

UC Riverside

UC Riverside Previously Published Works

Title

Modeling Atmospheric Lines by the Exoplanet Community (MALBEC) Version 1.0: A CUISINES Radiative Transfer Intercomparison Project

Permalink

<https://escholarship.org/uc/item/5bd558z2>

Journal

The Planetary Science Journal, 5(3)

ISSN

2632-3338

Authors

Villanueva, Geronimo L

Faucher, Thomas J

Kofman, Vincent

et al.

Publication Date

2024-03-01

DOI

10.3847/psj/ad2681



















Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Modeling Atmospheric Lines by the Exoplanet Community (MALBEC) Version 1.0: A CUISINES Radiative Transfer Intercomparison Project

Geronimo L. Villanueva¹ , Thomas J. Fauchez^{1,2,3} , Vincent Kofman^{1,2} , Eleonora Alei^{4,5,17} , Elspeth K. H. Lee⁶ , Estelle Janin⁷ , Michael D. Himes^{8,9} , Jérémy Leconte¹⁰ , Michaela Leung¹¹ , Sara Faggi^{1,2} , Mei Ting Mak¹² , Denis E. Sergeev¹² , Thea Kozakis¹³ , James Manners¹⁴ , Nathan Mayne¹² , Edward W. Schwieterman¹¹ , Alex R. Howe^{1,15} , and Natasha Batalha¹⁶ 

¹ NASA Goddard Space Flight Center, Greenbelt, MD, USA; geronimo.l.villanueva@nasa.gov

² Integrated Space Science and Technology Institute, Department of Physics, American University, Washington DC, USA

³ NASA GSFC Sellers Exoplanet Environments Collaboration, Greenbelt, MD, USA

⁴ ETH Zurich, Institute for Particle Physics & Astrophysics, Wolfgang-Pauli-Str. 27, 8093 Zurich, Switzerland

⁵ National Center of Competence in Research PlanetS, Switzerland

⁶ Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, CH-3012 Bern, Switzerland

⁷ School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85281, USA

⁸ Planetary Sciences Group, Department of Physics, University of Central Florida, USA

⁹ NASA Postdoctoral Program Fellow, NASA Goddard Space Flight Center, Greenbelt, MD, USA

¹⁰ Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France

¹¹ Department of Earth and Planetary Sciences, University of California, Riverside, CA, 92521, USA

¹² Physics and Astronomy, University of Exeter, Exeter EX4 4QL, UK

¹³ National Space Institute, Technical University of Denmark, Elektrovej 328, 2800 Kgs. Lyngby, Denmark

¹⁴ Met Office, FitzRoy Road, Exeter EX1 3PB, UK

¹⁵ Center for Research and Exploration in Space Science and Technology, NASA/GSFC, Greenbelt, MD 20771, USA

¹⁶ NASA Ames Research Center, Mountain View, CA 94035, USA

Received 2023 February 21; revised 2023 December 26; accepted 2024 January 17; published 2024 March 8

Abstract

Radiative transfer (RT) models are critical in the interpretation of exoplanetary spectra, in simulating exoplanet climates, and when designing the specifications of future flagship observatories. However, most models differ in methodologies and input data, which can lead to significantly different spectra. In this paper, we present the experimental protocol of the Modeling Atmospheric Lines By the Exoplanet Community (MALBEC) project. MALBEC is an exoplanet model intercomparison project that belongs to the Climates Using Interactive Suites of Intercomparisons Nested for Exoplanet Studies framework, which aims to provide the exoplanet community with a large and diverse set of comparison and validation of models. The proposed protocol tests include a large set of initial participating RT models, a broad range of atmospheres (from hot Jupiters to temperate terrestrials), and several observation geometries, which would allow us to quantify and compare the differences between different RT models used by the exoplanetary community. Two types of tests are proposed: transit spectroscopy and direct imaging modeling, with results from the proposed tests to be published in dedicated follow-up papers. To encourage the community to join this comparison effort and as an example, we present simulation results for one specific transit case (GJ-1214 b), in which we find notable differences in how the various codes handle the discretization of the atmospheres (e.g., sub-layering), the treatment of molecular opacities (e.g., correlated- k , line-by-line) and the default spectroscopic repositories generally used by each model (e.g., HITRAN, HITEMP, ExoMol).

Unified Astronomy Thesaurus concepts: [Radiative transfer simulations \(1967\)](#); [Exoplanet atmospheres \(487\)](#)

1. Introduction

Model intercomparison projects (MIPs) are the ideal tool to benchmark models and to track down numerical/algorithm issues and model dependencies. MIPs are widely used in the Earth Science community, with one of the most known MIPs being the Coupled Model Intercomparison Project (CMIP), currently in its sixth version (Eyring et al. 2016), which compares General Circulation Models (GCMs) in the simulation of Earth's past and current climate and the prediction of its future climate. Such model intercomparisons, which have been

running for almost three decades, are playing a crucial role in assessing global climate change, predicted by all models, and in forecasting possible future climate scenarios.

In Earth Sciences, many radiative transfer (RT) tool intercomparisons have been performed, such as the RAdiative transfer Model Intercomparison (RAMI; Pinty et al. 2001), whose activity focuses on the benchmarking of Earth forest canopy RT models, or more recently the work by Zawada et al. (2021), which compared seven different limb-scattering models. Another notable Earth Science RT MIP is the Correlated K -Distribution Model Intercomparison Project (CKDMIP) by Hogan & Matricardi (2020). This MIP benchmarks line-by-line calculations to evaluate the accuracy of existing correlated- k distribution (CKD) models, explores how accuracy varies with the number of pseudomonochromatic calculations, assesses how different choices made when CKD models are generated affect their accuracy, and generates open-

¹⁷ <https://nccr-planets.ch>



access data sets and software, enabling the development of new gas-optics tools. To date, results from this intercomparison are still being evaluated.

To our knowledge, one of the first RT MIPs to be proposed for the paleoclimate and exoplanet community was the Palaeoclimate and Terrestrial Exoplanet Radiative Transfer Model Intercomparison Project (PALAEOTRIP, Goldblatt et al. 2017). No results have been reported from this MIP yet. A few years later, Barstow et al. 2020 compared three radiative transfer codes, TauREx 2 (Waldmann et al. 2015b, 2015), NEMESIS (Irwin et al. 2008), and CHIMERA (Mai & Line 2019, and reference therein) for hot Jupiter and warm Neptune cases. Good agreements within 20–40 ppm of residual noise were found for the forward models, while cross-retrievals were shown to be consistent within 1σ . In the same year, Pincus et al. (2020) developed the Radiative Forcing Model Intercomparison Project (RFMIP) that uses benchmark calculations made with line-by-line models to identify parameterization errors in the representation of absorption and emission by greenhouse gases. Among the main findings is that agreement between line-by-line models is better in the longwave than in the shortwave, where various treatments of the water vapor continuum impact estimates of forcing by carbon dioxide and methane. Other intercomparisons (Yang et al. 2016; Baudino et al. 2017) have also revealed the importance of benchmarking radiative transfer and identified notable differences between resulting spectra.

More recently, Barstow et al. (2022) developed a retrieval challenge to prepare for the ARIEL mission with five different exoplanet retrieval codes (ARCIS; Min et al. 2020), NEMESIS (Irwin et al. 2008), Pyrat Bay (Cubillos & Blečić 2021), TauREx 3 (Al-Refaie et al. 2021) and POSEIDON (MacDonald & Madhusudhan 2017) analyzing the same synthetic data set of a hot Jupiter (clear and cloudy), a clear warm Neptune and a cloudy super-Earth. Very good agreements among them have been found, with the majority of the cases correctly retrieved. Harrington et al. (2022) presented the Bayesian Atmospheric Radiative Transfer (BART) retrieval code alongside a MIP framework called BARTTest, which included a detailed comparison with NEMESIS, CHIMERA, and TauREx from the Barstow et al. (2020) work. As part of the JWST Early Release Science (ERS) exoplanetary observations, the Transiting Exoplanet Community also explored and compared several RT models (Ahler et al. 2022).

A typical challenge of model intercomparisons is that, as the results converge and become more homogeneous between them, we may reach a situation of high precision (or rather, high robustness) but low accuracy. In that aspect, it is very important to employ models that originate and have foundations from different communities (e.g., Earth remote sensing, solar system astronomy, astrophysics), where the core elements of the different RT algorithms have been validated and consolidated by comparing to a broad range of data sets. As such, the general philosophy of the Climates Using Interactive Suites of Intercomparisons Nested for Exoplanet Studies (CUISINES) framework (and therefore of MALBEC) is to achieve high robustness between models while running them using their default configurations and explore why the results differ and how each model could be improved to reach higher accuracy. The new exoplanet MIP (exoMIP) presented here is called Modeling Atmospheric Lines By the Exoplanet Community (MALBEC). In MALBEC, we intend to go beyond

these previous RT model intercomparisons by including a larger set of models (\sim a dozen), a larger and more diverse set of exoplanets (from hot Jupiters to temperate terrestrials), and several observation geometries (transit spectroscopy, emission spectroscopy, and direct imaging). Each model scenario is called a “cask” within the MALBEC protocol. Currently, only forward modeling test cases are considered in MALBEC, yet we plan to integrate retrievals and noise modeling in future investigations. Comparing simulated spectra across models is a first step, and that allows us to determine notable spectroscopic differences, but identifying parametric physical or statistical distances between methods would ultimately require retrieval analyses that employ different forward models. Also to note is that none of these codes are radiative equilibrium (RE) or radiative convective equilibrium (RCE) models, and the here proposed RT tests operate with prescribed temperature, pressure, and chemical structures as input. RE and RCE models will be part of another exoMIP called “COD ACCRA” for Comparing One-Dimensional Climates of Convective Radiative Atmospheres available within the CUISINES framework.

1.1. Synergy with Other Model Intercomparison Projects

MALBEC is part of the CUISINES model intercomparison work group. CUISINES stands for Climates Using Interactive Suites of Intercomparisons Nested for Exoplanet Studies and provides a framework for different model intercomparison projects, as well as opportunities to find synergies (*pairings*) between different work groups. It currently consists of nine different projects, including chemical, circulation, energy balance, and planet-specific models.¹⁸ The framework started with the TRAPPIST-1 Habitable Atmospheres Intercomparison (THAI) project, which presented a trilogy of GCM studies investigating dry and wet climate for TRAPPIST-1e (Fauchez et al. 2022; Sergeev et al. 2022; Turbet et al. 2022). CUISINES was inspired by the THAI success and led to two international workshops and eight more exoMIPs at the time of writing. Three other exoMIPs—SAMOSA, CAMEMBER, and FILLET—have currently published protocol papers describing their efforts:

1. The Sparse Atmospheric Modeling Sample Analysis (SAMOSA) aims to provide guidance for what type of planets need to be modeled using full 3D general circulation models and where more simple 1D equivalent models are sufficient (Haqq-Misra et al. 2022). The parameters investigated are insolation and surface pressure, and it compares the outcome scenarios between the ExoCAM 3D GCM and ExoPlaSim 1D models (Paradise et al. 2022; Wolf et al. 2022).
2. Comparing Atmospheric Models of Extrasolar Mini-Neptunes Building and Envisioning Retrievals and Transits (CAMEMBER) focuses on intercomparing the results from eight different GCMs simulating the atmospheres of GJ1214b and K2-18b (Christie et al. 2022).
3. Functionality of Ice Line Latitudinal EBM Tenacity (FILLET) studies energy balance models, and in particular looks at climate “end” states using different

¹⁸ For more information, see <https://nexss.info/cuisines/>.

values for isolation, obliquity, starting temperatures, and atmospheric composition (Deitrick et al. 2023).

