UNIVERSITÀ DEGLI STUDI DI PADOVA

Physics and Astronomy Department "Galileo Galilei"
Master in Astrophysics and Cosmology

Atmospheric characterization of K2-18b

Thesis Supervisor

Dr. Tiziano Zingales

Thesis Co-supervisor

Prof. Luca Malavolta

Candidate

Ylenia Mascolo

Academic Year 2023/2024

Acknowledgements

I would like to express my gratitude to my Supervisor, Dr. Tiziano Zingales, for his valuable guidance and advice during the writing of this thesis. His availability and expertise have been fundamental to this work. A sincere thanks also to my Co-supervisor, Prof. Luca Malavolta, for his suggestion and the interest he has shown in this thesis project.

Abstract

This thesis investigates the atmosphere of K2-18b, a sub-Neptune exoplanet located in the habitable zone of an M-dwarf star. The study aims to characterize the chemical composition and the temperature-pressure (T-P) profile of the atmosphere to understand whether K2-18b has a primary or secondary atmosphere.

The research starts with an overview of small exoplanets, focusing on super-Earths and sub-Neptunes, two categories that include K2-18b. This thesis examines the gap in the radius distribution between these types of planets, exploring theories like photoevaporation and core-powered mass loss that might explain the transition between primary and secondary atmospheres. It continues by illustrating the various results relating to the atmosphere of K2-18b in the literature, concluding that the planet is a sub-Neptune with an H_2-rich atmosphere.

To verify the kinds of atmosphere of K2-18b, a complete analysis of the planet's atmospheric composition and temperature-pressure profile is conducted using data from the James Webb Space Telescope (JWST), through transit spectroscopy and transmission spectrum, already reduced. Various atmospheric models are simulated and compared, in particular molecular constant abundance and equilibrium chemistry models with isothermal and 4-point T-P profiles. For all the retrievals is dominant the presence of the CH_4 and CO_2 molecules. The best model, in this thesis, considers a constant molecular abundance profile and isothermal temperature pressure profile. The available transmission spectrum does not give enough information to choose a more complex model with a high statistical significance. The results suggest that K2-18b's atmosphere is likely in a transitional state, between a hydrogen-rich primary atmosphere and a secondary atmosphere containing heavier molecules.

The study concludes by considering the need for additional observations with higher precision and a larger wavelength range to confirm whether K2-18b has a hybrid or fully secondary atmosphere.

Contents

Chapter 1

Introduction

In the Universe there are several celestial objects worthy of note and interesting to study: among these are exoplanets. Exoplanets are extrasolar planets that orbit stars beyond our Solar System. The quest to uncover exoplanets traces its roots back centuries, with early astronomers speculating about the existence of other worlds. The first hypothesis of the existence of other worlds was traced to Greek philosophers.
Democrito ($\sim$ 460-270 A.C.) said:

> "In some worlds, neither the Sun nor the Moon is present, in others both are greater than those in our world and others they are great in number. There are worlds without creatures, plants or any humidity."

Epicuro ($\sim$ 341-270 A.C.) [29] added:

> "There is an infinity of worlds, someone is similar to our one, and others are different. There are no reasons to not believe that in other worlds there are animal species and vegetable species and all we see."

In the next centuries, some important astronomers exposed themselves assuming the existence of infinite planets, like Giordano Bruno (1548-1600) [14], who said:

> "Therefore, innumerable are the worlds and infinite the earths that orbit around those suns as we see the seven orbiting our Sun."

For this quote and his theories, judged heretical, Giordano Bruno was burnt alive.
Today we know that the Sun is only a single star in a peripheral position concerning the center of our Galaxy and that there are thousands of billions of galaxies. Most of them host hundreds of billions of stars or even more, and all of these stars host one or more planets [5]. Using the words of Edwin Hubble (1889-1953):

> "The scientific community has long supposed that if stars are suns (and vice versa!), and the Sun has planets, then it is highly probable that the other stars also have planets."

However, it wasn't until relatively recent times that technological advancements permitted the detection and characterization of exoplanets. One of the most important moments in the history of exoplanetary exploration occurred in 1992 when astronomers Aleksander Wolszczan and Dale Frail announced the presence of two planets orbiting a pulsar, PSR B1257+12 [92]. The unusual nature of these planets, orbiting a rapidly rotating neutron star rather than a main-sequence star, led to some scepticism in the scientific community.

The official discovery of the first exoplanet orbiting a main-sequence star was confirmed in 1995 by Michel Mayor and Didier Queloz [65]. They used the radial velocity method to analyse the stellar spectrum of the star 51 Pegasi and they found variations in the velocity, indicative of an unseen companion orbiting around it. The detected companion, called 51 Pegasi b, is a gas giant planet with roughly half the mass of Jupiter but with an incredibly close orbit to its host star which permits the revolution in just four days.

Since the first discovery in 1995, the number of techniques implemented for exoplanet discovery has increased as has the number of planets observed. To date, more than 5,600 exoplanets have been identified[1]. There are different methods scientists commonly use to discover exoplanets. The most successful techniques have been the transit and radial velocity methods. The first is based on the passage of a planet between its host star and the observer, causing a dim of the star's light, while the second is based on the periodic variation of the star relative to the centre of mass of the system. Another technique is Microlensing, which uses the gravitational distortion of the light as a planet passes between its host star and the observer. Direct imaging, when it is possible to take pictures of exoplanets using techniques that remove the overwhelming glare of the stars they orbit. Finally, Astrometry, when the presence of a planet causes the wobble of a star around in space concerning nearby stars [20]. There are other minor techniques used to discover exoplanets such as timing variation, orbital brightness modulation and disk kinematics.

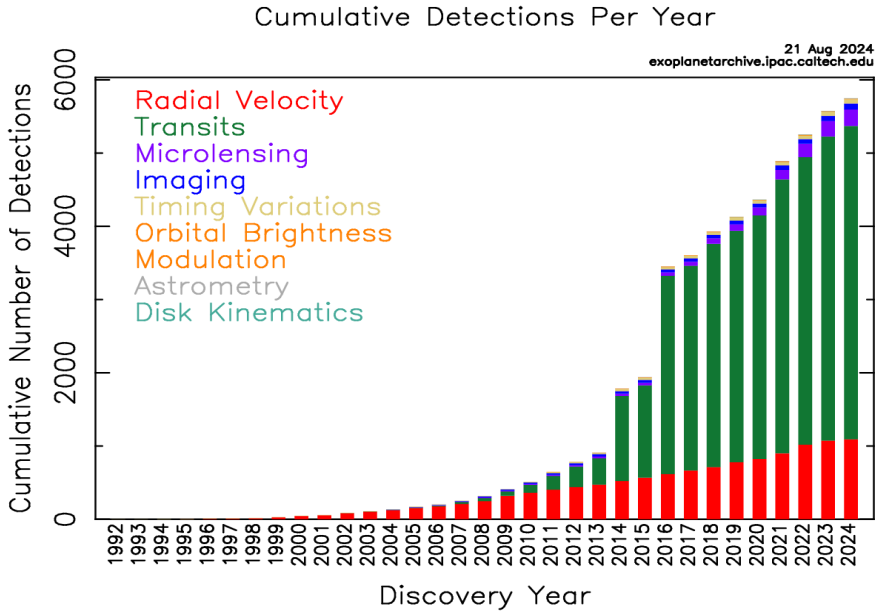

Figure 1.1: Cumulative detections of exoplanets per year, considering the various techniques implemented for the observations. Credits: NASA Exoplanet Archive[2].

As of today, thanks to the Kepler mission, we know that every star beyond our Solar System hosts at least one planet [5]. The number of planets discovered is growing rapidly, as visible in the Figure (1.1). These planets come in many different forms and follow a variety of orbits. In particular, it is possible to divide the exoplanets into giants and small planets.

[1] https://science.nasa.gov/exoplanets/
[2] https://exoplanetarchive.ipac.caltech.edu/exoplanetplots/

Gas giants are planets greater than Saturn with a radius $R > 0.6\ R_J$ [21]. A subclass is formed by the hot Jupiters, which orbit very close to their stars, making them extremely hot and subject to stellar winds. Ice giants, or Neptunian planets, are similar in size to Neptune or Uranus and typically have hydrogen and helium-dominated atmospheres with cores of rock and metals. Small planets have a radius smaller than Neptune's radius and can be divided into Earth-sizes, super-Earths and sub-Neptunes. The main difference between the small planets, in addition to size, is the possibility of having a heavy atmosphere, without volatile elements such as hydrogen and helium.

The best information for a first-order characterization of extrasolar planets is their bulk properties: mass, radius, and density. By comparing the relationships between these properties it is possible to observe how in the case of giant planets there is a linear relationship while in the case of small planets, in particular between sub-Neptunes and super-Earths, there are degeneracies, as visible in the Figure (1.2).

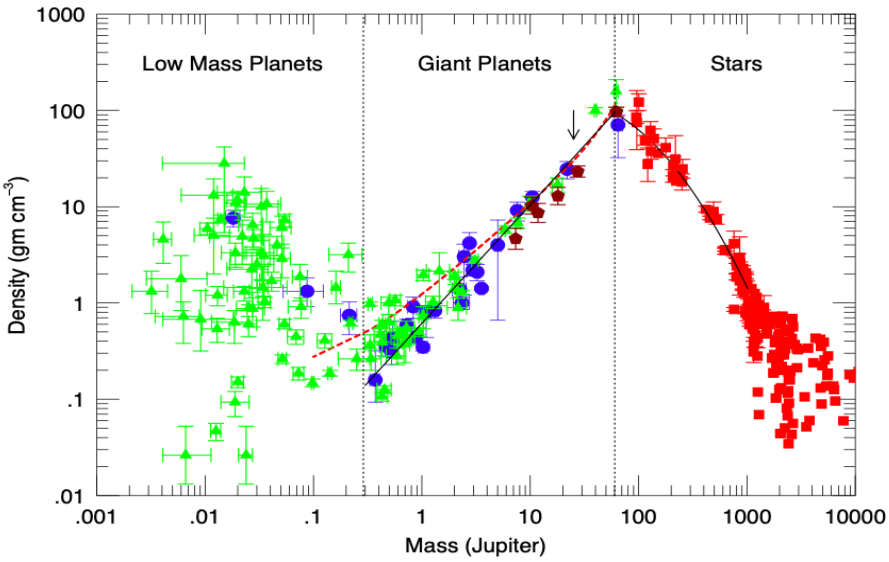

Figure 1.2: Mass-density relation for giants, low-mass planets and stars. Red squares represent stars, green triangles are Kepler's discoveries, blue dots are CoRoT exoplanets and brown pentagons are ground-based discoveries for high-mass giant planets. The dashed red line shows the mass-density relationship for H/He-dominated giant planets. The black central line represents a linear fit to the giant planets and brown dwarfs in the mass range $M = 0.35 - 60\ M_J$. The black curved line is made to the lower end of the stellar main sequence. Credits: [42]

An important characteristic of small planets is their habitability, which first depends on the position of a planet concerning its host star, and also on the star and the planet's features. Planetary characteristics that can influence the habitability of the planet depend particularly on its atmosphere [77]. The first detection of a planetary atmosphere was made by Charbonneau et al. (2002) [17]. They detected a photometric dimming during the transit of HD 209458 in a bandpass centred on the sodium resonance doublet at 589.3 nm and interpreted this dimming as absorption from sodium in the planetary atmosphere.

The information obtained about an exoplanetary atmosphere depends on the observations. The detections of exoplanetary atmospheric spectra are made using a variety of techniques: transit spectroscopy, direct imaging, and high-resolution spectroscopy, both from space as well as the ground.

The characterization of the atmospheres is possible thanks to various models capable of describing the physicochemical processes in exoplanetary atmospheres, spanning chemical and radiative processes, atmospheric dynamics, atmospheric escape, and clouds/hazes. The derived chemical abundances are also being used to investigate constraints on planetary formation and evolutionary processes, particularly related to the difference between primary and secondary atmosphere [60].

To investigate the atmospheric features observed in the planetary spectra, different atmospheric retrieval techniques are implemented. These permit to obtaining of detailed statistical estimates of atmospheric properties, like molecular and atomic species, elemental ratios (like He/H and C/O), pressure-temperature profiles, clouds or hazes.

There are two main reasons why it is essential to study exoplanetary atmospheres:

1. *planetary formation and evolution.* The various theories implemented to explain the formation of exoplanets consider the last step the collection of gas from the disk to form atmospheres, by growing planets. The chemical composition of a planet depends on the location and time of its formation in the disk as well as the relative amounts of gas and solids it accretes during its formation. Since the planet may migrate through the disk during formation, the net composition is governed by the cumulative accretion history of the planet over its migration pathway. Understanding planetary composition permits us to derive the formation history of the planet and the processes that occurred over time. Also, the comparison with the Solar System planets helps to understand the diversity of planetary environments and the various factors that influence atmospheric conditions, such as distance from the star, planetary mass, and geological activity [58].

