orbit (during commissioning), the prioritization analysis will need to be repeated to reflect the updated information on instrument precision.

Perhaps the weakest element in our analytical approach is our assumption that each target effectively has an atmosphere of average metallicity. This may be true if we consider the entire data set, but it is almost certainly not true for any particular target. In reality any given target may be skewed away from average metallicity, either high or low. This, of course, would affect the detectability of the atmosphere for that target, and the ranking. Our detection model provides a reasonable basis for ranking, but we must accept that there are uncertainties.

To date there have been less than 50 exoplanet atmospheres that have been characterized by transmission spectroscopy. Most of these studies have been at relatively low resolution and some only consist of a few data points gathered by multi-band photometry. *JWST* has the potential to increase the number of atmospheres characterized by at least an order of magnitude.

Our baseline run of the Sullivan catalog showed detection of over 1100 planet targets (atmospheres). This figure significantly overstates the realistic/practical target set that could be observed. The fraction of the 10 yr *JWST* mission devoted to exoplanet transmission spectroscopy is estimated to be on the order of 10%. This translates into roughly 8300 hr overall. Our analysis indicates that between 400 and 500 target/ atmospheres could realistically be detected and studied during the mission using NIRSpec G395M. Observations with other instrument/modes (e.g., NIRISS SOSS) could well add to this target/atmosphere characterization total.

It should also be noted, as we discussed in Section 1.3.1, that there have been several new studies of potential *TESS* planet yields since the Sullivan paper was published. The categorization and ranking results could well be different using these updated catalogs.

JWST will provide access to many important atmospheric spectral features in the infrared, including water vapor, methane, carbon dioxide, carbon monoxide, ammonia, sodium, potassium, and others. In addition, this access will be over a much wider wavelength range and at much higher resolution than what has been done previously (Madhusudhan 2019). This wealth of new spectral data will dramatically enhance our

knowledge of many exoplanet physical processes, including planetary formation, geology, climate, environment, and potential habitability (Kalirai 2018).

With *JWST* we will be in a position to begin to assess the demographics of planetary atmospheres and the relationship with the properties of the planet's host stars. In particular, determining planetary atmosphere metallicity can provide insight into formation scenarios. The wide wavelength range of *JWST* will allow us to determine overall atmospheric spectral shapes and to probe vertical energy redistribution. The effects of inversions and other temperature distributions can potentially be studied. We also expect to be able to determine atmospheric properties of smaller rocky planets orbiting M-dwarf host stars. It may in certain cases be possible to detect atmospheric biosignatures, but this will be challenging for *JWST* (Krissansen-Totton et al. 2018; Lustig-Yaeger et al. 2019). Unfortunately the characterization of the atmosphere of an Earth-like planet around a Sun-like star will be beyond the reach of *JWST*. For that we will have to wait for the next generation of space observatories (Gaudi et al. 2019a, 2019b; Roberge et al. 2019a, 2019b; The LUVOIR Team 2019).

5.2. Opportunities for Further Study

We have identified a number of opportunities to refine the analysis that we have described in this work and to consider other lines of investigation.

As we have mentioned, the JET code is slow when run on a typical home computer. An effort to make the code more efficient would be helpful. In particular, consideration of processing multiple targets with a parallel processing architecture would seem to be a worthwhile effort. In addition, a new open-source Python package, PLanetary Atmospheric Transmission for Observer Noobs (PLATON), described by Zhang et al. (2018), may reduce the computation time necessary for generating model transmission spectra. Unfortunately, this code was not available when we made the decision to use Exo-Transmit.

As we discussed, we have been able to implement the NIRSpec G395M and NIRISS SOSS (Order 1) with JET. Doing a full catalog run with NIRISS SOSS and then combining the output with our existing NIRSpec G395M baseline would cover a wide wavelength range (0.83–5.18 μm), and could yield interesting results.

In addition, for the sake of completeness, it would be helpful to expand the list of instrument/modes that JET could address. This would include the higher resolution modes for NIRSpec, the NIRCam grism (with F322W2 and/or F444W filters), the second order NIRISS SOSS mode, and MIRI LRS.