MALBEC serves several roles within the CUISINES framework: as a downstream framework to produce spectra and simulate observables from the 3D GCM exoMIPs (THAI, CAMEMBER, MOCHA), constrain the detectability of outputs from photochemical codes (PIE), and verify the radiative fluxes from energy balance models (FILLET). Central in all of these is the MALBEC configuration file (see below), which provides a standardized and accessible format for both input and output. Namely, “upstream” models that generate profiles of the atmospheric variables (e.g., GCMs, RCE models) can produce output files in the MALBEC format that can then be read by the “downstream” radiative transfer codes. This standardized file provides a bridge for testing the outputs of different models and is how we ensure the test cases in this protocol paper are interpreted similarly by all codes. Additionally, MALBEC can provide two metrics to other exoMIPs: the detectability of recorded differences between upstream codes and the spread or uncertainty in the downstream codes.

The objective of this paper is to present the MALBEC experimental protocol that will be used to compare 10 already participating models. A clear and descriptive protocol, separated from the result papers, as well as broad participation, is essential for a meaningful model intercomparison project. This paper is structured as follows. In Section 2, we outline the motivation for the MALBEC exoMIP, while Section 3 describes the physical and computational considerations that came across during the development of the study, as well as the input/output files. In Section 4 we describe the models currently on board MALBEC, and the experimental protocol is addressed in Section 5. The MALBEC repository, the output formats, standards, and archiving are discussed in Section 6. The summary is given in Section 7.

2. Motivation for the Model Intercomparison Project

RT models are used for a large variety of planetary science applications, from solar system objects (Villanueva et al. 2018, 2022) to exoplanets (Waldmann et al. 2015; Arney et al. 2017; Caldas et al. 2019; Fauchez et al. 2019; Lustig-Yaeger et al. 2019; Wunderlich et al. 2020; Konrad et al. 2022, and references therein). For exoplanet applications, the number of RT codes has been rapidly rising over the years, following the sharp increase in the number and diversity of planets discovered, alongside the development of various ground-based and space-based observatories either dedicated or significantly oriented toward exoplanet detection and characterization. In this context, it is essential to ensure that RT models are producing similar results to not potentially bias both the data prediction and interpretation.

A detailed model intercomparison study has many benefits both to users and developers of the code and to the broader exoplanetary community. The benefits of such a study are twofold: on the one hand, it helps to make an assessment of the differences in model predictions and potentially attach an uncertainty to the simulations, and on the other hand, it can help the different RT codes to be more consistent across the community, highlighting the most important functionalities, asserting what elements would benefit from further development, and overall improving every model fidelity.

By allowing each RT code to work with its built-in functions and specific engine, i.e., without constraining every single parameter nor forcing a particular analysis pipeline (see further sections for precision about the inputs), this intercomparison provides an objective image about each code’s specifics and characteristics. Note that this strategy is different from the one adopted by Barstow et al. (2020), where each model had to use exactly the same setup, and only the engines themselves were compared. MALBEC aims to identify the sources of disagreement among the participant codes, quantify the differences, and trigger discussions in a constructive environment toward improving the respective performances. As there are no ground truths in these types of simulations, understanding differences between codes can provide much insight and promote the unification of certain schemes and methods that increase RT fidelity. Importantly, the framework developed in MALBEC can be applied by any user who would want to work with one of these models as a benchmarked example, establishing a starting reference point.

3. Physical and Computational Considerations

The different radiative transfer models participating in MALBEC inevitably have different strategies to capture the physical phenomena, different emphases (e.g., direct imaging, transit, gas giants, Earth-like planets), as well as different computational descriptions of the RT phenomena. In the first part of this section, we discuss the physical processes underlying the RT calculations, with a particular focus on the processes where the codes differed, where they had to be aligned, or where the input standards were notably different. Subsequently, we came across a number of computational approaches that were substantially different for some of the participating codes (e.g., atmospheric layering and pressure/temperature definitions), which will be covered in Section 3.1. In some cases, small differences in the assumed constants (i.e., the size of the Sun, the gravitational constant) resulted in large differences in transit depths. In order to avoid discrepancies in constants, all simulation parameters defined in the MALBEC files are in absolute standard units.

3.1. Spectroscopy and Physical Phenomena

Throughout the wavelength range under investigation in MALBEC, many processes affect the radiation propagating through an atmosphere. Table 1 highlights these and in which spectral regions they are most relevant. Spectroscopic linelists are at the core of radiative transfer calculations, and in many cases, they are the dominant source of discrepancy between models. Considerations regarding the inclusion of isotopes, high-energy lines for warm atmospheres, lineshape models, and wing cutoffs can have substantial effects on the computed spectra. In addition, how these spectroscopic linelists are integrated into the RT model, such as line-by-line, cross sections, and correlated- k methods, can lead to different levels of fidelity.

Based on the spectroscopic parameters (including strength and energies) of the different rovibrational molecular transitions, the absorption of a particular gas can be determined for a temperature-pressure combination. The absorption lineshapes also depend on the properties of the dominant collisional partner in the atmosphere (e.g., air, CO₂, H₂, He). Different databases are available, with HITRAN (Gordon et al. 2022)

Table 1
Main Processes Impacting the Simulated Spectra of Planets and Investigated by This Investigation

Spectroscopic	Affected Region	Notes
Rayleigh Scattering	<1.5 μm	Different implementations exist
Molecular absorption:	All wavelengths	
1. Isotopes	Specific features and deuterium molar masses	e.g., HDO, CO ₂ isotopologues
2. Line-by-line	All wavelengths	Main sources, HITRAN, HITEMP, ExoMol
3. Correlated- <i>k</i>	All wavelengths	Each model implements these differently
4. Cross sections	Critical at UV/Optical	Complement linelists with cross sections
Collision-induced Absorption	Continuous	Main sources: HITRAN, MT_CKD, theory

being one of the most comprehensive databases, containing 55 molecules and 148 isotopologues at different collisional regimes, and about 300 in experimental cross sections. The focus of HITRAN is Earth’s atmospheric conditions, and for that reason molecular states and transitions that are relevant at different atmospheric temperatures and pressures are not always included. For higher temperatures, the HITRAN team created the HITEMP database (Rothman et al. 2010), which contains spectroscopic information about other H₂O, CO₂, CO (Li et al. 2015), N₂O (Hargreaves et al. 2019) and CH₄ (Hargreaves et al. 2020) molecules, particularly relevant for exoplanet atmospheres. Additionally, the ExoMol repository (Tennyson et al. 2016; Chubb et al. 2021) is specifically focused on exoplanets, brown dwarfs, and cool stars and hosts a large spectroscopic database of molecules, which are particularly suited for high temperature and non-LTE regimes. Specialized databases that combine several repositories also exist, such as HELIOS-K (Grimm et al. 2021).^{19,20,21}

RT codes have different ways of handling linelists, and lineshape models—including different assumptions regarding line wing cutoffs—can have a substantial impact on the computed spectra. The most direct method is to calculate the absorption by line-by-line shape modeling, which is very accurate but computationally expensive, and it is not always possible in the UV/optical since linelists typically do not accurately cover this region. Some RT codes rely on UV/optical cross section tables to complement the line-by-line/correlated-*k* tables. The MPI-Mainz provides extensive cross section data for many species in this spectral range (Keller-Rudek et al. 2013). For low-resolution studies (i.e., resolving powers less than 5000), the absorbance is often saved in correlated-*k* tables, which capture the distribution of absorbance in a spectroscopic bin in a selected number of points. Correlated-*k* tables are precalculated for a grid of temperature and pressure combinations and interpolated upon running RT calculations. Alternatively, cross section data can be used as well, which are precalculated absorption spectra for particular temperature-pressure conditions, but the RT calculation needs to be performed at very high spectral resolution to ensure high fidelity (Garland & Irwin 2019). Isotopes are available both in HITRAN and ExoMol. Even when using the same databases/methods, the inclusion and treatment of isotopologues among the RT codes vary greatly, mainly due to the assumed linelist and the considered isotopic abundances. For most isotopes, spectral shifts are minor and are typically lost under the main

isotope’s absorbance features (e.g., ¹³CH₄), but in some cases, the absorbance of the heavier isotopes is in absorbance minima of the main isotope and can thus be important to consider (e.g., HDO).

At short wavelengths, Rayleigh Scattering has a strong wavelength dependency, resulting in a smooth increase from the near-IR, visible to UV, following λ^{-4} . The magnitude depends on the polarizability of the atoms/molecules, but in principle, all (dominant) gases in an atmosphere scatter light in this way. The formulation used in RT codes is often the one described in Sneep & Ubachs (2005). The tabulated values are available from NIST,²² yet different codes implement this process in slightly different ways. In general we have encountered relatively small differences in this process across the “casks” compared here.

Primarily in the infrared, but also in the UV and in the visible, collision-induced absorption (CIA) arises from temporary bimolecular complexes, resulting in relatively broad features whose intensities scale quadratically with pressure. A description of the phenomena and how to calculate the corresponding absorption cross sections can be found in Karman et al. (2019). The HITRAN CIA repository²³ lists the most generally used CIA lists. Depending on the main emphasis of each RT model, the number of CIA databases and considered repositories varies greatly. RT models with an emphasis on Earth-like planets implement CIA lists well beyond HITRAN, including water continuum and capturing several N₂-x and O₂-X CIA bands, while gas-giant RT models mainly consider several theoretical and lab-based CIA databases for H₂-X and He-X pairs.

3.2. Computational Considerations, Definitions of Terminologies, and Input/Output Specifications

During the development of this protocol, we investigated a number of different treatments of the computational aspects of the radiative transfer calculations to understand the fundamental differences between codes. To ensure all models interpreted the different tests “(casks)” similarly, a common input file was agreed upon. This configuration file, called the MALBEC file, contains all the information required to calculate the different test cases in a consistent format across the different “casks” and can be ingested by each RT code within MALBEC.

Not all radiative transfer codes discretize the atmosphere in the same way, so a method was developed to homogenize the layering schemes (pressure and temperature versus height) for all RT codes. In addition to providing anchor points at specific

¹⁹ <https://dace.unige.ch/opacity> Planetary Spectrum Generator (PSG) databases (Villanueva et al. 2022).

²⁰ <https://psg.gsfc.nasa.gov/helptm.php>, or MAESTRO (Batalha & Rooney 2020).

²¹ <https://science.data.nasa.gov/opacities/app>

²² <https://cccbdb.nist.gov/pollistx.asp>

²³ <https://hitran.org/cia/>

Table 2
Radiative Transfer Models Participating in MALBEC and Their Point-of-contact (POC) and Notable Model References

Code	POC	References
APOLLO	Alex Howe	Howe et al. (2022)
BART	Michael D. Himes	Cubillos et al. (2022), Harrington et al. (2022)
...	...	Blecic et al. (2022)
gCMCRT	Elsbeth K.H. Lee	Lee et al. (2019, 2022)
petitRADTRANS	Eleonora Alei	Mollière et al. (2019)
PSG/PUMAS	Geronimo L. Villanueva	Villanueva et al. (2018, 2022)
Pytmosph3r	Jeremy Leconte	Caldas et al. (2019)
SMART	Michaela Leung, Samantha Gilbert, Gabrielle Suissa, Edward Schwieterman	Meadows & Crisp (1996)
SOCRATES	Mei Ting Mak, James Manners, Denis Sergeev, Nathan Mayne	Lines et al. (2018)
TauREx 3	Estelle Janin	Waldmann et al. (2015b, 2015), Al-Refai et al. (2021)
PICASO	Thea Kozakis	Batalha et al. (2019, 2021)

altitudes to describe the local atmospheric pressure, temperature, and molecular/aerosol abundances, we also provide polynomial fits to these quantities so that each code can discretize the atmosphere in any preferred manner. This is relevant both for the radiative transfer calculations in each layer and for the path lengths of light through each layer (i.e., ray tracing). Particularly for ray-tracing calculations in spherical refractive atmospheres, it is important to sample at sufficiently high resolutions, either by having enough layers or by creating sublayers. If insufficient layers are used, this may lead to large differences. In the configuration file, we provide “anchor” values, defining the specific pressure, temperature, and abundances at specific points in the atmosphere. These points quantify the properties of the atmosphere at that exact pressure point, yet when performing ray-tracing calculations to determine the path length of individual layers, each model may subdivide the light path between anchor points into much finer points by using a sublayering scheme in order to ensure an accurate sampling of the large variability of the parameters in-between anchor points. Additionally, we also provide analytical fits (third order versus \log_{10} pressure [bar]) to those profiles, so the user can create their own vertical discretization, ensuring accurate results with their respective model.