2. *habitability.* The search for biosignatures is a key aspect of astrobiology. Studying the atmospheres permits us to identify planets that could potentially support life. In particular planets in the so-called "habitable zone" [51], are capable of supporting liquid water, an important factor related to life on Earth. Besides water, other molecules are indicative of biological processes, such as oxygen (O_2), carbon dioxide (CO_2) or methane (CH_2) [76]. To constrain the habitability of a planet is also useful to derive information about its climate and weather patterns, in particular the variations in temperature, cloud formation, and wind patterns. In this way, it is possible to understand the long-term stability of the atmosphere.

Observations of exoplanets to study their atmospheres are done using space-based and ground-based surveys. The Wide Field Camera 3 on board the space telescope has been a crucial instrument to study exoplanetary atmosphere in transmission and emission spectroscopy. TESS and CHEOPS, on the other hand, can help to characterize exoplanetary atmosphere using photometry. The most recently launched telescope is the James Webb Space Telescope (JWST) which provides information about exoplanet atmospheres, especially in the infrared spectrum. In the late 2020s, new missions like PLATO, which will search for Earth-like planets, and ARIEL, which will focus on studying exoplanet atmospheres, are planned. On the ground, the Very Large Telescope (VLT) in Chile uses advanced instruments to detect specific gases in exoplanet atmospheres. The next generation of ground-based telescopes, such as the Extremely Large Telescope (E-ELT), will observe exoplanets with incredible precision to understand the atmospheres of potentially habitable exoplanets and detect signs of life [58].

This thesis is dedicated to a detailed analysis and description of the atmosphere of the small planet K2-18b. The aim is to determine whether K2-18b has a primary atmosphere, directly accreted from the protoplanetary nebula, or a secondary atmosphere, formed from outgassed volatiles. To achieve this, I have conducted a new analysis of the planet's atmospheric characteristics, focusing on identifying key atmospheric components and processes. The findings will be compared with previous studies, particularly those by Madhusudhan et al. (2023) [63] and Tsiaras et al. (2019) [88], to provide a complete understanding of K2-18b's atmosphere and its formation history.

In Chapter 2, I introduce the characteristics of small planets, focusing on the differences between super-Earths and sub-Neptunes to determine which category K2-18b falls into. I also examine the concepts of primary and secondary atmospheres, as this work aims to understand whether K2-18b has a primary or secondary atmosphere. Additionally, I gather and summarize existing research on this planet from the literature. This review is necessary to consolidate current knowledge and provide a solid foundation for the analysis presented in this work.

In Chapter 3, I illustrate the transit method and transit spectroscopy, key techniques used to obtain the atmospheric spectrum of K2-18b. I also offer a brief introduction to the fundamental physics governing planetary atmospheres, with a focus on concepts relevant to K2-18b. It provides the necessary background for analyzing the planet's atmospheric properties.

In Chapter 4, I explain the tool and methodology employed in the analysis of K2-18b's atmosphere, with a particular focus on the TauREx [2] framework. I consider the theoretical explanation of the Bayesian analysis and the nested sampling, focusing on the MULTINEST [32] algorithm.

In Chapter 5, I present the transit data acquired with the James Webb Space Telescope (JWST) and reduced by Madhusudhan et al. (2023) [63], to obtain the transmission spectrum of the planet. Then, I explain the detailed analysis of K2-18b's atmospheric spectrum. I describe the various models and retrievals used to characterize the atmosphere, including statistical comparisons to interpret the data. The goal is to understand the chemical composition, structure, and profiles of K2-18b's atmosphere.

In the final Chapter 6 I summarize the findings of this thesis. I present the conclusive results of the atmospheric characterization and discuss their implications.

Chapter 2

Small planets and K2-18b

2.1 Small planets

As of today, thanks to the Kepler mission, we know that about 50% of stars beyond our Solar System host a small planet in the habitable zone [5]. The small planets can be divided into Earth-sizes, super-Earths and sub-Neptunes.

Earth-size planets are similar in dimension to Earth, with a radius between 0.5 and 1.25 $R_\oplus$, composed of rock, silicate, water or Carbon, such as our planet [26]. Super-Earth planets are more massive than Earth but lighter than Neptune, between twice the size of Earth and up to 10 times its mass. These planets are typically believed to be rocky planets with a solid surface, similar to Earth, but they can also have significant amounts of water or ice. They generally have a radius between 1.25 and approximately 1.75 $R_\oplus$ with a secondary atmosphere lacking H and He [79]. Super-Earths can be divided into ocean planets and massive terrestrial planets [28]. Ocean planets have a lot of water, around 10% or more, either in liquid or solid form. The expected primary volatiles on ocean planets are H_2O, NH_3 and CO_2. In particular, ammonia is very sensitive to UV and can be photodissociated and converted into N_2 and H_2, but the produced hydrogen is subject to hydrodynamical escape and can move away a fraction of nitrogen [54]. Massive terrestrial planets, on the other hand, have less surface water or water within their rocks, making them more similar to Earth. Sub-Neptunes are a category of exoplanets that exhibit characteristics similar to Neptune but are smaller in size and mass. These planets typically have radii larger than that of Earth but smaller than Neptune, between 1.75 and ~ 3 $R_\oplus$. They consist of a rocky core surrounded by a substantial envelope of volatiles, such as water, hydrogen, and helium. In particular, the H/He envelope dominates a sub-Neptune planet's size regardless of the abundance of other elements. As a result, for these planets, any degeneracies between rock, water, and iron are secondary to the overall distribution of material between the H/He envelope and heavier elements [56]. These planets are often located farther from their host stars compared to super-Earths [79].

The occurrence rate of sub-Neptunes and super-Earths (0.5–4 $R_\oplus$) with orbital periods shorter than 50 days is 0.9 planets per star, while the occurrence rate of Earth-sizes (0.5–1.4 $R_\oplus$) is 0.51 planets per star [24]. Due to the higher occurrence rate, planets intermediate in size between the Earth and Neptune (1-4 $R_\oplus$ and ≤ 20 $M_\oplus$) with a shorter period ($P < 100$ days) are the most common type of planet discovered by exoplanet surveys [8].

A peculiarity related to the small planets is the presence of a planet occurrence rate deficit, or a gap, between 1.5 and 2.0 $R_\oplus$, centred at $\sim 1.6\ R_\oplus$ for M and K dwarfs, in the small and short-orbit (≤ 100 days) population. This gap is responsible for dividing super-Earths from sub-Neptunes [30]. The separation in the radius distribution between the bare cores and the planets that still hold an envelope occurs because the presence of the envelope has a large impact on the observed radius, as visible in the Figure (2.1).

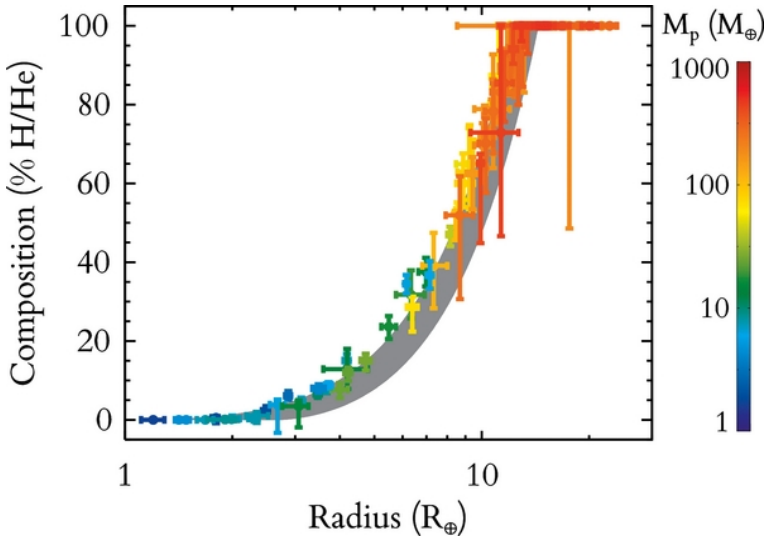

Figure 2.1: Correlation between H/He envelope fraction and planet radius for 200 transiting planets. The grey-shaded region shows the effect of varying the water abundance of the interior, which lowers the amount of H/He at a given radius. It is easy to observe that by increasing the radius (and the mass), from super-Earths to sub-Neptunes, the H/He envelope fraction increases. Credits: [56]

A possible explanation for this gap can be found in the photoevaporation of the H/He envelope of the short-orbit sub-Neptunes, due to the high energy radiation (XUV at $h\nu \sim 0.01$–1 keV) from their host stars. The atmospheric escape and, as a consequence, the separation in the radius distribution occurs in a relatively short timescale of the order of $\sim 10^5$ yr. The photoevaporation can explain the gap due to the H/He envelope loss of most super-Earths, although the additional core-powered mass-loss mechanism appears necessary in some planets. This loss is driven by the planet's internal luminosity, whose source is the planet's primordial energy from the formation. This energy can be comparable to or even higher than the gravitational binding energy of its atmosphere [30]. Similarly, the bolometric irradiation from the star drives the hydrodynamic outflow, albeit a cooler and slower one.

To discriminate which process between photoevaporation or core-powered mass-loss is dominant, it is necessary to indicate the position of the sonic surface, relative to the penetration surface of XUV photons, as indicated in the Figure (2.2). The sonic surface is the boundary within a planet's atmosphere where the speed of the outflow reaches the speed of sound. (i) Core-powered mass loss occurs when the sonic surface is positioned inward to the penetration surface of XUV photons, which thus do not affect the outflow. (ii) When the sonic surface appears outward from the XUV surface the absorption area of the planet to XUV photons becomes larger and the photoevaporation is enhanced. (iii) As the XUV luminosity increases, the upper atmosphere is XUV dominated and mass loss acts as classic photoevaporation [6].

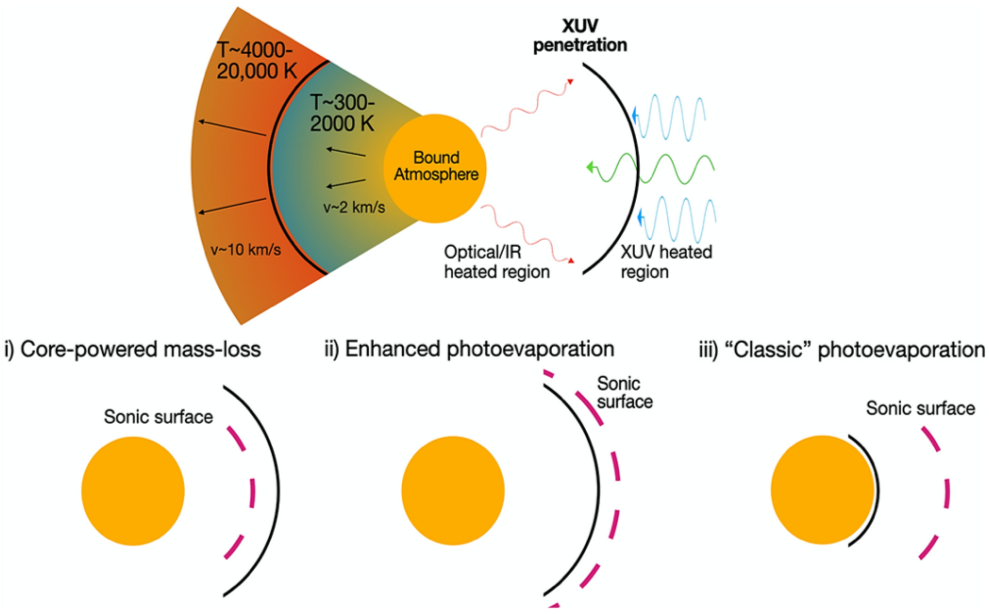

Figure 2.2: Scheme of the expected outflow structure for hydrodynamic mass loss and expected mass-loss regimes. The top panel shows the three layers of the planetary atmosphere. The bound atmosphere (yellow) is where the hydrodynamic outflow is weak. The middle layer (blue/green) is heated by cooling from the planet's interior (red photons) and the star's bolometric light (green photons). The outermost layer (orange) is heated by the star's XUV radiation (blue photons) and is several thousand Kelvin or more. The mass-loss regimes are shown from left to right as a function of increasing XUV luminosity. Credits: [56]

The theories illustrated above to explain the gap in the radius distribution are based only on pure evolution models, For these models the separation between super-Earths and sub-Neptunes can only be reproduced if the planet cores are rocky, implying a formation inside the water ice line for both classes of planets. Combining formation and evolution models indicate that both sub-Neptunes and super-Earths can also be water worlds formed beyond the water ice line and migrated inside [68].

Water words formed beyond the water ice line are typically located around M-dwarfs. This is because the water ice lines are closer to these stars, and planets around smaller stars tend to move inward more easily [68]. M-dwarf stars are the smallest stars capable of fusing hydrogen into helium and the most common type of main sequence. They define $\sim$ 70% of all stars in the Galaxy and have masses ranging from about 0.08 to 0.5 times that the mass of the Sun. M-dwarfs have temperatures between 2,400 and 3,700 K and are cool enough that their spectra are heavily veiled throughout by molecular bands of titanium oxide (TiO), hydrogen (H_2), water (H_2O), and vanadium oxide (VO) [77].

2.1.1 Formation and evolution

There are two main theories about how small planets form, both implicate the inward movement of material in gas-rich disks but at different scales. The drift model describes a close-in growth of pebbles, which form from dust grains and become large enough to partially decouple from the gas and drift inward. They often get trapped in pressure bumps in the inner disk, where they form planetary embryos that collide and migrate to form planets.