In Section 4 we discussed single target parameter variations of noise floor, detection threshold, etc. We could consider full runs with small parameter variations. This would provide a more thorough study of how small parameter variations could change overall target rankings.

Of course the most useful test of the analysis framework and code would be to run it on actual *TESS* (or other precursor) catalog data. *TESS* is delivering data sets periodically over its two year mission. It would be helpful to take these early data sets and format them (similar to the Sullivan catalog) or otherwise prepare them to be used as input to the JET code.

With further study we may find ways to refine our estimates of planet atmosphere metallicity that are better than simply taking the average of high and low bounds. This will allow us to prepare target rankings with less uncertainty in the results.

Our efforts so far have been focused on transmission spectroscopy, but exploring emission spectroscopy with our analysis approach would seem to be an interesting area for further study. This would be a major excursion from what we have done so far. The Exo-Transmit code is aimed at transmission spectroscopy, but there are other codes that have the capability to model atmospheric emission spectra as well as transmission; these include: PyDisort (Stamnes et al. 1988), NEMESIS (Irwin et al. 2008), CHIMERA (Line et al. 2013), ATMO (Tremblin et al. 2015), TauRex (Waldmann et al. 2015), HELIOS (Malik et al. 2017), PLATON, and petitRADTRANS (Molliere et al. 2019). PandExo can be used to simulate *JWST* emission (occultation) spectra.

There have been a number of delays in the *JWST* launch schedule; however, there is a positive aspect to the latest delay, in that there should be time to refine the approach to target optimization presented here, or by others, as well as time to incorporate actual *TESS* survey results into the analysis.

Facilities: JWST, *TESS*, Exoplanet Archive.

Software: `astropy` (The Astropy Collaboration et al. 2018), `Exo-Transmit` (Kempton et al. 2017), `Forecaster` (Chen & Kipping 2017), `matplotlib` (Hunter 2007), `numpy` (Oliphant 2015), `PandExo` (Batalha et al. 2017), `scipy` (Virtanen et al. 2019), `SpectRes` (Carnall 2017).

ORCID iDs

Charles D. Fortenbach ⬡ https://orcid.org/0000-0001-5286-639X

Courtney D. Dressing ⬡ https://orcid.org/0000-0001-8189-0233

References

Aronstein, D. L., Smith, J. S., Zielinski, T. P., et al. 2016, Proc. SPIE, 9904, 990409

Ballard, S. 2019, AJ, 157, 113

Barclay, T., Pepper, J., & Quintana, E. V. 2018, ApJS, 239, 2

Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., & Fletcher, L. N. 2015, MNRAS, 448, 2546

Batalha, N. E., & Line, M. R. 2017, AJ, 153, 151

Batalha, N. E., Mandell, A., Pontoppidan, K., et al. 2017, PASP, 129, 064501

Bean, J. L., Stevenson, K. B., Batalha, N. M., et al. 2018, PASP, 130, 114402

Beichman, C., Barrado, D., Belikov, R., et al. 2019, BAAS, 51, 58

Beichman, C., Benneke, B., Knutson, H., et al. 2014, PASP, 126, 1134

Benneke, B. 2013, PhD thesis, (Cambridge, MA: Massachusetts Institute of Technology)

Bouma, L. G., Winn, J. N., Kosiarek, J., & McCullough, P. R. 2017, arXiv:1705.08891

Carnall, A. C. 2017, arXiv:1705.05165

Carroll, B. W., & Ostlie, D. A. 2006, An Introduction to Modern Astrophysics (2nd ed.; London: Pearson)

Chen, J., & Kipping, D. 2017, ApJ, 834, 17

Crossfield, I. J. M., & Kreidberg, L. 2017, AJ, 154, 261

Deming, D., Seager, S., Winn, J., et al. 2009, PASP, 121, 952

Ferruit, P., Birkmann, S., Böker, T., et al. 2014, Proc. SPIE, 9143, 91430A

Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109

Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, SSRv, 123, 485

Gaudi, B. S., Seager, S., Mennesson, B., Kiessling, A., & Warfield, K. 2019a, Proc. SPIE, 11115, 111150M