The configuration file defines the following:

1. *Planetary and stellar parameters*: planetary radius [m], surface gravity [m s^{-2}], the phase at which the planet is observed, planet temperature [K], albedo, emissivity, star-planet distance [m], stellar type, stellar temperature [K] and stellar radius [m]. Albedo and emissivity are defined to be constant across the simulation wavelength range, and the surface scattering model is assumed to follow the Lambert diffuse principle.
2. *Opacity sources used*: linelists, CIAs, Rayleigh, etc.
3. *Atmosphere columns*: this entry specifies the variables included in the simulation and the anchor points. Pressures are in [bar], temperatures in [K], mean molecular weight in [g mol^{-1}], height in [km], and molecular abundances in [mol mol^{-1}].
4. *Layers*: the specifics of each layer, containing the pressure, temperature, mean molecular weight, altitude, and molecular abundances following the definitions described above.
5. *Analytical functions layering*: for a number of relevant species, third-order polynomial parameters describing the abundance with the \log_{10} of the pressure are given.
6. *Bottom of the atmosphere (BOA)*: the atmosphere is defined to start at the first anchor point, with surface properties as listed in the configuration file (surface temperature, albedo, emissivity). The planetary gravity and radius refer to this point.
7. *Top of the atmosphere (TOA)*: the atmosphere is defined to end at the last anchor point, and therefore the atmospheric radiative transfer is only computed between BOA and TOA.
8. *Spectral range and sampling*: four parameters define the range and resolution of each simulation. The minimum wavelength defines the center wavelength of the first “pixel,” while the maximum wavelength defines that of the last “pixel.” The separation between pixels can be a fixed wavelength if the unit is “ μm ,” while if the unit is “RP,” the separation assumes a constant resolving power ($\frac{\lambda}{\delta\lambda}$).
9. *Output units*: for transit simulations, the output model units are in $(\frac{R_{\text{planet}}}{R_{\text{star}}})^2$ while for direct-imaging simulations, the model outputs are in [$\text{W sr}^{-1} \text{m}^{-2} \mu\text{m}^{-1}$].
10. *Opacities*: the models should employ all opacity terms suitable for each test case; for instance, high-energy linelists would be beneficial when simulating hot atmospheres, as in case T3B (see below). This includes molecular opacity (for all isotopes), Rayleigh scattering/extinction, water continuum (e.g., MT_CKD), and CIA. Molecular opacities should be generated employing the most appropriate linelist available to each model for that collisional regime (e.g., H_2 , air broadening). All simulations as part of this intercomparison are at moderate resolutions (with resolving powers equal to or lower than 500), and therefore correlated- k or other precomputing methods are encouraged.

4. Description of the Models Participating in MALBEC

The MALBEC intercomparison is open to the whole community, with the framework of this initiative being defined among a group of diverse radiative-transfer teams and models (Table 2). The RT models are quite different in their scope and functionality, which is important when establishing and comparing results across a broad range of regimes, as required by the exoplanetary community. Importantly, not all models are required to address all intercomparison experiments, and in this

section we list the capabilities and characteristics of each of these initial RT models.

4.1. APOLLO

APOLLO is a 1D radiative transfer and Markov Chain Monte Carlo (MCMC) spectroscopic retrieval code designed for modularity and internal model intercomparisons. The main description of APOLLO can be found in Howe et al. (2022), and the model can be used to synthesize both transit and emission spectra. It is built on a philosophy of free retrieval that is agnostic about specific microphysics and as such allows the user to select from several options for prescriptions for temperature-pressure profiles, cloud structures, and molecular inventories. While multiple radiative transfer prescriptions are also available, by default, APOLLO simulates radiative transfer (reflected and emission) through an atmosphere using the two-stream approximation (Toon et al. 1989) with hemispheric closure (Meador & Weaver 1980). APOLLO uses the Freedman et al. (2014) cross section tables on a species-by-species basis. A weighted sum of the cross sections for each species is then computed based on the mixing ratios to compute optical depths layer by layer, including CIAs and aerosols. For transit spectroscopy, the starlight is assumed to be parallel. For emission spectra, the emitted light is integrated over an eight-point Gaussian quadrature. APOLLO can perform retrievals by an MCMC method using the `emcee` Python package (Foreman-Mackey et al. 2013). APOLLO also includes an “ensemble” method to rapidly generate and evaluate regularly spaced forward model grids.

4.2. BART

BART (Blecic et al. 2022; Cubillos et al. 2022; Harrington et al. 2022) is an open-source exoplanet retrieval code that pairs the `transit` RT code (Rojo 2006) with the Multi-Core Markov Chain Monte Carlo (MC3; Cubillos et al. 2017) Bayesian framework. BART performs 1D radiative transfer under the assumption of hydrostatic balance and LTE. For transmission spectra it performs ray tracing, and for emission spectra it assumes a plane-parallel atmosphere. Its programmatic design focuses on flexibility, enabling BART to replicate the setups of other codes and achieve agreement in results (e.g., Himes & Harrington 2022). The RT routines are implemented in C for speed, while BART is implemented in Python for ease of use and portability. Both `transit` and BART are generally executed via configuration files, where the user may specify the various parameters of the run, including the spectral range and resolution, atmospheric species and profiles, Rayleigh extinction and/or cloud models, and the observing geometry (transit, eclipse, or direct observations). At present, BART does not consider multiple scattering when computing emission or transmission spectra. `Transit`'s package to ingest linelists allows for HITRAN/HITEMP (Rothman et al. 2010; Hargreaves et al. 2020; Gordon et al. 2022), ExoMol (Tennyson et al. 2016; Chubb et al. 2021; via Repack, Cubillos 2017), TiO (Schwenke 1998), and the H₂O list of Partridge & Schwenke (1997). `Transit` utilizes a line-by-line opacity sampling approach to reduce the sizes of extensive linelists while still retaining an accurate RT calculation (see Rojo 2006; Harrington et al. 2022, and Cubillos et al. 2022 for more details on other computational cost-saving measures). In Figure 2 of test cask T3B, we utilize the HITEMP linelists for H₂O and CH₄;

HITRAN CIA databases for H₂-He, H₂O-H₂O, CH₄-CH₄, CH₄-He, and H₂-CH₄; and the H₂-H₂ CIA database of Borysow et al. (2001); Borysow (2002). During a retrieval, BART calculates high-resolution spectra and bins them according to the instruments used to observe the exoplanet. MC3 calculates the effective sample size of the posterior resulting from the inference as well as its induced uncertainty in credible regions determined from that posterior (see Harrington et al. 2022 for more details).

4.3. gCMCRT

gCMCRT is an open-source 3D spherical geometry hybrid ray tracing and Monte Carlo radiative-transfer code that uses GPUs to accelerate calculations. The main descriptions of gCMCRT can be found in Lee et al. (2019) and Lee et al. (2022). gCMCRT is primarily aimed at accurately postprocessing 3D GCM data, especially when multiple scattering from cloud particles is important. gCMCRT directly simulates the path of photon “packets” (usually $\sim 10^6$) through the prescribed 3D spherical geometry structure. This is a subgrid model, with packets traversing through the 3D volume of each cell. For transmission spectroscopy, packets are fired at the dayside of the planet and traced through the transmission limbs, which are then used to estimate the transmission spectra integral through random sampling. For emission spectra, packets originate in the 3D spherical volume itself and are traced toward a certain viewing angle (planetary phase). The fraction of energy escaping toward this direction is then tracked and used to calculate the emission spectra. For albedo calculations, the scattered fraction of the incoming starlight is calculated toward a certain viewing direction. gCMCRT can also perform multiple scattering in 3D geometry by sampling scattering phase functions (for example, the Rayleigh or Henyey-Greenstein phase functions) when a packet undergoes a scattering event. gCMCRT contains a well-tested opacity mixer (`optools`), which can mix and prepare gas phase abundances, correlated- k tables, line-by-line, CIA, Rayleigh, and cloud opacities and scattering properties. gCMCRT is able to perform high-resolution spectral modeling, including Doppler shifting of lines, allowing simulated GCM data to be used as part of high-resolution cross-correlation studies.

4.4. petitRADTRANS

petitRADTRANS (pRT for short) is a publicly available package to calculate transmission and emission spectra of clear and cloudy exoplanets (see Mollière et al. 2019, 2020, for more details). The core routines are written in FORTRAN and wrapped in a Python package. The model can run at high resolution ($\lambda/\Delta\lambda \leq 10^6$) using a line-by-line treatment, as well as at low resolution ($\lambda/\Delta\lambda \leq 10^3$) using a correlated- k approach. Opacities can be rebinned at any resolution needed by the user, either directly in petitRADTRANS (for high resolution) or through the `exo-k` package (Leconte 2021), which is called internally (for low resolution). The default set of opacities provided by the maintainers includes opacities from HITRAN/HITEMP and ExoMol, generally spanning a temperature range between 80 and 3000 K and a pressure range between 10^{-6} and 10^3 bar. Further opacity tables can be downloaded from ExoMol directly in a pRT-friendly format or calculated by the user through the ancillary scripts provided.

petitRADTRANS allows the user to consider different cloud treatments, such as gray clouds, wavelength-dependent power laws, or using opacity (in $\text{cm}^2 \text{g}^{-1}$) of real condensates of various compositions and shapes. Arbitrary opacity functions and particle setups can also be employed. Users can also provide low-res (pseudocontinuum) opacities to petitRADTRANS to be included in the calculation, thus customizing the cloud treatment. petitRADTRANS includes the calculation of Rayleigh scattering and collision-induced absorption. It was recently updated to include the surface direct and thermal scattering for rocky exoplanets (see Appendix A in Alei et al. 2022, for details). Specifically for the computation of multiple scattering of reflected starlight and thermal emission, pRT uses the Feautrier method, which is a third-order method that allows the treatment of the radiative transfer equation in the diffusive regime and for which pRT iterates to solve for the multiple-scattering source by employing the ALI and Ng acceleration. Users can specify wavelength-dependent surface reflectivity and emissivity, allowing the scattering treatment of various surface compositions of rocky exoplanets. In its latest version, it also allows the user to run retrievals using the nested sampling approach. Thanks to the MALBEC collaboration, opacities of additional isotopes for many atmospheric absorbers (e.g., CO_2 , CO , H_2O) were computed, as well as the collision-induced absorption tables for water (MT_CKD, Mlawer et al. 2012; Kofman & Villanueva 2021). These will soon be included in the main opacity repository of petitRADTRANS. For the purpose of this intercomparison, the version of petitRADTRANS was frozen to 2.6, though the package is in the process of being upgraded to version 3, with additional features and quality-of-life improvements.