The migration model defines large planetary cores that form further out in the disk, often near the snow line where pebble growth is faster. These cores then migrate inward, forming planets through collisions. In both models, growing planets collect gas from the disk to form atmospheres. After the gas disk dissipates, the atmospheres experience loss processes, as described above. The models predict different compositions for close-in planets: the drift model leads to rocky, water-poor planets, while the migration model results in a variety of volatile elements [6].

Planetary atmospheres can be classified as "primary" or "secondary", reflecting different formation scenarios. The formation of primary atmospheres is typically explained by the core-accretion theory, where a solid core attracts a gaseous envelope from the protoplanetary disk. Primary atmospheres directly originate from the surrounding nebula during a planet's formation. In contrast, secondary atmospheres form from volatiles outgassed from the planet's interior after its initial formation. For small planets, there is a strong expectation that their atmospheres potentially consist of a mix of primary and secondary components [60].

The primary atmosphere is composed mostly of light elements such as hydrogen and helium. This mixture of gases is similar to the composition of the star around which the planet orbits because it was developed at the initial stages of the formation of the star. The secondary atmosphere, instead, is composed of heavy elements due to internal volcanic outgassing or by accumulation of material from comet impacts, that remain after losing the light elements. This loss is because the light particles' velocity is higher than the escape velocity of the planet for two main reasons: their low mass and the surface temperature of the planet that increases the kinematic energy and, as a consequence, the velocity of the light particles. If the loss of the light elements is partial there is the formation of a hybrid atmosphere, composed of light and heavy elements, as it is schematized in the Figure (2.3). The secondary atmospheres are characteristic of planets close to their host star and they are thin compared to primary atmospheres [84].

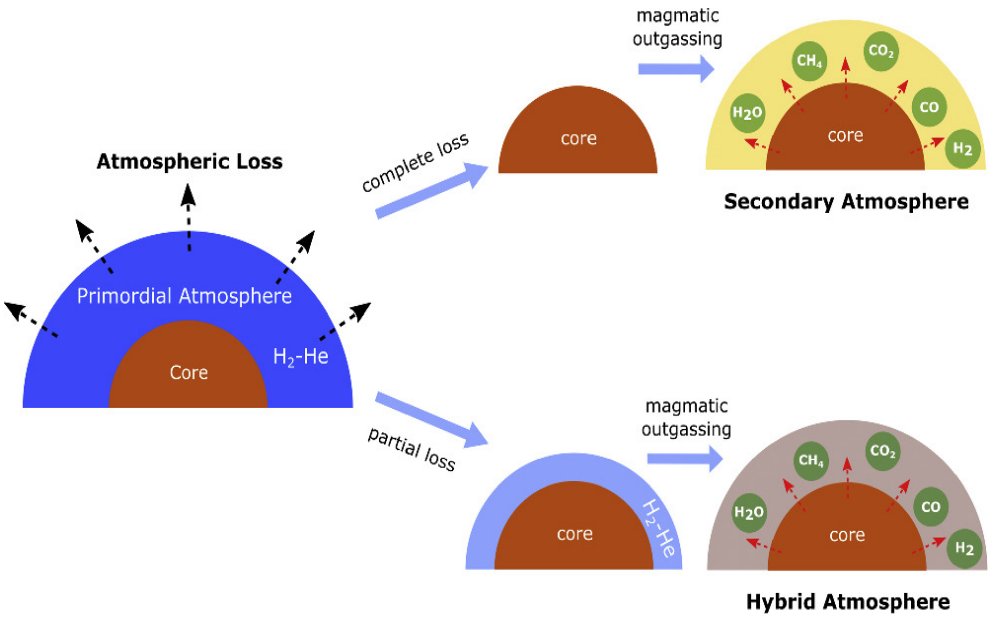

Figure 2.3: Scheme of primordial, secondary and hybrid atmosphere. The difference between the secondary and the hybrid one is due to the atmospheric loss of light elements. Credits: [84]

In our Solar System, the terrestrial planets have secondary atmospheres because of their vicinity to the Sun, mainly composed by outgassing from their mantles. Instead, gas giants like Jupiter and Saturn have primary atmospheres with compositions similar to the chemistry of the nebula out of which they formed [84]. Interestingly, Uranus and Neptune, despite being classified as gas giants, have atmospheres that are only $10 - 20\%$ hydrogen and helium by mass, suggesting a possible minor secondary component [60]. Hybrid atmospheres are not present in our Solar System, but they are a possible explanation for many exoplanetary atmospheres, particularly of super-Earths and sub-Neptunes.

2.1.2 Habitability

An important characteristic related to small planets concerns their distance from the host star, which can be responsible for classifying a planet as potentially habitable, as visible in the Figure (2.4). These planets are located in the so-called "habitable zone" (HZ), a range of orbits around the planet's star within which a planetary surface can support liquid water given sufficient atmospheric pressure [51].

The bounds of the habitable zone are based on Earth's position in the Solar System and the amount of radiant energy it receives from the Sun [50]. The inner edge of the habitable zone (IHZ) is defined considering the distance from the star at which the planet risks losing its entire water products to space in a runaway greenhouse effect. The outer edge (OHZ), instead, is based on the maximum CO_2 greenhouse limit. The traditional boundaries of the habitable zone are based on the assumption of a carbonate-silicate cycle similar to Earth's, where the rate of silicate weathering increases with the planet's surface temperature.

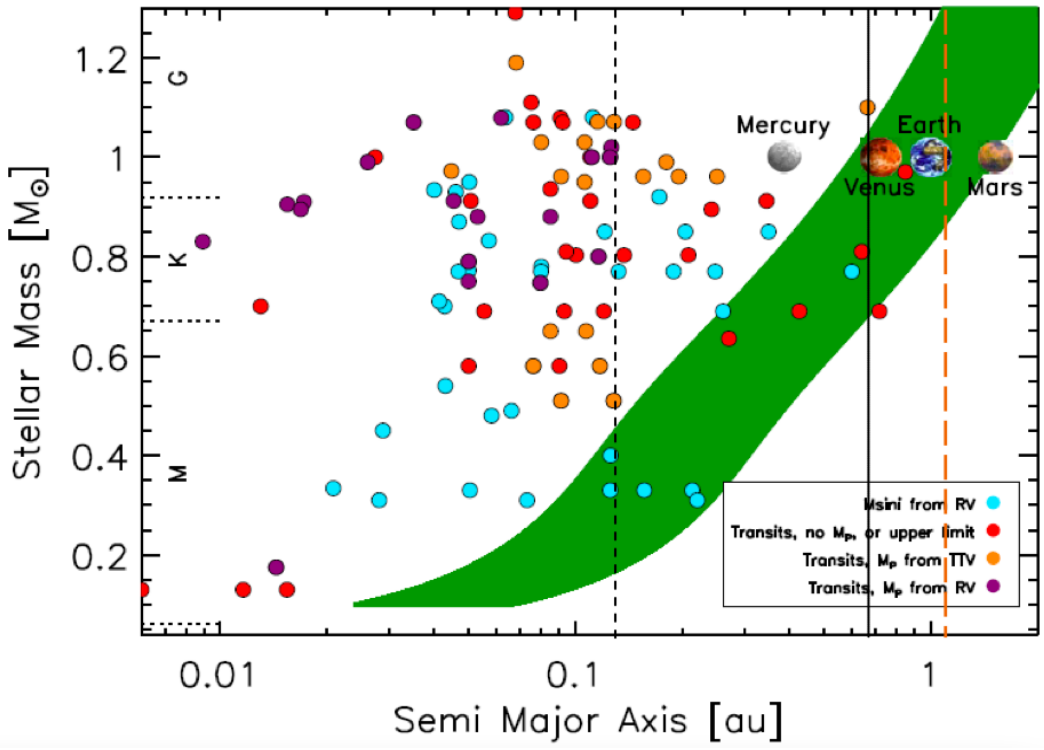

Figure 2.4: Distribution of super-Earth exoplanets for different host star masses in function to the semi-major axis. The green region represents the habitable zone. Credits: [73]

Due to the importance of liquid water to Earth's biosphere, the objects in the HZ are the best candidates to explore potentially habitable environments outside our Solar Systems, by analyzing their chemical composition and thermal structure.

Keeping liquid water on a planet's surface for a long time requires a stable climate with the right temperatures and pressures. The climate of a planet is mainly influenced by the amount of radiation it receives from its star and how the planet's atmosphere and surface respond to that energy. The global energy balance of a planet is described by its equilibrium temperature. However, the planet's atmosphere can affect its actual surface temperature, making the difference between a frozen planet and one with liquid water. Atmospheric molecules like carbon dioxide (CO_2), water vapour (H_2O), methane (CH_4), and ozone (O_3) can absorb a significant amount of incoming stellar radiation. In particular, dense CO_2 atmospheres can increase surface temperatures by absorbing near-infrared radiation, which could expand a planet's habitable area. However, too much CO_2 can lead to condensation and cooling effects, particularly near the outer edge of a star's habitable zone. Other ways to warm distant planets, such as thick hydrogen envelopes, could allow for liquid water at lower radiation levels, but these could also raise surface temperatures too high for habitability [77]. Another factor capable of affecting the planetary climate is the change in ocean/rocky fraction. In particular land planets orbiting Sun-like stars are less exposed to episodes of global-scale glaciation, so-called "snowball states", with respect to the water planets, due to the lower thermal inertia of land and drier atmospheres. Also, hot planets dominated by oceans might lose all their water before the silicate weathering process can stabilize the climate, making them less likely to support life [33].

The habitable zone varies depending on the type of star, as visible in the Figure (2.4). For example, the HZ around an M-type star is only about one-fifth to one-fifteenth the size of the HZ around a G-type star. This variation is not projected on the probability of finding small planets in the HZ of these types of stars. The estimation of the number of small planets in the habitable zones of stars in the Milky Way is around 45.5 billion, with 11.548 billion around M-type stars, 12.930 billion around K-type stars, 7.622 billion around G-type stars, and 5.556 billion around F-type stars [40].

Observing exoplanets within the habitable zone of M-dwarf stars presents several advantages. Firstly, the habitable zones around M-dwarfs are closer to the star, which means that planets in these zones complete their orbits more quickly. This increases the frequency of transit events, where the planet passes in front of its host star. Moreover, the smaller size and lower luminosity of M-dwarf stars mean that the transit signals of planets passing in front of them are relatively large compared to the planet's size. This makes it easier to detect and study the atmospheres of these planets. In particular, planets with Earth-like atmospheres around M-dwarf stars tend to absorb more radiation because their atmospheres have lower albedos, leading to higher surface temperatures [77].

Astrobiology

The field responsible for studying the life beyond Earth is the astrobiology. It focuses on what conditions are needed for life to exist and how it might develop, analysing the atmospheres. Life, as we know it on Earth, requires three things: liquid water, an environment for the formation of organic molecules from bioessential elements such as sulfur, phosphorous, oxygen, nitrogen, carbon, hydrogen, and an energy source, whether stellar or chemical [22].

The search for biosignature is a key aspect of astrobiology. The atmospheric investigation of the exoplanets focuses on identifying molecules present in the so-called "Earth-like" atmosphere that could indicate biological activity. An Earth-like atmosphere is, by definition, characterized by the presence of high-molecular-weight gases ($\mu >> 2$) that include condensable greenhouse gas (H_2O), noncondensable greenhouse gas (CO_2), and a noncondensable background gas (N_2). It is important to underline that an Earth-like atmosphere is not the only type of habitable atmosphere conceivable for small planets because an alternative possibility is an H_2-dominated atmosphere [75].

When looking at small planets, especially those orbiting M-dwarf stars, astrobiology allows us to understand if these planets could support life. M-dwarf planets often encounter difficult conditions, like strong radiation from their star and higher temperature differences between the day and night sides. The presence of oxygen-rich atmospheres could allow for complex life, though the extremes in temperature and radiation pose challenges. However, planets with thicker atmospheres or Earth-like rotation may offer more stable conditions.

On M-dwarf planets, life forms might employ mechanisms similar to Earth's "extremophiles", organisms able to survive at extreme temperatures, intense pressure, high radiation, and limited water availability. Another life-forms found on the Earth are the plants, which produce photosynthesis, the process where plants use light energy, carbon dioxide (CO_2), and water (H_2O) to produce carbohydrates and oxygen. The evolution of photosynthesis is thought to have started when organisms adapted from deep-sea hydrothermal vents to shallower waters, where solar energy could be utilized. On Earth, plants capture about 48% of incoming solar radiation, especially in the 400-700 nm range known as Photosynthetically Active Radiation (PAR). In contrast, M-dwarf stars emit ten times less PAR than on Earth at visible wavelengths, and their planets might have less effective photosynthesis due to lower photon availability. Despite this, adaptations in photosynthetic organisms, such as absorbing in near-infrared wavelengths, could potentially make photosynthesis possible on M-dwarf planets [77].