Gaudi, S., Seager, S., Kiessling, A., Mennesson, B., & Warfield, K. 2019b, BAAS, 51, 89

Gordon, C., & Trotta, R. 2007, MNRAS, 382, 1859

Greene, T., Batalha, N., Bean, J., et al. 2019, BAAS, 51, 61

Greene, T. P., Line, M. R., Montero, C., et al. 2016, ApJ, 817, 17

Grillmair, C. J., Carey, S. J., Stauffer, J. R., et al. 2012, Proc. SPIE, 8448, 84481I

Howe, A. R., Burrows, A., & Deming, D. 2017, ApJ, 835, 96

Hunter, J. D. 2007, CSE, 9, 90

Ingalls, J. G., Krick, J. E., Carey, S. J., et al. 2012, Proc. SPIE, 8442, 84421Y

Irwin, P., Teanby, N., de Kok, R., et al. 2008, JQSRT, 109, 1136

Kalirai, J. 2018, ConPh, 59, 251

Kaltenegger, L., & Traub, W. A. 2009, ApJ, 698, 519

Kass, R. E., & Raftery, A. E. 1995, J. Am. Stat. Assoc., 90, 773

Kempton, E. M.-R., Bean, J. L., Louie, D. R., et al. 2018, PASP, 130, 114401

Kempton, E. M.-R., Lupu, R., Owusu-Asare, A., Slough, P., & Cale, B. 2017, PASP, 129, 044402

Kreidberg, L., Bean, J. L., & Désert, J.-M. 2014, Natur, 505, 69

Krissansen-Totton, J., Garland, R., Irwin, P., & Catling, D. C. 2018, AJ, 156, 114

Line, M. R., Stevenson, K. B., Bean, J., et al. 2016, AJ, 152, 203

Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, ApJ, 775, 137

Louie, D. R., Deming, D., Albert, L., et al. 2018, PASP, 130, 044401

Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, AJ, 158, 27

Madhusudhan, N. 2018, Handbook of Exoplanets (Berlin: Springer), 2153

Madhusudhan, N. 2019, ARA&A, 57, 617

Malik, M., Grosheintz, L., Mendonça, J. M., et al. 2017, AJ, 153, 56

Miller-Ricci, E., & Fortney, J. J. 2010, ApJL, 716, L74

Miller-Ricci, E., Seager, S., & Sasselov, D. 2009, ApJ, 690, 1056

Molliere, P., Wardenier, J. P., van Boekel, R., et al. 2019, A&A, 627, A67

Morley, C. V., Kreidberg, L., Rustamkulov, Z., Robinson, T., & Fortney, J. J. 2017, ApJ, 850, 121

Muirhead, P. S., Dressing, C. D., Mann, A. W., et al. 2018, AJ, 155, 180

Oliphant, T. E. 2015, Guide to NumPy (2nd ed.; USA: CreateSpace Independent Publishing Platform)

Pickering, T. E., Pontoppidan, K. M., Laidler, V. G., et al. 2016, Proc. SPIE, 9910, 381

Roberge, A., Bolcar, M. R., & France, K. C. 2019a, Proc. SPIE, 11115, 111150O

Roberge, A., Fischer, D., & Peterson, B. 2019b, BAAS, 51, 199

Stamnes, K., Tsay, S.-C., Wiscombe, W., & Jayaweera, K. 1988, ApOpt, 27, 2502

Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, ApJ, 809, 77

The Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 3

The LUVOIR Team 2019, arXiv:1912.06219

Tremblin, P., Amundsen, D. S., Mourier, P., et al. 2015, ApJL, 804, L17

Virtanen, P., Gommers, R., & Oliphant, T. E. 2019, arXiv:1907.10121

Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, ApJ, 802, 107

Wall, J. V., & Jenkins, C. R. 2012, Practical Statistics for Astronomers (Cambridge Observing Handbooks for Research Astronomers) (Cambridge: Cambridge Univ. Press)

Winn, J. N. 2010, arXiv:1001.2010

Zhang, M., Chachan, Y., Kempton, E. M. R., & Knutson, H. A. 2018, PASP, 131, 034501