4.5. Planetary Spectrum Generator

The Planetary Spectrum Generator (PSG; Villanueva et al. 2018, 2022) is a publicly available radiative-transfer suite, generalized to any type of planetary atmosphere/surface across a broad range of wavelengths (50 nm to 100 μm , UV/Vis/near-IR/IR/far-IR/radio) and observational geometry (e.g., transit, direct imaging, in situ, orbiter, limb). The suite includes a 3D orbital calculator for most bodies in the solar system and all confirmed exoplanets, and also includes a broad range of atmospheric and chemical models. The ray-tracing integration across the atmosphere is performed in PSG via a sublayering spherical/refractive iterative algorithm, enabling high-fidelity results even when operating with a limited number of layers. Several radiative-transfer modules operate within PSG, including the Cometary Emission Model (CEM), a non-LTE exospheres solver, and PSGDORT, the multiple-scattering radiative-transfer module of PSG that employs the discrete ordinate method to model atmosphere/surface scattering processes. PSGDORT is based on DISORT v2.1 (Stamnes et al. 2000), which was adapted for non-LTE and optimized to operate with a variety of spectral grids (e.g., line-by-line, correlated- k , surface scattering grids) as employed by the PSG radiative transfer algorithm. PSGDORT also includes a correction for pseudospherical geometry, as described by Dahlback & Stamnes (1991). The PSG radiative transfer modules take into account Rayleigh and Raman scattering, refraction, non-LTE processes, molecular/atomic absorptions, and collision-induced absorption processes. Rovibrational absorption is calculated line by line for resolving powers

higher than 5000, and correlated- k tables are used for lower resolutions.

PSG operates with several databases, including the latest HITRAN/HITEMP distribution (linelist, CIAs, cross sections), GEISA, and ExoMol. For the PSG simulation results presented in Figure 2 of cask T3B, we utilized correlated- k tables based on the HITEMP linelists of H_2O and CH_4 . PSG complements the UV/optical range by integrating UV/optical cross sections for dozens of species collected from a range of spectral databases. Most of the UV cross sections originate from the MPI-Mainz Spectral Atlas (Keller-Rudek et al. 2013), which have been parsed, combined, and formatted to provide a comprehensive coverage of the 0.01–1 μm wavelength range. Additional cross sections include those of O_3 by Serdyuchenko et al. (2014), CO_2 by Venot et al. (2018), the Herzberg O_2 continuum bands as well as the O_2O_2 absorption bands (Wulf bands) by Fally et al. (2000), the Herzberg O_2 band system (Mérieu et al. 2001), and several O_2 -X CIAs (Fauchez et al. 2020). The MT_CKD water continuum is parameterized in PSG by computing CIAs tables for H_2O - H_2O and H_2O - N_2 from that water model (Mlawer et al. 2012; Kofman & Villanueva 2021). PSG also includes a retrieval framework, optimal estimation and Multinest methods, and a noise simulator. The tool can be operated in several ways: via a public web interface,²⁴ via a flexible application program interface (API) accessible from any scripting language (e.g., python, IDL), and it can also be operated by an installable suite via the Docker virtualization system. Additionally, all opacity tables, spectroscopic databases, and core fundamental source codes are publicly available at the site and the PSG GitHub repository.

4.6. Exo_k/Pytmosph3r

Originally developed to easily handle, convert, and interpolate radiative opacities from various sources, the `Exo_k` python library (Leconte 2021) now includes a 1D radiative transfer module able to model the transmission and reflection/emission spectra of a wide range of atmospheres. Thanks to its original philosophy, `Exo_k` can include any molecule or opacity source for which opacities can be provided in the form of either cross sections or correlated- k coefficients tables in one of the many formats handled by the library. In particular, data can be provided at any resolution as the library implements a novel algorithm able to bin-down even k -coefficient tables (Leconte 2021). High numerical efficiency is ensured by the use of the “just in time” compilation capability of the `numba` library.

The reflection/emission algorithm is an improved version of the two-stream framework from Toon et al. (1989) that accounts for multiple scattering and scattering asymmetry (see the appendix of Chaverot et al. 2022 for details). Mie scattering can be modeled for any type of aerosol by providing aerosol optical properties. The change in atmospheric scale height due to both gravity and mean molecular weight variations are taken into account, both in emission and transmission.

`Pytmosph3r` is the 3D implementation of `Exo_k` that can be used to model observables from both parameterized 3D atmospheric structures and outputs from a 3D global climate model (Caldas et al. 2019; Falco et al. 2022). Using `Exo_k` for

²⁴ <https://psg.gsfc.nasa.gov>

all the opacity calculations, it inherits all the aforementioned features. In transmission, rays of light are shot through the limbs of the planet, possibly crossing many atmospheric columns of the atmospheric model and thus accounting for its heterogeneity, both along and across the limb. In emission/reflection, the two-stream algorithm from the 1D model is used in each atmospheric column to compute the radiative intensity in the direction of the observer—thus accounting for limb darkening. Intensities are then weighted by a solid angle and integrated over the visible sphere to yield the observed flux at the desired observer’s location.

4.7. SMART

The Spectral Mapping and Radiative Transfer (SMART) code is a highly flexible 1D plane-parallel, line-by-line, multistream, and multiple-scattering radiative transfer model originally presented by Meadows & Crisp (1996) and Crisp (1997). SMART solves the radiative transfer equation using the discrete ordinates method through the DISORT FORTRAN code (Stamnes et al. 1988). SMART ingests gaseous absorption coefficients generated by its line-by-line absorption coefficients (LBL-ABC) companion model, as well as optional supplemental cross sections typically used to compensate for gaps in linelist opacity data, and separately sourced collisionally induced absorption (CIA) coefficients. LBL-ABC uses the HITRAN linelists to calculate absorption coefficients (Gordon et al. 2022). Usually, cross section data derived from the MPI-Mainz database (Keller-Rudek et al. 2013) or other sources is used for UV–Vis wavelengths where linelist data is often absent; however, cross section data can also be used at any wavelength where appropriate. The same code is used for both shortwave and longwave radiation. The version of SMART employed here uses the ray-tracing model of Robinson (2017) to simulate transit transmission observations, which includes the effects of refraction. SMART has been repeatedly verified against terrestrial solar system objects (e.g., Tinetti et al. 2005; Robinson et al. 2011; Arney et al. 2014) and has a long history of being used to generate exoplanet spectra (e.g., Charnay et al. 2015; Lincowski et al. 2018; Lustig-Yaeger et al. 2019). SMART is highly customizable with user-selected input gases and high internal resolution. The model can also be configured for clouds and aerosols.

4.8. SOCRATES

Suite Of Community RAdiative Transfer codes based on Edwards & Slingo (SOCRATES; Edwards & Slingo 1996) is an open-source radiative transfer code with a two-stream multiple-scattering solver for reflected starlight and thermal emission. It uses the correlated- k method to solve for gaseous absorption, with further details covered in Edwards & Slingo (1996) and Manners et al. (2022). Opacity data are sourced from a range of databases such as HITRAN, HITEMP, and ExoMol (see Amundsen et al. 2014, 2016; Goyal et al. 2018, for full details). For case T3B, correlated tables were constructed from ExoMol linelists. At the moment, the SOCRATES code was written to calculate transmission spectra for the global grid from the Unified Model (UM)—a 3D GCM developed and used by the Met Office, and at present, SOCRATES cannot calculate transmission spectra outside a GCM. However, a standalone version of this calculation is planned for a near-future release of SOCRATES. Within each

GCM column, the transmitted solar flux is calculated using a spherical shell geometry, replacing the standard plane-parallel assumption (a detailed description can be found in Lines et al. 2018). The total optical depth through the atmosphere is found by summing the mass along the slant path through each shell multiplied by the corresponding mass extinction coefficient for the layer. Each model column is treated independently, and each layer is assumed to be homogeneous with identical optical properties in that column. The fluxes are calculated in the column facing away from the star in the limb. The resulting fluxes in each column are added up to represent the full transmission spectrum. With this approach, the variations of fluxes in the limb perpendicular to the observer can be fully captured, but those along the line of sight can only be approximated. The assumption of a homogeneous spherical shell might also lead to a bias if there are variations across the terminator.

4.9. TauREx 3

Tau Retrieval for Exoplanets (TauREx) is an open-source program initially developed by Waldmann et al. (2015b, 2015) for both transmission and thermal emission spectroscopy. The current version in use—now completely Python-stacked—corresponds to TauREx 3 and is substantially described in Al-Refaie et al. (2021). TauREx is a fully Bayesian, line-by-line radiative transfer and retrieval framework. It can be used with any molecular linelists in hdf5 and in pickle format, including the latest data from the Exomol project (Tennyson et al. 2016; Chubb et al. 2021), HITEMP (Rothman et al. 2014) and HITRAN (Gordon et al. 2017). Its modular, object-oriented architecture, combined with an extensive TauREx 3 Python library, makes it particularly flexible and customizable to the user. TauREx models a 1D atmosphere and computes the path integral from a single cross section where each (interpolated and weighted) opacity has been fused. The opacity integral is computed based on a finite difference scheme. In addition, TauREx provides a full implementation of Rayleigh and Mie scattering via extinction, as well as simple (gray) cloud models. It has been particularly tailored for the modeling and interpretation of transit spectra, so it only considers extinction processes, and it does not solve for scattering emission when computing thermal emission. The user builds both transmission and emission models by defining a temperature profile, pressure parameters, a chemistry model, and contributions to the optical depth (molecular absorption, CIA, Rayleigh, Mie scattering, simple clouds). The wavenumber grid of the forward model is selected at runtime from whichever loaded opacity with the highest resolution. Every other opacity is then resampled to the chosen opacity’s grid before any computation begins; this grid becomes the native grid of the forward model.

In this study, the T-P profile, as well as chemical abundances at each layer, were provided as inputs (thus not requiring the use of any TauREx chemical model). For the TauREx simulation results presented in Figure 2 of case T3B, we used ExoMol absorption cross sections for H₂O (Polyansky et al. 2018) and CH₄ (Yurchenko et al. 2017), CIA opacities for H₂–H₂ (Fletcher et al. 2018), H₂–He (Abel et al. 2011), CH₄-X (from HITRAN), as well as Rayleigh scattering contributions (Cox 2015).

Table 3
List of MALBEC Radiative Transfer “Casks” and Experiments

Cask	Planet	Geometry	Description	Aerosols	exoMIP
T0A	GJ1214b	Transit	Layering and ray-tracing test	Clear	...
T1A	TRAPPIST-1e	Transit	Compact terrestrial atmosphere	Clear	THAI
T1B	TRAPPIST-1e	Transit	Cloudy terrestrial atmosphere	Simple	THAI
T3A	GJ1214b	Transit	Isothermal mini-Neptune	Clear	CAMEMBERT
T3B	GJ1214b	Transit	Extended clear atmosphere	Clear	CAMEMBERT
T3C	GJ1214b	Transit	Cloudy extended atmosphere	Simple	PIE
T4A	Barstow et al. (2020)	Transit	Hot gas-rich atmosphere (model 0)	Clear	MOCHA
T4B	WASP-18b	Transit	Ref. to HST/JWST observations	Clear	MOCHA
D1A	TRAPPIST-1e	Phase 0/90/180	Compact atmosphere (as T1A)	Clear	THAI
D3B	GJ1214b	Phase 0/90/180	Deep atmosphere (as T3B)	Clear	CAMEMBERT
D5A	Earth-like	Phase 0/90/180	Modern Earth atmosphere	Clear	SAMOSA/FILLET
D5B	Earth-like	Phase 0/90/180	Cloudy modern Earth	Water-ice	CREME

Note. The exoMIP column describes which partner exoMIP these cases will be coordinated with.