2.2 Planet K2-18b

One of the most interesting targets for the atmospheric characterization of a M-dwarf habitable zone exoplanet is K2-18b. This small planet was first discovered in 2015 during the K2 mission Campaign 1 of the Kepler Space Telescope [34]. It orbits an M-dwarf star, named EPIC 201912552 or K2-18. The star is a M2.8 type star with a $K_{mag} = 8.9$. It is located in the constellation Leo with coordinates right ascension 11h 30m 14.7s and declination +07° 18' 05" and its distance from Earth is approximately 124 light-years [66]. The bulk characteristics of the star K2-18 are reported in the Table (2.1). The K2-18 planetary system has a gyrochronological age of 2.4 ± 0.6 Gyr. This age is in agreement with solar-like metal abundance and the star's low U, V, and W space motions of -1.37, -5.14, -12.40 km/s [39].

Table 2.1: Stellar properties of K2-18, by Montet et al. (2015) [66].

Parameters of star K2-18				
Mass ($M_\odot$)	Radius ($R_\odot$)	T_{eff} (K)	[Fe/H] (dex)	Distance (pc)
$0.413^{+0.043}_{-0.043}$	$0.394^{+0.038}_{-0.038}$	3503^{+60}_{-60}	$0.09^{+0.09}_{-0.09}$	34^{+4}_{-4}

The planetary system contains for sure the planet K2-18b but additional studies have hypothesised the presence of a second planet named K2-18c [18]. This candidate planet seems to be a warm super-Earth, smaller than K2-18b, with a minimum mass $m\sin(i)$ of 5.62 ± 0.84 $M_\oplus$ and orbits much closer to the star. The information about the planet K2-18c was obtained through radial velocity measurements using HARPS, which display a 9-day signal related to the presence of a second planet in the system [19]. K2-18c is a non-transiting planet, in particular, studies suggest that its orbit is slightly tilted compared to K2-18b's, which could explain why is not possible to see it passing in front of its star [18].

K2-18b, instead, is a transiting planet which was discovered using the transit method. A theoretical description of this technique is reported in Section (3.1). During the K2 Campaign 1 mission, two transit occurrences of the planet were identified [34]. From these data, it was possible to deduce parameters like the ratio between planet and star radius, the period and the planet's magnitude. To refine the characteristics of K2-18b, additional parameters were verified and fine-tuned by observing a third transit event that was compared with the two obtained by the K2 mission. The identification of a third transit was done using the Spitzer/IRAC instrument [10]. It permitted the determination of parameters such as the radius of the planet, its orbital inclination, and its semi-major axis. The most recent mass value of the planet was obtained using precision radial velocity measurements taken with a technique known as "line-by-line". This method is designed to extract radial velocity information from high-resolution spectra remaining robust against outliers. It was applied to archived data from the HARPS and CARMENES spectrographs and, by combining data from both, it was possible to enhance the accuracy of the mass and the period of K2-18b [72] and to obtain the equilibrium temperature, adopting an Earth-like albedo of 0.3. The parameters of this planet are reported in the Table (2.2).

Table 2.2: Parameters of the planet K2-18b. Mass and period measured by Radica et al. (2022) [72]; radius, semi-major axis and orbital inclination by Benneke et al. (2017) [10]; the ratio of planet and star radius and magnitude by Foreman-Mackey et al. (2015) [34]; planet density by Cloutier et al. (2017) [18] and equilibrium temperature by Cloutier et al. (2019) [19].

Parameters of planet K2-18b	
Mass ($M_\oplus$)	8.63 ± 1.35
Period (days)	32.9396 ± 10^{-4}
Radius ($R_\oplus$)	2.28 ± 0.03
$R_p/R_\star$	$0.051^{+0.004}_{-0.006}$
Semi-major axis (AU)	0.143 ± 0.006
Orbital inclination (deg)	$89.578^{+0.008}_{-0.009}$
Density (g/cm^3)	3.3 ± 1.2
Kepler mag	12.47
Temperature (K)	265 ± 5

The planet K2-18b has a distance from the star such that it falls within its habitable zone. Its orbital period is about 33 days around an M-dwarf [72], which is cooler and smaller than our Sun, with only 2.53% of the Sun's luminosity. This allows K2-18b to receive a flux from the star of 1368^{+114}_{-107} W m^{-2}, similar to what Earth receives from the Sun [11].

This confirms the position of K2-18b in the habitable zone of its star.

Considering the size of the planet there are some doubts about its nature. In particular, its radius of 2.28 $R_\oplus$ [10] falls very close to the upper limit of the gap (between 1.5 and 2.0 $R_\oplus$) responsible for distinguishing a planet between a super-Earth or a sub-Neptune, as described in Section (2.1). Another, and more recent, category which could be used to describe K2-18b is the class of potentially habitable planets known as "Hycean worlds". These planets have water-rich interiors with oceans under hydrogen-dominated atmospheres. Their densities are between those of rocky super-Earths and more extended sub-Neptunes. In particular, they are larger than Earth-like habitable planets and have radii between 1-2.6 $R_\oplus$ for masses between 1-10 $M_\oplus$. These planets have rocky cores of at least 10% of their mass.

Hycean planets also extend the traditional habitable zone (HZ), as visible in the Figure (2.5), since they can remain habitable with little or no stellar radiation. These so-called "Cold Hycean" worlds could be habitable even without direct sunlight. Some may also have permanent night sides that remain habitable, despite high temperatures on the day side—these are known as "Dark Hycean" planets [62]. The planet K2-18b is marked in the Figure (2.5) by a red arrow. It can be observed that it is located at the edge of the habitable zone for terrestrial planets and Hycean worlds.

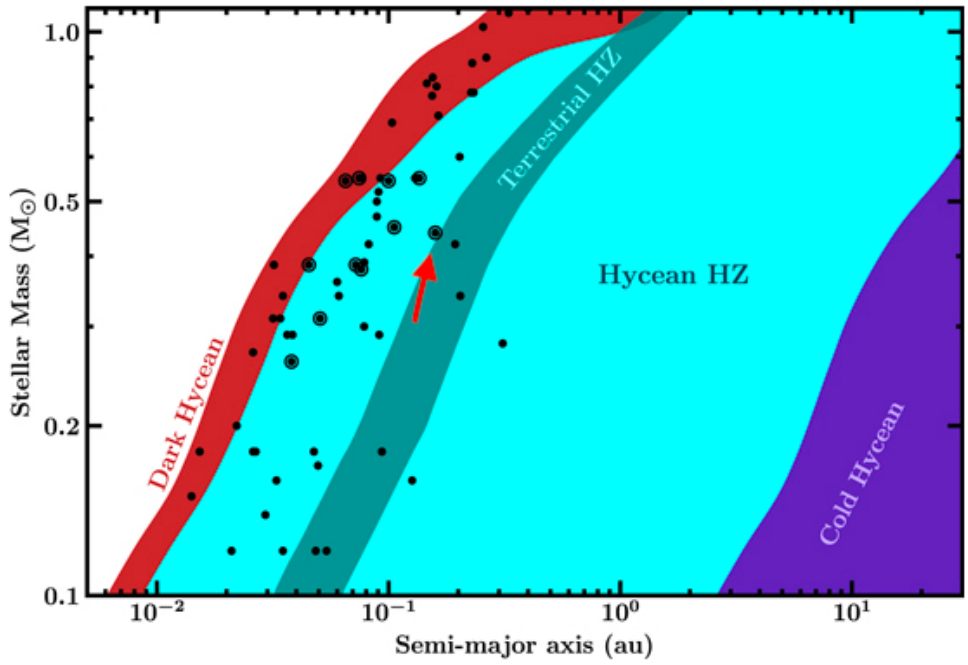

Figure 2.5: Habitable zones for different types of planets. The red arrow indicates the position of the planet K2-18b. Credits: [62]

The atmospheric composition of a Hycean world is mainly dominated by hydrogen and helium (H_2/He), along with water (H_2O). In addition to these molecules, methane (CH_4) and ammonia (NH_3) are also likely to be present. Since these planets are located in the habitable zone and have larger radii and lower gravity compared to rocky super-Earths, they offer better chances for detecting potential biomarkers like dimethyl sulfide (DMS), carbon disulfide (CS_2), methyl chloride (CH_3Cl), carbonyl sulfide (OCS), and nitrous oxide (N_2O).

A Hycean planet typically has an atmosphere rich in hydrogen, a water layer that accounts for 10% to 90% of its mass, a habitable surface, and an iron and rocky core making up at least 10% of its mass. These characteristics suggest that K2-18b could be a strong candidate for classification as a Hycean world.

Given its size and position, K2-18b could be classified as a super-Earth, a sub-Neptune, or a Hycean planet. To determine the most accurate classification, it is necessary to analyze the planet's atmosphere by reviewing studies in the literature related to K2-18b.

2.2.1 Review of K2-18b's atmosphere in existing studies

The composition and structure of K2-18b's atmosphere have been topics of intense assumption and discussion. The study of this atmosphere was conducted using transit spectroscopy, deriving the planet's transmission spectrum from information obtained through transits in front of the star. The theoretical discussion regarding transit spectroscopy is reported in Section (3.2).

The first data used to obtain a transmission spectrum of K2-18b were collected by the Wide Field Camera 3 (WFC3) on board the Hubble Space Telescope (HST). Nine transits of K2-18b were observed within 3 years. Each transit was observed during five HST orbits, with the G141 infrared grism in a wavelength range of 1.1-1.7 μm. The atmospheric analysis using these data was conducted by Tsiaras et al. (2019) [88] and Benneke et al. (2019) [11].

Tsiaras et al. (2019) [88] investigated three scenarios to explain the atmospheric spectrum obtained from data: (1) a cloud-free atmosphere containing only H_2O and H_2–He, (2) a cloud-free atmosphere containing H_2O, H_2–He and N_2 and (3) a cloudy atmosphere containing only H_2O and H_2–He. They conducted the analysis using the TauREx [1] framework. The statistical values achieved were too similar to distinguish between the three scenarios but were enough strong to confirm the presence of water vapour in the atmosphere in all the cases. In particular, for the H_2O and H_2–He case, the abundance of H_2O was $20-50\%$. These results indicated that a non-negligible fraction of the atmosphere was still made of H_2–He.

Benneke et al. (2019) [11] extended the HST/WFC3 observations with two Spitzer transit observations between 3.15–3.94 μm, one Spitzer observation from 3.96–5.02 μm, and a Kepler/K2 transit observation from 0.43–0.89 μm. Their model considered molecules such as H_2O, CH_4, CO, CO_2, NH_3, HCN, and N_2, along with H_2/He. They also included a cloud layer assumed to be opaque below the cloud top pressure, where large droplets would form. The analysis, conducted using the SCARLET atmospheric retrieval framework [9], revealed a water absorption feature at 1.4 μm in the HST/WFC3 data, indicating absorption above the clouds in the mid-atmosphere. The retrieval model showed that the data were best explained by a hydrogen-dominated atmosphere, which agrees with the planet's density reported in the Table (2.2).

Using the transmission spectrum reduced by Benneke et al. (2019) [11], Madhusudhan et al. (2020) [61] retrieved the atmospheric properties of K2-18b by applying an adaptation of the AURA retrieval code [71]. They conducted retrievals for four model configurations: (1) a full model including inhomogeneous clouds and hazes, (2) a clear atmosphere, (3) an atmosphere with an opaque cloud deck but no hazes, and (4) an atmosphere with inhomogeneous clouds but no hazes.

They confirmed the high-confidence detection of water in an H_2-rich atmosphere as reported by Benneke et al. (2019) [11] and Tsiaras et al. (2019) [88], but they did not find strong evidence for clouds/hazes in the atmosphere.
In addition to analyzing the atmosphere of K2-18b, they also analyzed its internal structure using its bulk properties and hypothesizing three possible composition and origin scenarios of the planet from the results obtained: a rocky world with a massive rocky core overlaid by an H/He envelope. Such a scenario is consistent with either H_2 outgassing from the interior or accretion of an H_2-rich envelope during formation; a sub-Neptune with a non-negligible H/He envelope in addition to significant H_2O and core mass fractions such as 45% Earth-like core, akin to canonical models for Neptune and Uranus; and a water world full of water with a minimal H_2-rich atmosphere.

Changeat et al. (2021) [16] also retrieved three similar scenarios, underlying the difference between primary and secondary atmospheres.

1. Icy/Water World with a secondary atmosphere composed of hydrogen and helium with a mean molecular weight mainly explained by water vapour (up to 50% in volume mixing ratio).

2. Super-Earth with a secondary atmosphere composed of a small amount of water vapour, and a higher mean molecular weight due to one or more gases like nitrogen, along with hydrogen and helium.

3. Sub-Neptune with a cloudy primary atmosphere mostly made of hydrogen and helium, with a mean molecular weight < 4 amu, and small amounts of water vapour.

They demonstrated that the James Webb Space Telescope (JWST) and the future ARIEL space telescope will be able to differentiate between the three cases.