4.10. PICASO

PICASO, the Planetary Intensity Code for Atmospheric Scattering Observations, is a radiative transfer code originally designed for reflected light planetary spectra (Batalha et al. 2019) and has since been expanded to include calculations of transmission and emission planetary spectra (Batalha & Rooney 2020; Batalha et al. 2021), 3D spectral modeling (Adams et al. 2022), phase curves (Robbins-Blanch et al. 2022), 1D climate modeling (Mukherjee et al. 2023), as well as fitting spectroscopic data to models (The JWST Transiting Exoplanet Community Early Release Science Team et al. 2023). The code is written in Python and publicly available on GitHub,²⁵ and has been used in a wide range of studies for exoplanets and brown dwarfs (e.g., Mukherjee et al. 2021; Foote et al. 2022; Kozakis et al. 2022; Lew et al. 2022). PICASO is based on codes originally described in McKay et al. (1989). Marley et al. (1999) and Cahoy et al. (2010) built off of methods from Toon et al. (1977, 1989) to compute reflected starlight and thermal emission as observed at full phase, with updates from Cahoy et al. (2010) for observations at any phase angle and including multiple scattering via the two-streams approximation.

The code is accompanied by a database of opacities resampled down to $R = 20,000$ and $60,000$ from line-by-line calculations on the order of $R \sim 10^6$ (see methodology in Gharib-Nezhad et al. 2021 for details). This database spans wavelengths from 0.3 to $14 \mu\text{m}$ and can either be downloaded in its entirety or queried using `sqlite3`. The resampled opacity database draws from multiple databases such as HITRAN and ExoMol, along with many other sources for linelists, CIAs, and cross sections, with the code providing helpful tools to extract the relevant references for individual model runs. If desired, PICASO additionally provides functions for the creation of correlated- k tables for individual gases. Within the code there are options to add clouds in the form of gray cloud coverage or cloud model output

(Rooney et al. 2022), as well as user-supplied parameterized arbitrary Rayleigh scattering.

5. The MALBEC Protocol and Experiments

The MALBEC model intercomparison contains three sets of experiments: (1) core, (2) transit, and (3) direct-imaging experiments, with a total of 12 cases (Table 3). All simulations are to be carried across a wide spectral range ($0.2\text{--}20 \mu\text{m}$) at moderate resolutions ($RP = 200$), and each model is asked to provide results only in regimes and spectral regions for which the models have been previously operated.

5.1. The Core Experiments

The core experiments (T0A, T3B) are defined specifically to test basic assumptions within each model and to further assist with the homogenization of the results. The purpose of the MALBEC exoMIP is not for all models to agree but to understand how certain aspects in the parameterization of the models impact the simulations. As such, issues with layering, opacities, refraction, linelists, and scattering considerations impact the results. As long as each model is parameterized as close as possible to the MALBEC configuration file, then each model result will reveal how parameterizations and modeling issues impact the spectra.

One of the most common differences in RT models is how each code discretizes and defines atmospheric layering. For instance, some models expect values describing the average properties across the layer, while others expect the properties at the border regions. If the number of layers is finite and reduced, then differences in these assumptions can lead to substantial errors. Such type of error primarily originates from the path length the light traces through the atmosphere. To investigate this, the T0A experiment explores how transit is modeled for the same atmosphere but when considering different numbers of layers (20, 50, 100, 1000). It is demonstrated that if the RT model implements some sort of sublayering algorithm (e.g., internally dividing the input layers to better capture the change

²⁵ <https://github.com/natashabatalha/picaso>

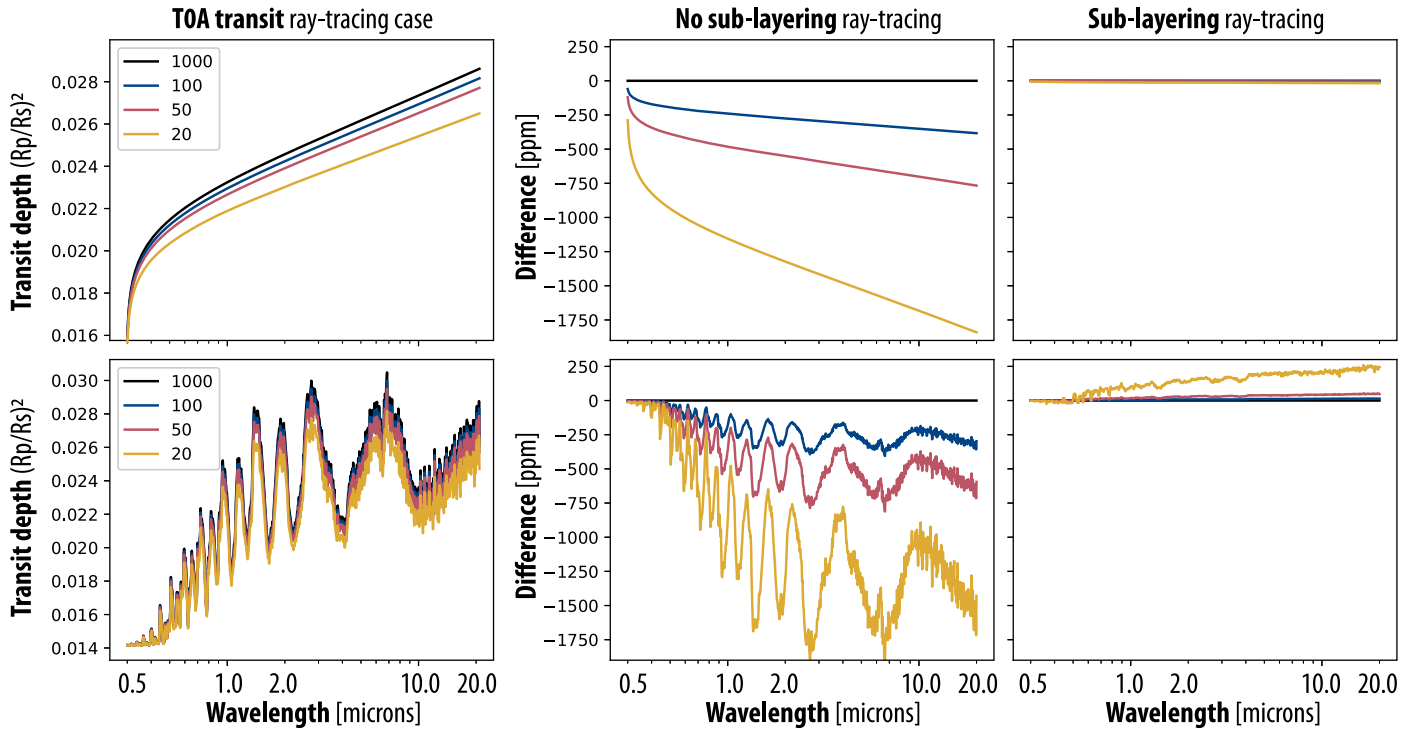


Figure 1. Comparison of the effects of atmospheric discretization and ray tracing on the PSG computation of transit spectra as part of the TOB test cask. The left panels show transit for a hypothetical extended atmosphere assuming a parameterized opacity function (top) and for realistic water opacities (bottom). The different traces were computed considering different number of layers (20, 50, 100, 1000) for the 100 bar to 1 nbar range. The middle panel shows differences with respect to the 1000 layer solution, which for 50 layers the differences can reach up to 500 ppm at $2 \mu\text{m}$. On the other hand, if the RT model implements a sublayering algorithm in the ray-tracing scheme, these discretization effects can be largely mitigated (right panels).

of densities and opacities), these effects can be largely mitigated and practically removed, as we can see in Figure 1 right panels where the transit depth difference between the 1000 layer cases and the lower layer cases tends toward 0 ppm when sublayering ray tracing is performed. As noted by Malik et al. (2017), sublayering can also have a strong effect on the numerical convergence of radiative equilibrium (RE) and Radiative Convective Equilibrium (RCE) models, which are used to establish and model the thermal structure of planetary atmospheres.

For the MALBEC intercomparison, the configuration files provide anchor values, values at the border of the layers, and analytical fits to each parameter across the atmosphere, so the modelers can explore and optimize the calculations of the layering as needed. For instance, if the models require layer quantities (not border values), providing straight average values of the two border values would be then more accurate than providing border values.

The TOA is a highly simplified atmosphere designed to test the treatments of layering and sublayering of the atmosphere and investigate the effects this has on the final absorbance spectrum. It is based on the atmosphere of GJ1214b, where the temperature decreases linearly from 1000 K at 100 bar to 400 K at the top of the atmosphere at 1 nano-bar. The experiment explores two scenarios: a case with a simple parameterized water cross section, increasing with wavelength, and a case employing realistic water opacities. The simple cross section increases exponentially from $10^{-26} \text{ cm}^2 \text{ molecule}^{-1}$ at $0.3 \mu\text{m}$ to $8 \cdot 10^{-19} \text{ cm}^2 \text{ molecule}^{-1}$ at $20 \mu\text{m}$, and is designed to probe different altitudes with wavelength and to highlight the layering effects in the final transmission spectrum. More details about the cross sections employed in this test can be found in the

MALBEC repository. In Figure 1, we show the comparison between the parameterized water absorption (top panels) and using the water cross section in the bottom panels (H_2O ; No CIA, no continuum, no Rayleigh for simplicity). This serves to show that our simple opacity covers a dynamic range similar to water.

The effects of not including sublayering in the RT models can be quite noticeable, as shown in the middle panels of Figure 1, reaching absolute errors beyond 500 ppm at $2 \mu\text{m}$ on the transit radius when no sublayering is included in the model. Interestingly, the error is not an offset, but more of a trend: the higher in the atmosphere the transit probes, the larger the error. This manifests in smaller errors at weaker water absorption (windows) versus where absorption is stronger (bands). This discretization and ray-tracing error can be largely mitigated when employing a sublayering algorithm, as shown in the right panels of Figure 1 and computed with PSG, which subdivides each layer into 10 sublayers when computing the ray-traced column densities.