Other simulations of K2-18b's atmosphere, using data from Benneke et al. (2019) [11] and Tsiaras et al. (2019) [88], were conducted by Blain et al. (2021) [13] and Bézard et al. (2022) [12]. These studies found that methane (CH_4) absorption could either dominate or be comparable to H_2O absorption within the analyzed spectral range. They also suggested that the atmosphere may be highly enriched in metals, with levels ranging from 65 to 500 times the solar metal abundance when compared to Tsiaras et al. (2019) [88], or between 100 and 200 times the solar level when compared to Benneke et al. (2019) [11], assuming a solar C/O ratio. The simulations also indicated the possibility of either thick or minimal H_2O-ice clouds. Liquid water clouds might form on planets like K2-18b but only if they receive no more than 80% of the stellar radiation the planet currently gets. Whether clouds are present or not, K2-18b's Bond albedo was estimated to be approximately 0.02.

The most recent and extended dataset used to obtain the transmission spectrum of K2-18b was collected by the James Webb Space Telescope (JWST), using NIRISS SOSS and NIRSpec G395H instruments in the 0.9-5.2 μm wavelength range. A description of the data reduction is reported in Section (5.1).
Madhusudhan et al. (2023) [63] performed the retrieval of the atmosphere using the AURA retrieval code [71], considering molecules such as H_2O, CH_4, NH_3, HCN, CO and CO_2, in-homogeneous clouds/hazes at the day-night terminator region and five molecules, DMS, CS_2, CH_3Cl, OCS and N_2O, that have been suggested to be promising biomarkers in Hycean worlds. The mass, radius, and equilibrium temperature of K2-18b, obtained from the analysis, were consistent with multiple models, hypothesised by Madhusudhan et al. (2020) [61] and Changeat et al. (2021) [16].

However, the chemical composition of the atmosphere allowed differentiation between various scenarios. Madhusudhan et al. (2023) [63] reported strong detection of CH_4 and CO_2, both having high volume mixing ratios of around 1% in a hydrogen-rich atmosphere. Instead, they detected only upper limits on the abundances of H_2O, NH_3, CO, and HCN. The low level of water in the gas phase at pressures below about 100 mbar fits with the idea of condensation caused by a cold trap in the troposphere, similar to what happens in Earth's stratosphere. This suggests that while H_2O might be abundant below its condensation region, the transit observations of the planet's terminator region may not be deep enough to detect it. These results allowed a higher mean molecular weight, typical of a secondary atmosphere, and aligned with chemical expectations for a cold ocean under a thin hydrogen-rich atmosphere. Additionally, there was possible evidence for DMS (dimethyl sulfide), which has been suggested as a biomarker in both terrestrial and Hycean planets. These findings strengthen the idea that K2-18b could be a Hycean world with a secondary atmosphere and might even support biological activity.
The exclusion of a sub-Neptune model was because, in this model, any carbon and nitrogen compounds formed in the deep atmosphere would be converted back into their more stable forms, CH_4 and NH_3, and then transported to the upper atmosphere where are observable. One limitation of the analysis of Madhusudhan et al. (2023) [63] was that the photochemical models for sub-Neptunes assumed ideal gas behaviour and did not account for how a primordial hydrogen atmosphere might interact chemically with a supercritical water layer or silicate magma deep in the atmosphere. Similarly, the results were not compatible with a rocky planet having a shallow, hydrogen-rich atmosphere sitting above a solid surface, despite the observed CH_4 and CO_2 seeming consistent with such a model. Though some atmospheric CO_2 could be explained by interaction with a deep silicate mantle, this model does not typically predict a 1% CO_2 mixing ratio from such high-pressure conditions, nor does it account for the lack of nitrogen recycling back to NH_3. Overall, the planet's bulk density and the atmospheric composition indicated that K2-18b is a Hycean world rather than a rocky or volatile-rich planet with a deep hydrogen atmosphere, or a rocky planet with a thin hydrogen atmosphere.

Due to the limitation in the photochemical models used by Madhusudhan et al. (2023) [63], Wogan et al (2024) [91] simulated K2-18b as both a Hycean planet and a gas-rich sub-Neptune utilising more complex photochemical models. Their findings predicted an atmosphere with 4% CH_4 and almost 0.1% CO_2, which match the JWST data. In this scenario, CH_4 and CO_2 were formed deep in the atmosphere through thermochemical processes and were then brought up to the upper layers where are observable. The results of Madhusudhan et al. (2023) [63], instead, showed about 1% CH_4. There are two possible explanations for this result: either K2-18b is a Hycean planet with methane-producing life, or it is a sub-Neptune with no solid surface. In particular, a Hycean planet would need biogenic methane, similar to early Earth's microbial life, or an unknown source to maintain it against photochemical destruction.
Keeping a stable, temperate climate on a Hycean K2-18b is difficult due to the possibility of a runaway greenhouse effect unless clouds reflect enough starlight. Also, a thin hydrogen atmosphere might escape due to XUV radiation, and volcanism cannot replace it because the thick ice and ocean layer would prevent silicate melting. Wogan et al (2024) [91] favoured the sub-Neptune interpretation with respect to the Hycean world of the planet K2-18b because it does not need a biosphere or other unknown source of methane to explain the data.

From the JWST observations, CO_2 was detected in K2-18b's atmosphere, while ammonia (NH_3) was not, suggesting the possible presence of a liquid-water ocean under its hydrogen-rich atmosphere [63]. The atmosphere's features were visible considering a pressure level of 0.1 bar, but the surface of the planet could not be directly observed. This leaves uncertainty about whether the planet has a rocky, Neptune-like, or water-rich composition. In a deep hydrogen-rich atmosphere, NH_3 would usually be transported to the upper regions if the planet had a Neptune-like interior. However, if there is a liquid-water ocean, NH_3 could dissolve into it and become hard to detect. An alternative explanation for the low NH_3 levels is that K2-18b has a hydrogen-rich atmosphere over a magma ocean. Nitrogen can dissolve in magma, which would explain the low ammonia levels. The interaction between the atmosphere and a silicate magma ocean could mimic the effects seen in a planet with a shallow surface, stopping the recycling of NH_3 back into the atmosphere. Shorttle et al (2024) [78] explored the impact of a magma ocean on the atmospheric chemistry of the sub-Neptune K2-18b. They found that the lack of NH_3 in the atmosphere can be explained by nitrogen dissolving in magma under certain conditions, especially when a thick hydrogen atmosphere interacts with a molten surface. This magma ocean model matches the JWST data for the planet within 3σ, making it as likely an explanation as the presence of a liquid-water ocean. Further observations, particularly in the region above 4 μm where CO_2 and CO features are strong, confirmed the magma ocean theory.

The latest studies conducted to analyze the atmosphere of K2-18b, reported above, establish that this planet can be classified as a sub-Neptune having a hydrogen-rich atmosphere. However, it remains to be clarified whether the atmosphere of K2-18b is primary or secondary. In this thesis work, the objective is to characterize the atmosphere of K2-18b and verify whether it is primary or secondary. To do this I used the most recent data collected by JWST and reduced by Madhusudhan et al. (2023) [63] and implemented a retrieval using TauREx [1], the same framework as Tsiaras et al. (2019) [88], described in Section (4.1). The various retrievals used to simulate the atmosphere are reported in Section (5.2). Analyzing the atmosphere can help us better understand the formation history of the planet K2-18b.

Chapter 3

Exoplanetary Transit and Atmospheric Physics

3.1 Transit Method

The transit method is used to detect exoplanets by observing a drop in the star's flux, caused by a planet that passes in front of it. The first transiting exoplanet was observed in 1999 [43] and today more than 4400 exoplanets[1] were discovered using this technique. In this section, I describe the transit method geometry and parameterization [70].

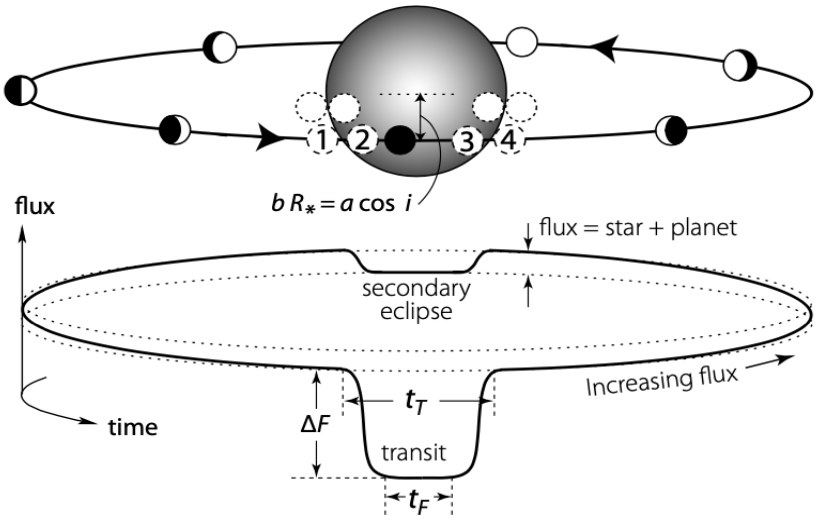

Figure 3.1: Scheme of the transit of a planet in front of its host star. This passage causes a drop (ΔF) in the stellar flux. The time between the first and fourth contact star-planet, labelled as 1 and 4 marks, is the total transit duration t_T, while the time between the second and third contact, labelled as 2 and 3 marks, corresponds to the transit flat time t_F. The formula $bR_\star = a\cos i$ represents the impact parameter b, where $R_\star$ is the radius of the star, a is the semi-major axis of the planet's orbit and i is its inclination. Credits: [70]

[1] https://exoplanet.eu/

It is possible to distinguish two main phases of the transit of a planet: primary and secondary transit, as visible in the Figure (3.1). A transit, or primary transit, occurs when the planet partially obscures the star, passing in front of it during inferior conjunction. The first and fourth contact describe, respectively, the moment in which the planet enters and leaves the stellar disk. The second and third contacts refer to the moments when the planet completely overlaps the star soon after the ingress and soon before the egress, respectively. Grazing transits happen when the projections of the planet and the star never fully overlap. An occultation, or secondary eclipse/transit, happens when the planet passes behind its star.

To determine the probability of detecting a transit, we consider the geometry of the system and the observer's position. A planet will be observed transiting if the observer is within a cone with an opening angle defined by $\sin\theta = (R_\star + R_p)/r$ for transits including grazing, or $\sin\theta = (R_\star - R_p)/r$ for only full transits, where $R_\star$ is the star's radius, R_p is the planet's radius, and r is the distance between the planet and the star at the time of transit. For a system with known eccentricity e and argument of periapsis ω, often derived from radial velocity measurements, the transit probability P_{tra} and the occultation probability P_{occ} are given by:

$$P_{tra} = \left(\frac{R_\star \pm R_p}{a}\right)\left(\frac{1 + e\sin\omega}{1 - e^2}\right) \qquad P_{occ} = \left(\frac{R_\star \pm R_p}{a}\right)\left(\frac{1 - e\sin\omega}{1 - e^2}\right) \tag{3.1}$$

Where a is the semi-major axis of the planet's orbit. In the case where the planet's radius is much smaller than the star's radius ($R_p \ll R_\star$) and the orbit is circular ($e = 0$), the transit and occultation probabilities simplify to:

$$P_{tra} = P_{occ} \sim 0.005\left(\frac{R_\star}{R_\odot}\right)\left(\frac{a}{1\,\mathrm{AU}}\right)^{-1} \tag{3.2}$$

This formula indicates that the probability of detecting a transit is proportional to the ratio of the star's radius to the planet's semi-major axis.

Several parameters can be used to characterize a transit: the impact parameter b, the transit depth ΔF, the transit duration t_T, and the transit flat time t_F.

The impact parameter b represents the projected distance between the centres of the planet and the star, during mid-transit. By the formula for the impact parameter is possible to derive the orbital inclination i of the planet:

$$bR_\star = a\cos i \tag{3.3}$$

where a is the semi-major axis of the planet's orbit and $R_\star$ is the star's radius.

The transit depth ΔF represents the decrease in the star's flux during a planetary transit. It is related to the radii of the star $R_\star$ and the planet R_p by the formula:

$$\Delta F = \frac{R_p^2}{R_\star^2} \tag{3.4}$$

The larger the planet, the greater the flux variation of the star. By measuring ΔF during the transit and knowing the star's radius, it is possible to determine the planet's radius.

The transit duration, t_T, is the time interval between the first and fourth contacts of the planet with the star. This duration represents the fraction of the orbital period, P, during which the projected distance between the centres of the star and the planet is less than the sum of their radii. The transit duration is directly proportional to the radius of the star and it is a function of the planetary orbital inclination i: it decreases to zero as the orbit becomes more inclined away from edge-on. The formula is:

$$t_T = \frac{P}{\pi} \arcsin\left(\frac{R_\star}{a} \left[\frac{(1 + R_p/R_\star)^2 - (a/R_\star \cos i)^2}{1 - \cos^2 i} \right]^{1/2} \right) \sim$$

$$\sim \frac{P R_\star}{\pi a} \sqrt{\left(1 + \frac{R_p}{R_\star}\right)^2 - \left(\frac{a}{R_\star} \cos i\right)^2} \tag{3.5}$$

The last approximation is valid for small inclination angles ($\cos i \ll 1$).
The impact parameter b and the scaled semi-major axis $a/R_\star$ are related to the transit duration t_T and ingress/egress duration τ through specific formulas. These relationships are given by:

$$b^2 = 1 - \sqrt{\Delta F} \frac{t_T}{\tau} \qquad\qquad \frac{R_\star}{a} = \frac{\pi}{\Delta F^{1/4}} \frac{\sqrt{t_T \tau}}{P} \left(\frac{1 + e \sin \omega}{\sqrt{1 - e^2}} \right) \tag{3.6}$$

In these formulas, ΔF represents the transit depth, P is the orbital period, e is the orbital eccentricity, and ω is the argument of periastron. The impact parameter influences the transit's characteristics. A smaller impact parameter b indicates a transit that is deeper and longer. Instead, a larger impact parameter results in a transit that is shallower and smoother. These equations are valid under the assumptions that the planet radius R_p is much smaller than the stellar radius $R_\star$, and $R_\star$ is much smaller than the semi-major axis a. Additionally, the condition $t_T \ll \tau$ is generally considered for small planets or non-grazing transits, where the ingress and egress duration are short compared to the total transit duration.