T3B is a more advanced core experiment, although still with a limited amount of absorbers, integrating a complex temperature profile, varying molecular abundances, and all possible opacity effects (e.g., Rayleigh scattering, CIA, molecular absorptions). In Figure 2, we show preliminary results for this case computed using the models reported in Table 2 and in Section 4, in which we consider a base pressure of 1 bar, a planetary diameter of 36318 km, and a stellar diameter of 306108 km (more simulation details available at the MALBEC CKAN repository). One can see a general agreement yet notable differences at certain spectral regions remain. Many of the differences originate from the considered linelists—Assumptions regarding linelists, CIAs, and isotopic species

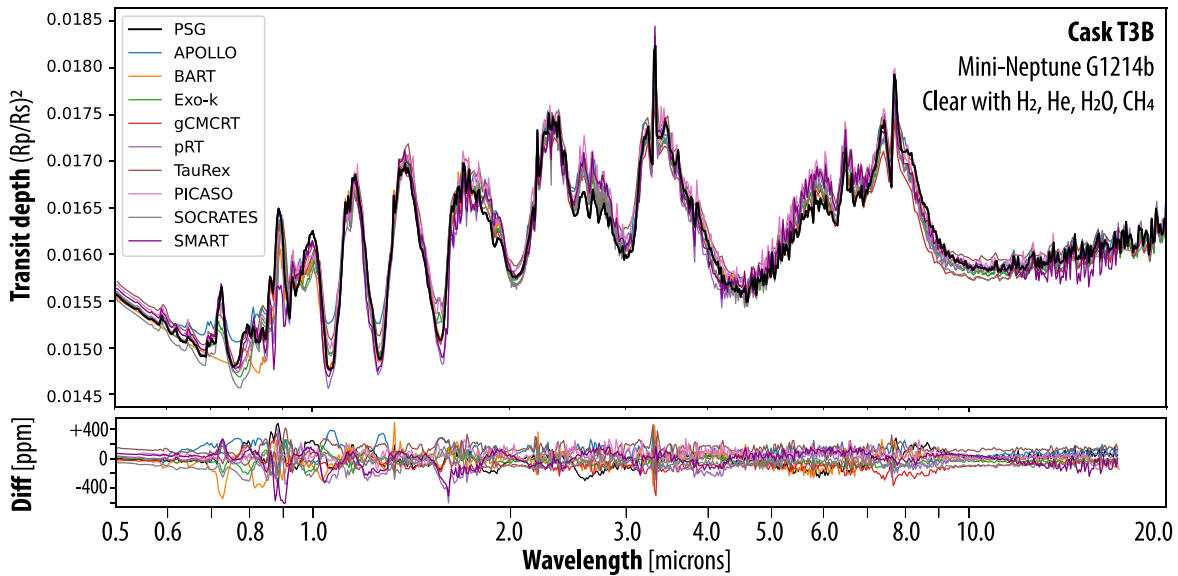


Figure 2. Comparison of preliminary simulation results for test cask T3B. The RT models were run using their typical and core functionalities (e.g., linelists, ray tracing, CIAs) with minimum adaption for this experiment, and were configured to take into account the MALBEC input parameters as closely as possible. There is general agreement among the models, with the valleys/peaks in most cases being reproduced by all the models. However, there are specific features in the UV/Optical and at certain wavelengths where the models disagree, which is probably related to the employed linelists and the available CIAs for each RT model. Further revised results with improved parameterizations for each RT model will be reported in a subsequent investigation. See detailed comparisons of the near-IR region and the optical region in Figures 3 and 4, respectively. The differences between the models, as shown in the lower panel, are relative to the mean of all the presented spectra.

can lead to notable differences (Fortney et al. 2020; Niraula et al. 2022). A detailed report exploring this in detail and for the other “casks” will be reported soon following this initial protocol report.

5.2. The Transit Simulations

The transit experiments (T0A, T1A, T1B, T3A, T3B, T3C, T4A, T4B) are meant to test the functionalities of the models for the interpretation of transit spectra, which is one of the prime techniques to characterize exoplanets. The resulting spectra in this mode will be primarily defined by the considered opacity processes and the ray-tracing algorithm employed by each model. As such, we have defined a set of experiments to test these models at extreme regimes, highly compact and dense atmospheres (e.g., TRAPPIST-1e; Gillon et al. 2017), extended and light atmospheres (e.g., GJ1214b; Charbonneau et al. 2009), and hot Jupiters (e.g., Barstow et al. 2020). By testing transit simulations at different temperatures, we expect to better understand the differences due to various linelist databases (e.g., HITRAN versus HITEMP versus ExoMol), while the different ambient compositions (e.g., high molecular weight atmospheres versus H₂-rich) would allow us to explore the effects on the assumptions regarding collisional broadening and CIAs. Results from these experiments are primarily tailored to understand the impact of the RT models when interpreting HST, JWST, and ARIEL transit data.

As we explore more in detail the origin of the discrepancies of the spectra presented in Figure 2, linelists and the methods considered to model opacities by each model are the dominating factors. The case of T3B is interesting in that it simulates a hydrogen-rich atmosphere at relatively high temperatures (350–900 K), requiring specializing the linelists. On the other hand, models primarily developed to model terrestrial planets and temperate planets are set to operate with HITRAN (mostly air collisional parameters and tailored for <300 K), which is limited in operability for this specific case.

When including the millions of transitions active at these high temperatures, correlated-*k* and precomputed high-resolution opacity tables are needed, yet how these databases are computed, the source linelists, and the way they are handled by each RT model can lead to substantial differences. As users use these models in their default operational regimes, they should be mindful of the limitations of the linelists configured for the simulation. Detailed comparisons between linelists and opacity modeling methods (e.g., line-by-line, correlated-*k*, and cross sections) have been discussed elsewhere as previously discussed, and as part of this model intercomparison, we primarily want to identify spectral regions in which models tend to disagree and why.

In Figure 3, we show a zoom of the near-IR region, in which one can see notable differences between the models and different levels of transit intensity, particularly near the “valley” regions. These valley regions are primarily affected by water and methane opacity effects. In order to demonstrate the effects of linelists in this spectral region of this specific test case, we ran one specific RT model (PSG) but employed different linelists and methods. If measured at extremely high resolution, the transit spectrum would look substantially spiky, as shown with the thin gray trace in panels (b) and (c), yet as we convolve the spectrum to a lower resolution (RP = 200), the corresponding spectrum (LBL-HITRAN) looks substantially smooth, with small wings and baseline considerations in the line-by-line analysis notably affecting the resulting convolved spectrum. The high-energy linelist (HITEMP), which is more suitable for this test case than HITRAN, fills substantially more of the valleys though, while the cross sections by Freedman et al. (2014), which rely on a different set of linelists, show a general better agreement to HITEMP than to HITRAN. The line-by-line and correlated-*k* methods, when considering the same source linelist and resolution, should in principle agree perfectly, yet small spectral grid interpolation effects do lead to

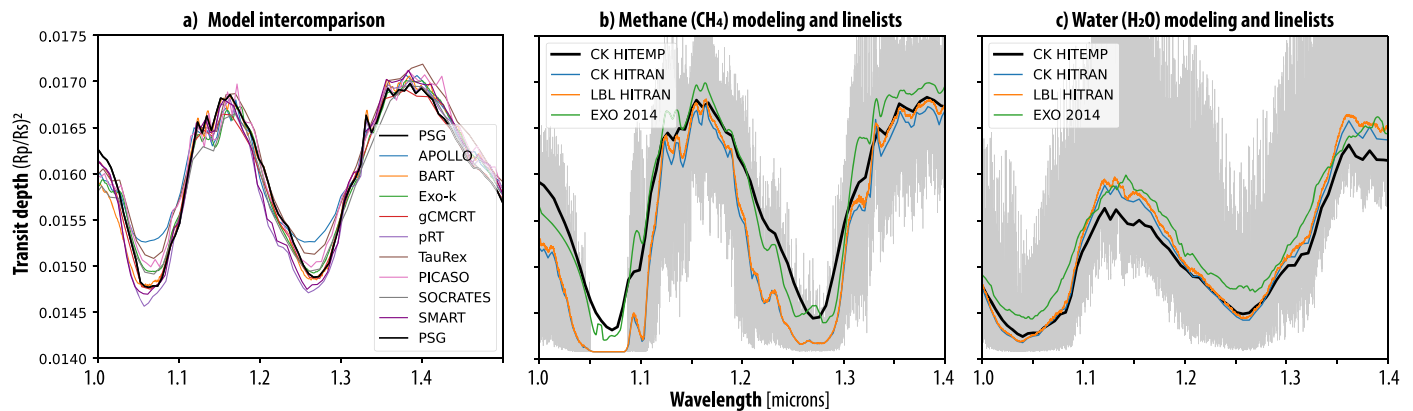


Figure 3. Comparison of model results in the near-IR and simulations computed employing different linelists—HITRAN, HITEMP, and EXO (Freedman et al. 2014)—and opacity algorithms line-by-line (LBL) and correlated- k (CK). There is relatively good agreement in the model intercomparison, yet notable differences are observed in the valleys at 1.05 and 1.3 μm regions. Water and methane are the main absorbers at these wavelengths, and how the opacities for these species are treated has a notable effect on how full the valleys look. Panels (b) and (c) show PSG simulations employing different methods and linelists for methane and water, respectively. The thin underlying gray trace is the high-resolution line-by-line spectrum as computed with the HITRAN database, and one can observe substantial differences depending on the choice of opacity database.

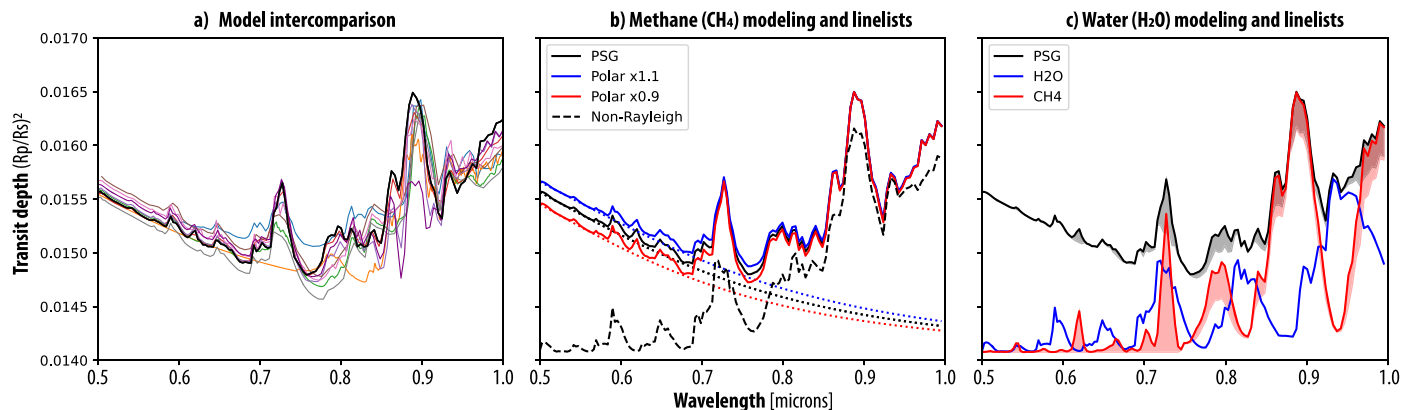


Figure 4. Comparison of preliminary simulation results for test case T3B. The RT models were run using their typical and core functionalities (e.g., linelists, ray tracing, CIAs) with minimum adaption for this experiment and were configured to take into account the MALBEC input parameters as closely as possible. There is general agreement among the models, with the valleys/peaks in most cases being reproduced by all the models. However, there are specific features in the UV/Optical and at certain wavelengths where the models disagree, which is probably related to the employed linelists and the available CIAs for each RT model. For instance, differences at the near-IR valleys are in most cases related to the considered CH_4 linelists (e.g., HITRAN [less absorption], HITEMP [intermediate], ExoMol [more absorption]). Further revised results with improved parameterizations for each RT model will be reported in a subsequent investigation.

subtle differences, as shown for LBL-HITRAN and CK-HITRAN.

A particularly challenging spectral region is the optical and UV, in which linelists normally provide incomplete coverage, and where different RT models treat molecular opacities and Rayleigh scattering in different ways. As such, simulation results differ substantially in this spectral region, as panel (a) in Figure 4 shows. Some RT models (e.g., PSG) compute Rayleigh cross sections as a summation of the individual molecular cross sections, and based on the individual polarizability of the encompassing molecules. Subtle differences in how each model handles Rayleigh scattering or the assumed molecular polarizability values would lead to differences in the generated spectra. In panel (b), we perturb the PSG model by scaling the polarizability values by 0.9 and 1.1, and one obtains a level of difference comparable to the observed model differences. More complex is the treatment of molecular opacities at these wavelengths, since only sparse and limited molecular spectroscopy is available. At these high frequencies, the energy structure of molecules becomes increasingly nonquantized into a broad swath of energy levels. As such, classical linelists do not accurately characterize these

high-frequency spectral regions, and laboratory measurements of molecular cross sections are typically needed to fully capture the opacities at these wavelengths. For instance, the line-by-line and correlated- k simulations in PSG are complemented in this spectral range with laboratory cross sections collected from a range of spectral databases, with most of these cross sections originating from the MPI-Mainz Spectral Atlas (Keller-Rudek et al. 2013). The shaded regions in panel (c) show how this complements, which is particularly notable for methane near 0.7 μm . As better and more complete spectroscopy becomes available over time for a wide range of collisional and excitation regimes, the RT models will need to be updated and adapted to ingest these. More details regarding the databases, the methods, and the lineshape considerations employed by each model, as well as the associated implications and recommendations, will be presented in a subsequent paper that discusses these findings in detail.