The geometry of the light curve, observed during a planetary transit, depends on the flat time t_F, the time between the second and third contact. Assuming a trapezoidal geometry, we derive the following equation:

$$\left(\frac{t_F}{t_T}\right)^2 = \frac{(R_\star - R_p)^2 - a^2 \cos^2 i}{(R_\star + R_p)^2 - a^2 \cos^2 i} \tag{3.7}$$

Where t_F is the duration of the flat part of the light curve, t_T is the total duration of the transit, $R_\star$ and R_p are the radii of the star and planet respectively, a is the semi-major axis, and i is the inclination of the planet's orbit. In particular, when $i = 90°$, this simplifies to:

$$\frac{t_F}{t_T} = \frac{R_\star - R_p}{R_\star + R_p} \tag{3.8}$$

However, transit light curves are not perfectly flat and trapezoidal because they depend on the star's properties, the wavelength of observations, and the planet's atmospheric absorption. First, stars are not uniformly illuminated. Variations in surface brightness are caused by factors such as dark stellar spots, bright faculae, plages, flares, and stellar atmosphere. These factors affect the observed light curve.

In particular, the stellar atmosphere is related to the so-called limb darkening, the decrease in apparent brightness from the centre to the edge of a star. This phenomenon in stars occurs because the optical depth becomes unity at colder, less bright layers towards the limb. The planet blocks less light at the limb, resulting in a less pronounced transit drop. Giant stars show less pronounced limb darkening, and in some cases, may even show limb brightening. This difference arises from the structure of their atmospheres: giant stars have extended, less dense atmospheres where the temperature gradient is not as vertical as in dwarf stars. As a result, the decrease in temperature towards the limb is less significant, leading to a more uniform brightness across the stellar disk. Limb darkening is typically described using functions of $\mu = \cos\theta$, where θ is the angle between the stellar surface normal and the observer's line of sight. The radial dependence of intensity, $I(r)$, is often approximated by the following fourth-order Taylor series expansion:

$$I(r) = 1 + \sum_{n=1}^{4} c_n(1 - \mu^{n/2}), \tag{3.9}$$

where $I(0) = 1$. The coefficients c_n depend on the star's effective temperature, luminosity class and metallicity.

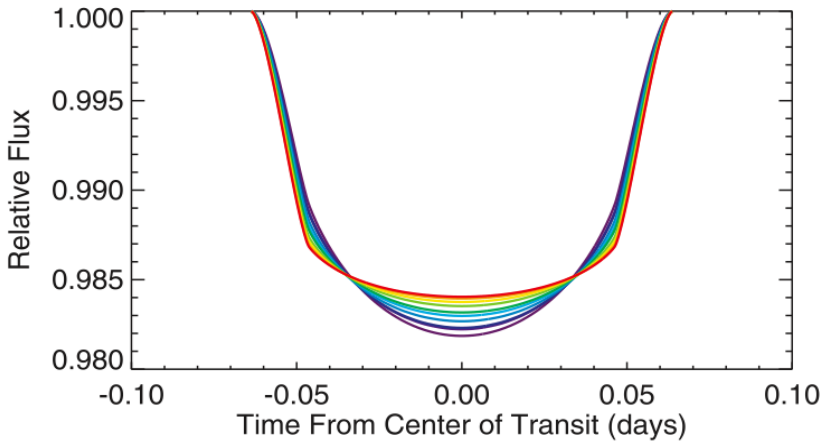

Figure 3.2: Limb-darkened light curves of the exoplanet HD 209458b, across ten bandpasses ranging from 293 to 1019 nm, plotted using a model with four-parameter nonlinear limb-darkening coefficients. Credits: [52]

The wavelength of observations can affect the limb-darkened transit light curves. Observations at shorter wavelengths tend to smooth out the trapezoidal shape of the transit light curve, making the geometric features less precise, as visible in the Figure (3.2). For this reason, it is better to observe at longer wavelengths, in the near-infrared or infrared regions of the spectrum. These longer wavelengths permit us to obtain better the trapezoidal form. One of the best instruments for transit observations is the James Webb Space Telescope (JWST), as it operates in the wavelength range 0.6-28 μm that includes the infrared part of the spectrum [7].

Another factor that distorts the light curve is the absorption of the starlight by the planet's atmosphere. To understand how the atmosphere can change the light curve I need to introduce the concept of transit spectroscopy and radiative transfer.

3.2 Transit spectroscopy

Transit spectroscopy is a technique based on the transit method, useful for getting information on the planet's atmosphere. It permits us to characterize a planet's atmosphere in two ways: through the transmission spectrum during primary transit and the emission spectrum during secondary transit. These two approaches, schematized in the Figure (3.3), allow complementary information about the atmosphere: the transmission spectroscopy provides information on the chemistry, temperature and pressure gradient; the occultation spectroscopy measures the planet's reflected light (albedo) and permits the obtaining of atmosphere thermal information.

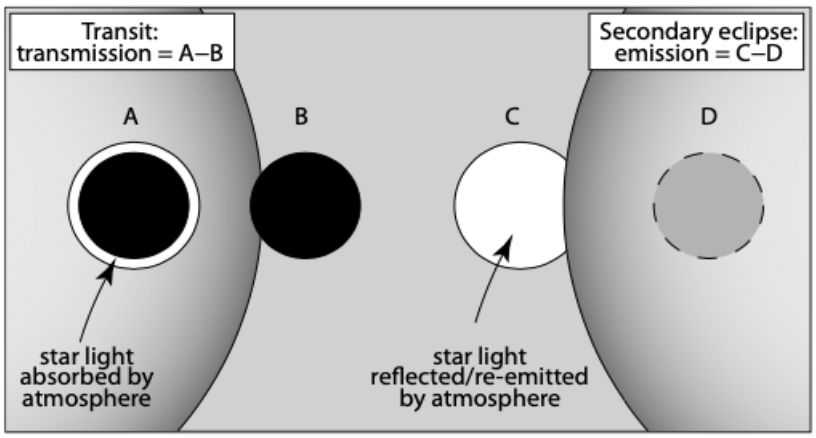

Figure 3.3: Geometry of transmission (left) and emission spectroscopy (right). During the primary transit, a portion of the starlight passes through the planet's atmosphere and it is absorbed. The starlight is reflected from the planet's atmosphere. Credits: [70]

The transmission spectrum is the result of the passage of starlight through the planet's atmosphere at the boundary between its day and night sides. At wavelengths where the atmosphere is more opaque due to absorption by atoms or molecules, the planet blocks more stellar flux. The first detection of an absorbing species was the sodium resonance doublet at 589.3 nm in the hot Jupiter HD 209458b [17]. The absorption of a portion of the starlight causes variations in the transit depth of the planet in the function of the wavelengths, forming the so-called transmission spectrum, which represents the variation that occurs from the transmission of stellar flux through the planet's atmosphere. To measure these variations, it is necessary to bin the light curve in wavelength into spectrophotometric channels and to fit separately the light curve from each channel with a transit model.

The area of the planetary atmosphere analyzed during transit is approximately an annulus with a radial dimension of nH, where H is the atmospheric scale height and n is the number of scale heights crossed at wavelengths with high opacity. The scale height H is the change in altitude over which the pressure drops by a factor of e. Assuming hydrostatic equilibrium and using the ideal gas law it is given by:

$$H = \frac{k_B T_{eq}}{\mu_m g_p} \tag{3.10}$$

Where k_B is Boltzmann's constant, T_{eq} is the equilibrium temperature, μ_m is the mean molecular weight of the atmosphere, and g_p is the planet's surface gravity.

The fractional contribution of the transmission signal, δ, is the ratio of the area of the annulus to the area of the star, calculated as:

$$\delta = \frac{(R_p + nH)^2}{R_\star^2} - \frac{R_p^2}{R_\star^2} \sim \frac{2nR_pH}{R_\star^2} \tag{3.11}$$

Where R_p is the radius of the planet, and $R_\star$ is the radius of the star. This fractional contribution of the transmission signal is used to quantify the star's light that passes through the planet's atmosphere. From the equation (3.11) ideal candidates have high equilibrium temperatures, small host stars, low surface gravity, and low mean molecular mass as hydrogen dominated. These typical candidates are the hot Jupiters and their amplitude of spectral features is just <0.1%, instead, for Earth-like planets, the amplitude can be two to three orders of magnitude smaller, depending on the host star [53].

The emission spectrum represents the light emitted by the planet's day-side atmosphere. To derive this spectrum, it is necessary to disentangle the star and planet spectra. Before the planet moves behind the star, we see a combined spectrum of both the star and the planet. The star spectrum is measured during the eclipse of the planet. As a consequence, the planet's emission spectrum is isolated subtracting the star's spectrum from the combined one. In this way, we observe a brightness that can be attributed to the planet's thermal emission and the reflected light of the star. It is possible to make some considerations on the reflected and emitted wavelengths: the reflected light is prominent at optical wavelengths, while emitted light, which comes from the planet's heat, dominates in the infrared. This is due to the planet's lower temperatures concerning the star. The contrast between the planet's and the star's brightness can improve as the wavelength increases and this makes infrared observations the most suitable ones for the atmosphere [58].
From the emission spectrum is possible to derive the Bond albedo of the planet, crucial to the equilibrium temperature of the planet. This temperature results from a balance between the incident radiation from the host star and that absorbed by the planet or its atmosphere. Ignoring additional heat sources, such as tidal deformation, radiogenic decay, and the greenhouse effect, and considering only the incident starlight, we can use the Stefan–Boltzmann law to express the equilibrium temperature as follows:

$$T_{eq} = T_* \left(\frac{R_*}{2a}\right)^{1/2} [f(1 - A_B)]^{1/4} \tag{3.12}$$

In this equation, R_*, and T_* represent the radius, and temperature of the host star, respectively. The Bond albedo, A_B, indicates the fraction of the incident radiation that is reflected by the planet. The star-planet separation is denoted by a. The heat redistribution factor, f, varies based on the efficiency of heat distribution across the planet's surface: $f = 1$ if the heat is evenly distributed across the entire surface of the planet, which occurs in the case of fast rotation or efficient atmospheric circulation; $f = 2$ if the heat is only emitted by the diurnal hemisphere, which is typical for planets in synchronous rotation [70]. From the equilibrium temperature it is possible to derive the typical size of the emission signal, given by:

$$\frac{F_p}{F_\star} = \frac{B_\lambda(T_{eq})}{B_\lambda(T_\star)} \left(\frac{R_p}{R_\star}\right)^2 \tag{3.13}$$

where $F_p/F_\star$ is the planet-to-star flux ratio, $B_\lambda(T_\star)$ is the blackbody radiation of the star at temperature $T_\star$, $B_\lambda(T_{eq})$ is the blackbody radiation of the planet at temperature T_{eq}, and $R_p/R_\star$ is the planet-to-star radius ratio. In general, the planet is cooler than the star, so the flux ratio is larger at longer wavelengths [53].

3.3 Atmospheric physics

Atmospheric characterization is crucial to constrain the formation and evolution models and shed light on the physical phenomenon in exoplanetary atmospheres. In the following paragraphs, I have reported the theoretical information of atmospheric physics, following the books by Liou (2002) [55] and by Andrews (2010) [3].

3.3.1 Atmospheric composition

The atmosphere can be defined as a mixture of atoms and molecules that can be approximate to ideal gases. Following the ideal gas law, the partial pressure P_i and the partial volume V_i of a gas S_i are given by:

$$P_i = n_i \frac{k_B T}{V} \qquad\qquad V_i = n_i \frac{k_B T}{P} \tag{3.14}$$

where n_i is the number of gas molecules, k_B is Boltzmann's constant, V and P are the volume and the pressure occupied by the gas molecules with a temperature T. If there are n_i molecules of gas S_i, it is possible to define the mass mixing ratio μ_i as:

$$\mu_i = \frac{m_i n_i}{m} \tag{3.15}$$

where m_i multiplied by n_i represents the mass of a specific gas S_i and m is the total mass of the gases. From the equations (3.14) and (3.15), the mean molecular mass μ is defined:

$$\mu = \frac{m}{n} = \sum m_i \frac{n_i}{n} \tag{3.16}$$

where $n = \sum n_i$ is the total number of molecules and $m = \sum m_i n_i$ is the total mass of the gasses. The mean molecular mass is one of the most important parameters to characterize the composition of a planetary atmosphere because it permits to distinguish between primary or secondary atmosphere.

3.3.2 Chemical equilibrium in atmosphere

The molecular abundances in the atmospheres can be extracted through a heuristic parameterization, directly from the data, making explicit assumptions about the chemistry of the atmosphere. An example is the constant profile, where the molecular mixing ratios are uniform for all the atmospheric layers. Another possibility is the self-consistent parameterization that depends on an underlying chemical model to produce molecular abundances. This model sets the initial abundance of elements in the atmosphere and iterates until convergence. One possible self-consistent chemical model considers the chemical equilibrium approximation, which assumes chemical reaction timescales are faster than the dynamical timescale and neglects photochemical and cosmic-ray processes [1]. This approximation can describe the composition of an atmosphere if it is sufficiently hot and dense to guarantee that chemical reactions occur faster than any other process at work. This is the case of hot exoplanets with atmospheric temperatures higher than 2000 K. In the case of cooler planets, the chemical equilibrium is a good starting point to understand the main atmospheric constituents [59].

To compute the chemical composition of an atmosphere in equilibrium it is necessary to minimize the Gibbs free energy of the system given an initial atomic abundance.

The Gibbs free energy has a nonlinear dependence on the number density of species in the system, which are nonnegative. The number densities n_i for the gas molecule S_i, with $i \in S$ set of all gas molecules, composed by the chemical elements E_j, with $j \in E$ set of all elements, are determined in dissociative equilibrium by the law of mass action and the change conservation of the element j. The law of mass action is given by:

$$n_i = K_i \prod_{j \in E} n_j^{\nu_{ij}} \tag{3.17}$$

where K_i is the mass action constant and ν_{ij} are the coefficients of the stoichiometric matrix, that represents the conservation of mass of the gas molecules.

The mass action constant of gas molecule S_i, in natural logarithm and dimensionless form, is given by:

$$\ln K_i(T) = -\frac{\Delta_d G_i(T)}{RT} \tag{3.18}$$

where T is a given temperature, R is the universal gas constant and $\Delta_d G_i(T)$ is the Gibbs free energy of dissociation, defined as:

$$\Delta_d G_i(T) = \Delta_f G_i(T) - \sum_{j \in E} \nu_{ij} \Delta_f G_j(T) \tag{3.19}$$

where $\Delta_f G_i(T)$ are the Gibbs free energies of formation.
The charge conservation is given by:

$$\epsilon_j n_{\langle H \rangle} = n_j + \sum_{i \in S} \nu_{ij} n_i \tag{3.20}$$

where ϵ_j is the relative elemental abundance concerning hydrogen and $n_{\langle H \rangle}$ indicates the sum of all hydrogen nuclei per unit volume. The stoichiometric coefficients ν_{ij} are nonnegative integers if $j \neq 0$. Instead, if $j = 0$, ν_{i0} is a negative integer number for positively charged species, and ν_{i0} is a positive integer number for negatively charged species. Otherwise, ν_{i0} is null [80].

The equations (3.17) and (3.20) form a system of coupled nonlinear algebraic equations, which must be solved to find the chemical composition of an atmosphere in equilibrium. Various algorithms have been designed to solve the system of equations. One is the Newton-Raphson method [23] in higher dimensions that solves the equations (3.17) and (3.20) simultaneously. Another possibility is to decompose the equation system into a set of equations, each of them in one variable.

3.3.3 Radiative transfer equation

To explain how the radiation emitted by a star changes when it passes through the planet's atmosphere, it is necessary to introduce the radiative transfer process. This phenomenon describes the interaction between the radiation and a material that in our case is represented by the planet's atmosphere. The following equation drives the process:

$$dI_\lambda = -k_\lambda \rho I_\lambda ds + j_\lambda \rho ds \tag{3.21}$$

Where dI_λ is the intensity of radiation I_λ after the passage through the atmosphere of a thickness ds, k_λ represents the mass extinction cross-section (in units of area per mass) for radiation of wavelength λ, ρ is the density of the material and j_λ is the source function coefficient. For simplicity, it is defined the source function $J_\lambda = j_\lambda/k_\lambda$, so the general radiative transfer equation (3.21) becomes:

$$\frac{dI_\lambda}{k_\lambda \rho ds} = -I_\lambda + J_\lambda \tag{3.22}$$

This formula defines that the propagation of radiation through a medium is affected by scattering, absorption and emission processes. Now, it is necessary to explain in detail all the quantities that play a role in the radiative transfer equation.

3.3.4 Scattering, absorption and emission

When the radiation passes through the atmosphere, it is attenuated by scattering and absorption processes. This attenuation is called extinction. The mass extinction cross-section k_λ represents the amount of energy removed from the original radiation by the particles, due to scattering and absorption, and it is the sum of the mass scattering cross-section σ_s and the mass absorption cross-section σ_a:

$$k_\lambda = \sigma_s + \sigma_a \tag{3.23}$$

Scattering

Scattering is a physical phenomenon that forces the deviation of the radiation from the initial trajectory by localized particles in the medium through which it passes. This process depends on the particle size, which is described by the size parameter:

$$x = \frac{2\pi r}{\lambda} \tag{3.24}$$

It is defined as the ratio of the particle circumference $2\pi r$, considering spherical particles, to the incident wavelength λ. When $x << 1$ there is the Rayleigh scattering, and when $x \geq 1$ there is the Lorenz–Mie scattering. In the planet's atmosphere, the typical sizes of the particles range from 10^{-4} μm for gas molecules to 1 μm for aerosols and 100 μm for ice crystals. In general, the assumption of independent scattering is possible, when the radiation scattered by molecules and particles has the same wavelength as the incident light because they scatter as if all other particles did not exist. In case this assumption is not possible, a particle can scatter the light that has already been scattered by other particles, producing multiple scattering. The consideration of multiple scattering is included in the radiative transfer equation in the source function coefficient j_λ.

The scattering cross-section (in units of area) represents the amount of incident energy that is removed from the original radiation due to the scatter of a molecule that redistributes isotropically the radiation in a spherical area of radius r whose center is the scatterer. The scattering cross section per single molecule is defined as:

$$\sigma_s = \frac{f}{F_0} = \frac{\alpha^2 128\pi^5}{3\lambda^4} \tag{3.25}$$

Where f is the scattered flux, F_0 is the incident flux density, α is the polarizability and λ is the wavelength. Polarizability is a characteristic of matter that acquires an electric dipole moment when subjected to an electric field, as in the case of radiation:

$$\alpha = \frac{3}{4\pi N_s}\left(\frac{m^2-1}{m^2+2}\right) \tag{3.26}$$

Where N_s is the number of molecules per unit volume and m is the refractive index of molecules, which is a parameter associated with the velocity change of the radiation in a medium to the vacuum. The general formula for scatter intensity, valid for the Rayleigh scattering and the Lorenz–Mie scattering, is given by:

$$I(\theta) = I_0 \frac{\sigma_s}{r^2}\frac{P(\theta)}{4\pi} \tag{3.27}$$

Where I_0 is the incident radiation, σ_s is the scattering cross section, r is the radius of the sphere in which the radiation is redistributed after the scatter (also defined as the distance between the scatter and the observer), and $P(\theta)$ is the phase function which represents the angular distribution of scattered radiation. Scattering is related to absorption, which also is responsible for the attenuation of the radiation.

Absorption

Absorption is defined as the process in which a photon is absorbed by an electron in the ground state of an atom. The absorption of a photon allows the electron to jump into a more energetic orbital, i.e. an excited state. Then, when the electron decays to a lower energy state, it emits a photon. Each jump from a lower to an upper energy level is defined as an electronic transition and it is characterized by the absorption or emission of a specific wavelength, related to a precise spectral line.

For the molecules, it is possible to add two other types of transition: vibrational and rotational, depending on how the molecules can store the energy. The vibrational energy is due to the vibration of atoms in molecules concerning their equilibrium position. The rotational energy is related to the rotation of an atom about an axis through its centre of gravity. The three main transitions explained above occur in different parts of the spectrum: the transitions between electronic states are less than 1 μm in the ultra-violet and visible regions, vibrational states are between 1 and 20 μm in the infra-red domain and the rotational states are of the order of $10^2 - 10^4$ μm in the far infrared and microwave regions. In general, vibrational and rotational energies do not modify separately but change during a transition, producing complex vibration–rotation spectral bands in the infra-red [3]. The energy level transitions are possible only at certain discrete wavelengths λ_n, which are related to the absorption coefficient a_λ by:

$$a_\lambda = \sum_n S_n f_n(\lambda - \lambda_n) \tag{3.28}$$

Where S_n are constants called the line strengths and $f_n(\lambda - \lambda_n)$ are the normalized line-shape functions that consider the broadening of the spectral lines.
It is possible to describe three main broadening processes: the natural broadening due to the Heisenberg uncertainty principle, that no particle has both a fixed velocity and position, the pressure broadening due to collisions between the molecules, and the thermal Doppler effect represented by the difference in thermal velocities of atoms and molecules.

The natural broadening is negligible as compared to the other two processes, that are combined in the upper atmosphere. In the lower atmosphere, collision broadening is dominant because of the pressure effect.

Another phenomenon related to absorption is collision-induced absorption (CIA) which is an inelastic process that concerns symmetric molecules. During inelastic collisions involving these molecules, a transient dipole is created and this causes a collision-induced absorption. CIA is related to symmetric molecules, such as O_2, N_2, H_2, and CH_4 because they have no permanent electric dipole like magnetic dipole and electric quadrupole moments, that can contribute to the absorption [74].

At this point, it is possible to define the absorption cross-section:

$$\sigma_a = \frac{a_\lambda}{N_0} \tag{3.29}$$

Where a_λ is the absorption coefficient and $N_0 = 2.687 \times 10^{19}$ is the Loschmidt's number in the unit of particles cm^{-3}, at the temperature of 273 K and pressure of 1013 mbar.

Emission

The absorption of radiation by molecules leads to emission, associated with blackbody radiation. This radiation is described by the Planck function, which relates the radiated monochromatic intensity to the wavelength and the temperature:

$$I_\lambda = B_\lambda(T) = \frac{2hc^2}{\lambda^5(e^{hc/k\lambda T} - 1)} \tag{3.30}$$

where h is the Planck's constant, k is Boltzmann's constant, c is the velocity of light, T is the absolute temperature and λ is the wavelength considered. To obtain the total radiant intensity of a blackbody it is necessary to integrate the Planck function from 0 to ∞, which corresponds to the entire wavelength domain. In this way, the intensity is:

$$B(T) = \frac{2\pi k^4}{15c^2h^3}T^4 = bT^4 \tag{3.31}$$

A characteristic of a blackbody is its isotropic radiation, so the flux density emitted is described by the Stefan-Boltzmann law, given by:

$$F = \pi B(T) = \sigma T^4 \tag{3.32}$$

where σ is the Stefan–Boltzmann constant.

3.3.5 Radiative transfer for plane-parallel atmosphere

One of the most used models to physically describe the planet's atmosphere is considering a plane parallel configuration, a one-dimensional approximation in which the variation of the radiant intensity and atmospheric parameters, such as temperature and gaseous profiles, is possible only in the vertical direction z, that can be representative of height or pressure. This approximation allows us to perform a first approximation of an exoplanetary atmosphere, it is intuitive and easy to implement with a numerical model. The radiative transfer equation for this structure is given by:

$$\mu \frac{dI_\lambda(\tau, \mu, \phi)}{d\tau} = I_\lambda(\tau, \mu, \phi) - J_\lambda(\tau, \mu, \phi) \tag{3.33}$$

Where $\mu = \cos(\theta)$, θ is the the zenith angle, ϕ is the azimuthal angle and $d\tau = -k_\lambda \rho \, dz$. The τ parameter is the optical depth, that represents the attenuation power of molecules of a specific wavelength of the incident radiation. It is defined as the integral over the distance normal to the planes of stratification of the mass extinction cross-section k_λ multiplied by the number density ρ of the atmosphere:

$$\tau(z) = \int_z^\infty k_\lambda \rho \, dz \tag{3.34}$$

To derive the emergent outward intensity for a planet's atmosphere it is necessary to integrate the equation (3.33), upon multiplying by a factor $e^{-\tau/\mu}$, that represents the monochromatic transmittance.

After the integration, the outward intensity at the top of the atmosphere (TOA, $\tau = 0$) is given by:

$$I_\lambda(\tau = 0, \mu, \phi) = I_\lambda(\tau_s, \mu, \phi) e^{-\tau_s/\mu} + \int_0^1 d\mu \int_0^{\tau_s} J_\lambda(\tau, \mu, \phi) e^{-\tau/\mu} d\tau \tag{3.35}$$

Where τ_s is the optical depth at the planetary surface. The zenith angle θ is in a range between 0 and $\pi/2$ for the emergent intensity, so $0 \leq \mu \leq 1$. The term to the left represents the wavelength-dependent intensity at the top of the atmosphere for which the optical depth becomes null, and the first and second terms to the right are respectively the planetary surface contribution (attenuated to the top) and the atmospheric contribution.