5.3. The Direct-imaging Simulations

For the direct-imaging experiments (D1A, D3B, D5A, D5B), the simulations are to be carried out when the planets are in

quadrature (phase 90), the prime phase employed by coronagraphy, and sampling the night-hemisphere (phase 180) to test thermal emissions, and as seen from the star (phase 0, geometric albedo). These modes will allow us to better separate issues when modeling thermal and reflected light, each being susceptible differently to the assumptions regarding multiple-scattering, observational angle, disk discretization, and spherical considerations. The simulations are to be carried out by employing the highest fidelity available for each model, including multiple scattering, the highest number of scattering/Legendre terms possible, and realistic stellar and surface flux models. The tests are designed with different levels of complexity. D1A is the simplest test: it has no aerosols, a generally thin atmosphere (with N_2 , CO_2 , and H_2O starting at 1 atm), and the reflected and thermal emissions are mostly dominated by the modeling of the surface, which is assumed to be simply a Lambert scattering solid surface with a constant albedo and emissivity. Test D3B explores how models handle optically thick atmospheres, in which the deep bottom is defined by opacity functions and not by a solid surface, and it will test how each model handles deep gas atmospheres as those of gas giants. The D3B atmosphere is free of aerosols, so the opacity terms are relatively simple to model and defined by molecular extinction and Rayleigh scattering, yet the treatment of reflected radiation for the quadrature phase will explore how the models discretize and sample the disk across a wide range of illumination regimes in an optically thick multiple-scattering non-Lambertian deep atmosphere. The D5A test explores a much more spectrally complex atmosphere of Earth, with strong signatures of H_2O , O_3 , O_2 , CH_4 , CO_2 , and CO , in an ambient air mixture (O_2/N_2). In particular, this test will be used to analyze the treatment of specialized UV cross sections and the modeling of Rayleigh multiple scattering across models and RT algorithms. Case D5B is based on D5A but adds the complexity of clouds, and the test is designed to produce a realistic Earth spectrum as targeted by future observatories. One type of cloud is considered in this test, yet the size distribution is assumed to vary across altitude together with its abundance. A specific target of this test is to explore how each model handle handles size distributions of aerosols, optical constants, and multiple scattering for a wide range of incidence and emission angles.

We expect more differences in the reflected component of the spectrum, in which the effects of multiple scattering are more notable and where differences in the treatment of UV opacities and stellar models/templates could become significant. These experiments are specifically tailored to test the differences of RT models in the characterization of large exoplanets with JWST via secondary transit/phase analysis, of gas giants with the Roman Space Telescope, and of Earth-like planets via direct imaging. In particular, future flagship observatories (e.g., Habitable Worlds Observatory, LIFE mission) have as a prime target the characterization of Earth-like planets via direct imaging, so better understanding the current state of RT modeling of these planets is of great importance when designing these future telescopes.

6. The MALBEC Repositories

The input MALBEC files and all the output data resulting from the different simulations and analyses are to be accessible at the MALBEC permanent repository²⁶ by participating scientists and the exoplanet community. We have prepared

the MALBEC input files for each of the experiments and provided preliminary results as computed with at least one of the participating RT models. Ultimately, the results for all models will be made available for public access upon the publication of the results. Prepublication access can be requested by contacting the authors. Additional inputs and scripts related to the analysis of the data and production of plots for the publications will be made available on the MALBEC GitHub repository.²⁷

7. Summary

In this paper, we presented the protocol for the MALBEC model intercomparison project aiming to compare radiative transfer models used by the exoplanet community. MALBEC is designed to have a low entrance barrier across three sets of experiments, therefore allowing for broad participation with no requirement to participate in *all* experiments. The sets of experiments allow us to test numerous configurations of exoplanet RT calculations, such as transmission and emission spectroscopy, as well as direct imaging, for terrestrial planets, mini-Neptunes, and hot Jupiters. The goal is to highlight any differences arising from different opacity calculations, line broadening, collision-induced absorption, Rayleigh scattering, layering treatment, etc. In the context of JWST and while preparing for future observatories, differences in the number of hours of integration to detect molecules of interest will be quantified using the spectra provided by each RT model combined with a single noise model. User-friendly MALBEC ASCII configuration files containing all the required parameters have been created for each of the cases, associated with analytical profiles for some of them. Output files for at least one RT model are provided to guide the community in the generation of spectra for each experiment. MALBEC will also serve as a bridge between other CUISINES exoMIPs, such as the Photochemical model Intercomparison for Exoplanets (PIE) to provide atmospheric profiles with multiple gas abundances, the TRAPPIST-1 Habitable Atmosphere Intercomparison (THAI) for TRAPPIST-1e clear and cloudy modern Earth-like profiles, and (CAMEMBER) for GJ1214b profiles. Participation in MALBEC is open to anyone, with monthly team meetings and annual Building a Unified Framework For Exoplanet Treatments (BUFFET) workshops, with the third edition planned for late 2024.

Acknowledgments


















MALBEC belongs to the CUISINES metaframework, a Nexus for Exoplanet System Science (NExSS) science working group. G.L.V., T.J.F., V.K., and S.F. acknowledge support from the GSFC Sellers Exoplanet Environments Collaboration (SEEC), which is funded in part by the NASA Planetary Science Divisions Internal Scientist Funding Model. E.A.'s work has been carried out within the framework of the NCCR PlanetS supported by the Swiss National Science Foundation under grants 51NF40_182901 and 51NF40_205606. E.L. is supported by the SNSF Ambizione Fellowship grant (#193448). Material produced using Met Office Software. We acknowledge the use of the Monsoon2 system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, a strategic partnership between

²⁶ <https://ckan.emac.gsfc.nasa.gov/organization/cuisines-malbec>

²⁷ <https://github.com/projectcuisines/malbec>

the Met Office and the Natural Environment Research Council. His work was supported by a Leverhulme Trust research project grant (RPG-2020-82), a Science and Technology Facilities Council Consolidated grant (ST/R000395/1), and a UKRI Future Leaders Fellowship (MR/T040866/1). M.T.M. and N.J.M. acknowledge funding from the Bell Burnell Graduate Scholarship Fund, administered and managed by the Institute of Physics, which made this work possible. M.D. H. was supported by the NASA Fellowship Activity under NASA grant 80NSSC20K0682 and by an appointment to the NASA Postdoctoral Program at the NASA Goddard Space Flight Center, administered by Oak Ridge Associated Universities under contract with NASA. M.L. and E.W.S. acknowledge support from the Virtual Planetary Laboratory, which is funded via NASA Astrobiology Program grant No. 80NSSC18K0829, and the NASA Interdisciplinary Consortia for Astrobiology Research (ICAR) Program via grant Nos. 80NSSC21K0905 and 80NSSC21K0594. A.R.H. was supported by NASA under award number 80GSFC21M0002 through the CRESST II cooperative agreement. We sincerely thank Duncan Christie for his help with generating SOCRATES spectral files for the T3B case.

ORCID iDs

Geronimo L. Villanueva  <https://orcid.org/0000-0002-2662-5776>
 Thomas J. Fauchez  <https://orcid.org/0000-0002-5967-9631>
 Vincent Kofman  <https://orcid.org/0000-0002-5060-1993>
 Eleonora Alei  <https://orcid.org/0000-0002-0006-1175>
 Elspeth K. H. Lee  <https://orcid.org/0000-0002-3052-7116>
 Estelle Janin  <https://orcid.org/0000-0003-0475-8479>
 Michael D. Himes  <https://orcid.org/0000-0002-9338-8600>
 Jérémy Leconte  <https://orcid.org/0000-0002-3555-480X>
 Michaela Leung  <https://orcid.org/0000-0003-1906-5093>
 Sara Faggi  <https://orcid.org/0000-0003-0194-5615>
 Mei Ting Mak  <https://orcid.org/0000-0002-1624-3360>
 Denis E. Sergeev  <https://orcid.org/0000-0001-8832-5288>
 Thea Kozakis  <https://orcid.org/0000-0002-3868-2129>
 James Manners  <https://orcid.org/0000-0003-4402-6811>
 Nathan Mayne  <https://orcid.org/0000-0001-6707-4563>
 Edward W. Schwieterman  <https://orcid.org/0000-0002-2949-2163>
 Alex R. Howe  <https://orcid.org/0000-0002-4884-7150>
 Natasha Batalha  <https://orcid.org/0000-0003-1240-6844>

References

Abel, M., Frommhold, L., Li, X., & Hunt, K. L. C. 2011, *JPCA*, **115**, 6805
 Adams, D. J., Kataria, T., Batalha, N. E., Gao, P., & Knutson, H. A. 2022, *ApJ*, **926**, 157
 Ahrer, E.-M., Alderson, L., Batalha, N. M., et al. 2022, *Natur*, **614**, 649
 Alei, E., Konrad, B. S., Angerhausen, D., et al. 2022, *A&A*, **665**, A106
 Al-Refaie, A. F., Changeat, Q., Venot, O., Waldmann, I. P., & Tinetti, G. 2021, *ApJ*, **932**, 123
 Amundsen, D. S., Baraffe, I., Tremblin, P., et al. 2014, *A&A*, **564**, A59
 Amundsen, D. S., Mayne, N. J., Baraffe, I., et al. 2016, *A&A*, **595**, A36
 Arney, G., Meadows, V., Crisp, D., et al. 2014, *JGRE*, **119**, 1860
 Arney, G. N., Meadows, V. S., Domagal-Goldman, S. D., et al. 2017, *ApJ*, **836**, 49
 Barstow, J. K., Changeat, Q., Chubb, K. L., et al. 2022, *ExA*, **53**, 447
 Barstow, J. K., Changeat, Q., Garland, R., et al. 2020, *MNRAS*, **493**, 4884
 Batalha, N., & Rooney, C. 2020, natashabatalha/picasso: Release v2.1, Zenodo, doi:10.5281/zenodo.4206648
 Batalha, N., Rooney, C., & MacDonald, R. 2021, natashabatalha/picasso: Release 2.2, v2.2.0, Zenodo, doi:10.5281/zenodo.5093710
 Batalha, N. E., Marley, M. S., Lewis, N. K., & Fortney, J. J. 2019, *ApJ*, **878**, 70

Baudino, J.-L., Mollière, P., Venot, O., et al. 2017, *ApJ*, **850**, 150
 Blečić, J., Harrington, J., Cubillos, P. E., et al. 2022, *PSJ*, **3**, 82
 Borysow, A. 2002, *A&A*, **390**, 779
 Borysow, A., Jorgensen, U. G., & Fu, Y. 2001, *JQSRT*, **68**, 235
 Cahoy, K. L., Marley, M. S., & Fortney, J. J. 2010, *ApJ*, **724**, 189
 Caldas, A., Leconte, J., Selsis, F., et al. 2019, *A&A*, **623**, A161
 Charbonneau, D., Berta, Z. K., Irwin, J., et al. 2009, *Natur*, **462**, 891
 Charnay, B., Meadows, V., Misra, A., Leconte, J., & Arney, G. 2015, *ApJL*, **813**, L1
 Chaverot, G., Turbet, M., Bolmont, E., & Leconte, J. 2022, *A&A*, **658**, A40
 Christie, D. A., Lee, E. K. H., Innes, H., et al. 2022, *PSJ*, **3**, 261
 Chubb, K. L., Rocchetto, M., Yurchenko, S. N., et al. 2021, *A&A*, **646**, A21
 Cox, A. 2015, *Allen's Astrophysical Quantities* (Berlin: Springer)
 Crisp, D. 1997, *GeoRL*, **24**, 571
 Cubillos, P., Harrington, J., Lored, T. J., et al. 2017, *AJ*, **153**, 3
 Cubillos, P. E. 2017, *ApJ*, **850**, 32
 Cubillos, P. E., & Blečić, J. 2021, *MNRAS*, **505**, 2675
 Cubillos, P. E., Harrington, J., Blečić, J., et al. 2022, *PSJ*, **3**, 81
 Dahlback, A., & Stamnes, K. 1991, *P&SS*, **39**, 671
 Deitrick, R., Haqq-Misra, J., Kadoya, S., et al. 2023, *PSJ*, **4**, 39
 Edwards, J. M., & Slingo, A. 1996, *QJRM*, **122**, 689
 Eyring, V., Bony, S., Meehl, G. A., et al. 2016, *GMD*, **9**, 1937
 Falco, A., Zingales, T., Pluriel, W., & Leconte, J. 2022, *A&A*, **658**, A41
 Fally, S., Vandaele, A. C., Carleer, M., et al. 2000, *JMoSp*, **204**, 10
 Fauchez, T. J., Turbet, M., Villanueva, G. L., et al. 2019, *ApJ*, **887**, 194
 Fauchez, T. J., Villanueva, G. L., Schwieterman, E. W., et al. 2020, *NatAs*, **4**, 372
 Fauchez, T. J., Villanueva, G. L., Sergeev, D. E., et al. 2022, *PSJ*, **3**, 213
 Fletcher, L. N., Gustafsson, M., & Orton, G. S. 2018, *ApJS*, **235**, 24
 Foote, T. O., Lewis, N. K., Kilpatrick, B. M., et al. 2022, *AJ*, **163**, 7
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
 Fortney, J. J., Visscher, C., Marley, M. S., et al. 2020, *AJ*, **160**, 288
 Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, *ApJS*, **214**, 25
 Garland, R., & Irwin, P. G. J. 2019, arXiv:1903.03997
 Gharib-Nezhad, E., Iyer, A., Line, M. R., et al. 2021, 2021 Int. Symp. on Molecular Spectroscopy (Virtual) (Urbana, IL: Univ. Illinois)
 Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, *Natur*, **542**, 456
 Goldblatt, C., Kavanagh, L., & Dewey, M. 2017, *GMD*, **10**, 3931
 Gordon, I. E., Rothman, L. S., Hargreaves, R. J., et al. 2022, *JQSRT*, **277**, 107949
 Gordon, I. E., Rothman, L. S., Hill, C., et al. 2017, *JQSRT*, **203**, 3
 Goyal, J. M., Mayne, N., Sing, D. K., et al. 2018, *MNRAS*, **474**, 5158
 Grimm, S. L., Malik, M., Kitzmann, D., et al. 2021, *ApJS*, **253**, 30
 Haqq-Misra, J., Wolf, E. T., Fauchez, T. J., Shields, A. L., & Kopparapu, R. K. 2022, *PSJ*, **3**, 260
 Hargreaves, R. J., Gordon, I. E., Rey, M., et al. 2020, *ApJS*, **247**, 55
 Hargreaves, R. J., Gordon, I. E., Rothman, L. S., et al. 2019, *JQSRT*, **232**, 35
 Harrington, J., Himes, M. D., Cubillos, P. E., et al. 2022, *PSJ*, **3**, 80
 Himes, M. D., & Harrington, J. 2022, *ApJ*, **931**, 86
 Hogan, R. J., & Matricardi, M. 2020, *GMD*, **13**, 6501
 Howe, A. R., McElwain, M. W., & Mandell, A. M. 2022, *ApJ*, **935**, 107
 Irwin, P., Teanby, N., de Kok, R., et al. 2008, *JQSRT*, **109**, 1136
 Karman, T., Gordon, I. E., van der Avoird, A., et al. 2019, *Icar*, **328**, 160
 Keller-Rudek, H., Moortgat, G. K., Sander, R., & Sörensen, R. 2013, *ESSD*, **5**, 365
 Kofman, V., & Villanueva, G. L. 2021, *JQSRT*, **270**, 107708
 Konrad, B. S., Alei, E., Quanz, S. P., et al. 2022, *A&A*, **664**, A23
 Kozakis, T., Mendonça, J. M., & Buchhave, L. A. 2022, *A&A*, **665**, A156
 Leconte, J. 2021, *A&A*, **645**, A20
 Lee, E., Taylor, J., Grimm, S. L., et al. 2019, *MNRAS*, **487**, 2082
 Lee, E. K. H., Wardenier, J. P., Prinoth, B., et al. 2022, *ApJ*, **929**, 180
 Lew, B. W. P., Apai, D., Zhou, Y., et al. 2022, *AJ*, **163**, 8
 Li, G., Gordon, I. E., Rothman, L. S., et al. 2015, *ApJS*, **216**, 15
 Lincowski, A. P., Meadows, V. S., Crisp, D., et al. 2018, *ApJ*, **867**, 76
 Lines, S., Manners, J., Mayne, N. J., et al. 2018, *MNRAS*, **481**, 194
 Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, *AJ*, **158**, 27
 MacDonald, R. J., & Madhusudhan, N. 2017, *MNRAS*, **469**, 1979
 Mai, C., & Line, M. R. 2019, *ApJ*, **883**, 144
 Malik, M., Grosheintz, L., Mendonça, J. M., et al. 2017, *AJ*, **153**, 56
 Manners, J., Edwards, J. M., Hill, P., & Thelen, J.-C. 2022, SOCRATES (Suite of Community RAdiative Transfer codes based on Edwards and Slingo) Technical Guide, Met Office, UK, <https://code.metoffice.gov.uk/trac/socrates>

- Marley, M. S., Gelino, C., Stephens, D., Lunine, J. I., & Freedman, R. 1999, *ApJ*, **513**, 879
- McKay, C. P., Pollack, J. B., & Courtin, R. 1989, *Icar*, **80**, 23
- Meador, W. E., & Weaver, W. R. 1980, *JATIS*, **37**, 630
- Meadows, V. S., & Crisp, D. 1996, *JGR*, **101**, 4595
- Mérienne, M.-F., Jenouvrier, A., Coquart, B., et al. 2001, *JMoSp*, **207**, 120
- Min, M., Ormel, C. W., Chubb, K., Helling, C., & Kawashima, Y. 2020, *A&A*, **642**, A28
- Mlawer, E. J., Payne, V. H., Moncet, J. L., et al. 2012, *RSPTA*, **370**, 2520
- Mollière, P., Stolker, T., Lacour, S., et al. 2020, *A&A*, **640**, A131
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, **627**, A67
- Mukherjee, S., Batalha, N. E., Fortney, J. J., & Marley, M. S. 2023, *ApJ*, **942**, 71
- Mukherjee, S., Fortney, J. J., Jensen-Clem, R., et al. 2021, *ApJ*, **923**, 113
- Niraula, P., de Wit, J., Gordon, I. E., et al. 2022, *NatAs*, **6**, 1287
- Paradise, A., Macdonald, E., Menou, K., Lee, C., & Fan, B. L. 2022, *MNRAS*, **511**, 3272
- Partridge, H., & Schwenke, D. W. 1997, *JChPh*, **106**, 4618
- Pincus, R., Buehler, S. A., Brath, M., et al. 2020, *JGRD*, **125**, e2020JD033483
- Pinty, B., Gobron, N., Widlowski, J.-L., et al. 2001, *JGR*, **106**, 11,937
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, *MNRAS*, **480**, 2597
- Robbins-Blanch, N., Kataria, T., Batalha, N. E., & Adams, D. J. 2022, *ApJ*, **930**, 93
- Robinson, T. D. 2017, *ApJ*, **836**, 236
- Robinson, T. D., Meadows, V. S., Crisp, D., et al. 2011, *AsBio*, **11**, 393
- Rojo, P. M. 2006, PhD thesis, Cornell Univ.
- Rooney, C. M., Batalha, N. E., Gao, P., & Marley, M. S. 2022, *ApJ*, **925**, 33
- Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, *JQSRT*, **111**, 2139
- Rothman, L. S., Gordon, I. E., Rothman, L. S., & Gordon, I. E. 2014, 13th Biennial HITRAN Conf., Zenodo, doi:10.5281/ZENODO.11207
- Schwenke, D. W. 1998, *FaDi*, **109**, 321
- Serdychenko, A., Gorshelev, V., Weber, M., Chehade, W., & Burrows, J. P. 2014, *AMT*, **7**, 625
- Sergeev, D. E., Fauchez, T. J., Turbet, M., et al. 2022, *PSJ*, **3**, 212
- Sneep, M., & Ubachs, W. 2005, *J. Quant. Spec. Radiat. Transf.*, **92**, 293
- Stamnes, K., Tsay, S.-C., & Laszlo, I. 2000, DISORT, a General-Purpose Fortran Program for Discrete-Ordinate-Method Radiative Transfer in Scattering and Emitting Layered Media: Documentations of Methodology NASA Technical Report, Version 1.1, NASA
- Stamnes, K., Tsay, S. C., Wiscombe, W., & Jayaweera, K. 1988, *ApOpt*, **27**, 2502
- Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2016, *JMoSp*, **327**, 73
- The JWST Transiting Exoplanet Community Early Release Science Team, Ahler, E. M., Alderson, L., et al. 2023, *Natur*, **614**, 649
- Tinetti, G., Meadows, V. S., Crisp, D., et al. 2005, *AsBio*, **5**, 461
- Toon, O. B., McKay, C. P., Ackerman, T. P., & Santhanam, K. 1989, *JGR*, **94**, 16287
- Toon, O. B., Pollack, J. B., & Sagan, C. 1977, *Icar*, **30**, 663
- Turbet, M., Fauchez, T. J., Sergeev, D. E., et al. 2022, *PSJ*, **3**, 211
- Venot, O., Bénilan, Y., Fray, N., et al. 2018, *A&A*, **609**, A34
- Villanueva, G. L., Liuzzi, G., Faggi, S., et al. 2022, Fundamentals of the Planetary Spectrum Generator (Greenbelt, MD: NASA Goddard Space Flight Center)
- Villanueva, G. L., Smith, M. D., Protopapa, S., Faggi, S., & Mandell, A. M. 2018, *JQSRT*, **217**, 86
- Waldmann, I. P., Rocchetto, M., Tinetti, G., et al. 2015b, *ApJ*, **813**, 13
- Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, *ApJ*, **802**, 107
- Wolf, E. T., Kopparapu, R., Haqq-Misra, J., & Fauchez, T. J. 2022, *PSJ*, **3**, 7
- Wunderlich, F., Scheucher, M., Godolt, M., et al. 2020, *ApJ*, **901**, 126
- Yang, J., Leconte, J., Wolf, E. T., et al. 2016, *ApJ*, **826**, 222
- Yurchenko, S. N., Amundsen, D. S., Tennyson, J., & Waldmann, I. P. 2017, *A&A*, **605**, A95
- Zawada, D., Franssens, G., Loughman, R., et al. 2021, *AMT*, **14**, 3953