A simplification of the plane parallel configuration is possible considering only an absorbing and emitting medium in local thermodynamic equilibrium, in which the scattering effects are negligible compared to the other processes. Under this assumption, absorptions and emissions are considered symmetrical to the azimuthal angle ϕ and the source function J_λ is assumed as the Planck intensity $B_\lambda(T)$ at a given temperature T. Considering these premises, the thermal radiation of an atmosphere is given by the Schwartzschild equation, which corresponds to the equation (3.33) modified:

$$\mu \, \frac{dI_\lambda(\tau, \mu)}{d\tau} = I_\lambda(\tau, \mu) - B_\lambda(T) \tag{3.36}$$

In this way, the total integrated radiation at the top of the atmosphere becomes:

$$I_\lambda(\tau = 0, \mu) = B_\lambda(T_s) e^{-\tau_s/\mu} + \int_0^1 d\mu \int_0^{\tau_s} B_\lambda(T) e^{-\tau/\mu} d\tau \tag{3.37}$$

Where T_s is the surface temperature of the planet.

3.3.6 Clouds in atmosphere

A possible component that can be found in an exoplanet's atmosphere is the presence of clouds. The term "clouds" indicates condensates that grow from an atmospheric species when the partial pressure of a gas, indicated by the formula (3.14), exceeds its saturation gas pressure. The presence of clouds can affect the atmospheric thermal profile and the spectra of scattered and emitted radiation of a planet. These objects are defined as grey clouds and they are characterized by incomplete absorption and emission of radiation.

The ratio of a grey body's thermal radiation to a black body's thermal radiation at the same temperature is called the emissivity of the grey body. The emissivity of a black body is unity, while that of a grey body is between 0 and 1, and it is constant at all temperatures and throughout the entire range of wavelengths.

Cloud particles scatter incident light, raising the albedo of a planet, so a cloudy planet is brighter in infrared by scattering more light back to space than a cloudless planet. If the particles are assumed to be sufficiently far from each other and the distance between them is much greater than the incident wavelength, the scattering by one particle can be treated independently of other particles and it is called independent scattering. The single-scattering albedo $\tilde{\omega}$ is defined by:

$$\tilde{\omega} = \frac{\beta_s}{\beta_e} \qquad \text{or} \qquad 1 - \tilde{\omega} = \frac{\beta_a}{\beta_e} \tag{3.38}$$

It represents the rate of radiation that experiences the scattering event. β_s is the scattering coefficient for cloud particles, β_a is the absorption coefficient for cloud particles and β_e is the extinction coefficient. They can be defined considering the particle size distribution a and assuming that the size range of particles is from a_1 to a_2:

$$\beta_{s,a,e} = \int_{a_1}^{a_2} \sigma_{s,a,e}(a) n(a)\, da \tag{3.39}$$

Where $n(a)$ is the parameter that describes the size spectrum of the cloud particles, $\sigma_e(a)$ $\sigma_s(a)$ and $\sigma_a(a)$ are respectively the extinction, the scattering and the absorption cross sections. The scattered light contribution to the planet's spectrum from clouds is generally a bright, flat continuum. At lower spectral resolution, in particular for terrestrial exoplanets, clouds cause the molecular features less detectable, strongly affecting the determination of their atmospheric composition [64].

Clouds also decrease the depth to which light can penetrate the atmosphere, reducing the amount of atmospheric gas encountered by the incident photons before being scattered. As there are fewer opportunities for those photons to be absorbed by atoms and molecules, the gaseous absorption features are shallower for a cloudy planet compared to a clear one.

Another important effect due to the clouds regards the increase of the apparent radius of the planet, observable in the transmission spectrum. The impact of clouds on the transmission spectra depends strongly on their atmospheric height because the spectra give information only about the atmospheric features above the cloud layer [27].

A mathematical description of the clouds' effect involves considering a plane-parallel cloud layer and modifying the radiative transfer equation (3.36) adding the source function J_λ associated with clouds scattering. The equation (neglecting the variables in parenthesis) can be rewritten as:

$$\mu \frac{dI_\lambda}{dz} = -\beta_a(I_\lambda - B_\lambda) - \beta_s(I_\lambda - J_\lambda) = -\beta_e(I_\lambda - S_\lambda) \tag{3.40}$$

Where β_a is the absorption coefficient for cloud particles, β_s is the scattering coefficient for cloud particles and $\beta_e = \beta_a + \beta_s$ is the extinction coefficient defined as the inverse of the mean free path of a photon before scattering or absorption. The term S_λ is the source function involving absorption and scattering processes and it is defined as:

$$S_\lambda = \frac{\beta_a B_\lambda + \beta_s J_\lambda}{\beta_e} = (1 - \tilde{\omega}) B_\lambda + \tilde{\omega} J_\lambda \tag{3.41}$$

where $\tilde{\omega}$ is the single-scattering albedo defined by the formula (3.38).

3.3.7 Temperature-pressure profile

Another important information about the atmosphere of a planet is its thermal structure which represents the temperature of each atmospheric layer. The variation of the temperature T with height z is defined lapse rate Γ and it is given by:

$$\Gamma(z) = -\frac{dT}{dz} \tag{3.42}$$

It represents the rate of decrease of temperature with height, so $\Gamma > 0$. If $\Gamma < 0$ the temperature increases with height and the layer in which happens is called inversion layer.

The thermal structure of an atmosphere is defined by the temperature-pressure (T-P) profile which is a function that maps the atmospheric pressure to its corresponding temperature at each atmospheric layer [37]. For an atmosphere at rest, in static equilibrium is valid the hydrostatic balance equation, relates the variation of the pressure dP with the variation of height dz:

$$\frac{dP}{dz} = -\rho g_p = -\frac{P\mu_m}{k_B T} g_p \tag{3.43}$$

Where ρ is the gas density equals to $P\mu_m/(k_B T)$ from the ideal gas law, g_p is the gravitational force of the planet, P is the pressure, μ_m is the mean molecular mass of a gas, k_B is the Bolzmann's constant and T is the temperature. If T is a function of height z, it is possible to integrate the equation (3.43) considering P_0 and $z = 0$ respectively the pressure and the height at the surface of the planet:

$$P = P_0 \exp\left(-\frac{\mu_m g_p}{k_B}\int_0^z \frac{dz}{T(z)}\right) \tag{3.44}$$

It is possible to distinguish different T-P profiles, the simplest one is the isothermal profile, in which the temperature $T(z) = T_0$ is constant along the vertical direction z. In this case, the equation (3.44) becomes:

$$P = P_0 \exp\left(-\frac{\mu_m g_p}{k_B}\frac{z}{T_0}\right) = P_0 e^{-z/H} \tag{3.45}$$

Where H is the pressure scale height, as defined in the equation (3.10).

Another possible profile is the N-point, characterised by a division of the atmosphere in the N layer indicated with l. Each layer is considered between two surface z_l and z_{l-1} with respectively pressure of P_l and P_{l-1}. In hydrostatic equilibrium the pressure $P_{l-1} > P_l$ decreases with height $z_l > z_{l-1}$. Integrating the equation (3.43) we obtain:

$$P_{l-1} = P_l \ \exp\left(-\frac{\mu_m g_p}{k_B \bar{T}}(z_l - z_{l-1})\right) \tag{3.46}$$

$\bar{T}$ is the mean temperature within the layer considered and it is defined as:

$$\bar{T} = \frac{\int_{P_l}^{P_{l-1}} T d(\ln P)}{\int_{P_l}^{P_{l-1}} d(\ln P)} \tag{3.47}$$

Where T is the temperature as a function of pressure P. If T is constant, the equation (3.46) describes an isothermal profile for each atmospheric layer.

3.3.8 Transit 1D atmospheric model

The above discussion on atmospheric physics, and in particular on the plane-parallel configuration, applies to primary transits and secondary eclipses. In the case of transmission spectrum, the transit depth at wavelength λ is modified considering an additional term a_λ, corresponding to the wavelength-dependent atmospheric depth:

$$a_\lambda = 2\int_0^{z_\infty} (R_p + z)(1 - e^{-\tau(z)})dz \tag{3.48}$$

Where R_p is the planet radius, z_∞ is defined as the altitude at the top of the atmosphere and $\tau(z)$ is the wavelength-dependent global optical depth and corresponds to the sum of the optical depth of each molecule:

$$\tau(z) = \sum_{m=1}^{N} \tau_m(z) \tag{3.49}$$

where N is the total number of the absorber m. $\tau_m(z)$ is the optical depth of the absorber m and it is given by:

$$\tau_m(z) = \int_z^{z_\infty} \sigma_{a,m}(z)\chi_m(z)\rho_N(z)dz \tag{3.50}$$

where $\sigma_{a,m}$ is the absorption cross-section, χ_m the column density and ρ_N is the number density. z is the altitude profile at each atmospheric layer l, it starts from $z_0 = 0$ and it is defined considering the equation (3.46) as:

$$z_l = z_{l-1} + \Delta z_l \qquad \text{with} \qquad \Delta z_l = -\frac{k_B \bar{T}}{\mu_m g_p} \ln\left(\frac{P_l}{P_{l-1}}\right) \tag{3.51}$$

Where $\bar{T}$ is the mean temperature within the layer, defined by the formula (3.47) and P_l is the pressure at each layer l. In this way, the wavelength-dependent transit depth becomes:

$$\Delta F_\lambda = \frac{R_p^2 + a_\lambda}{R_\star^2} \tag{3.52}$$

Where $R_\star$ is the host star radius. This equation describes the transmission forward model.

Instead, to derive the emission spectrum it is necessary to integrate the equation (3.37) for μ, to obtain the radiation at the top of the atmosphere $I_\lambda(\tau = 0)$. At this point, the emission spectrum is given by:

$$\frac{F_p}{F_\star} = \frac{I_\lambda(\tau = 0)}{I_\star}\left(\frac{R_p}{R_\star}\right)^2 \tag{3.53}$$

Where F_p and $F_\star$ are the fluxes of the planet and the star, $I_\star$ is the stellar intensity, R_p and $R_\star$ are the radii of the planet and the star [2].

Chapter 4

Exoplanet Atmospheric Retrieval

In this chapter, I present an overview of the tools and methodologies employed in the analysis of exoplanetary atmospheres, with a particular focus on the TauREx framework. It begins by introducing the structure and functionalities of TauREx3 [1], including its forward model and retrieval frameworks, and how these features work together to simulate transmission spectra. It continues considering the theoretical explanation of the Bayesian analysis and the nested sampling, focusing on the MULTINEST algorithm [32].

4.1 TauREx

Upon obtaining the planet's transmission spectrum at low resolution, the next passage is to simulate a possible atmosphere that correctly describes the observed data. To compute different retrievals I use TauREx[1], Tau Retrieval for Exoplanets, a line-by-line radiative transfer fully Bayesian retrieval framework for exoplanetary atmospheres [90].

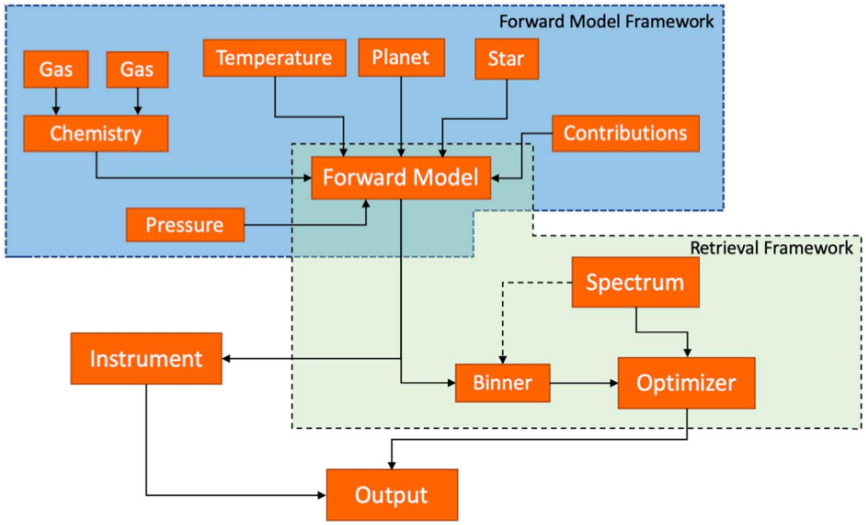

Figure 4.1: The design of TauREx3. The two-framework structure is the Forward model framework in light blue and the Retrieval framework in light green. Each orange box describes a class from Tables (4.1) and (4.2). Solid arrows describe outputs if flowing out, and inputs if flowing in. A dashed arrow describes the creation of an object. Credits: [2]

[1] https://pypi.org/project/taurex/

In this work, I use TauREx3, the last generation of the TauREx exoplanet atmospheric retrieval framework, built with modularity as an all-in-one package for retrievals. TauREx3 consists of two frameworks, visualized in the Figure (4.1): the forward model framework, which builds and computes a forward model, and the retrieval framework, which fits a forward model against an observation. Each framework has different classes, reported in Tables (4.1) and (4.2), all connected by the `ForwardModel` class.

4.1.1 Forward model framework

The forward model framework aims to build a forward model considering different atmospheric components, with dependencies with each other, such as a chemistry profile which requires temperature-pressure (T-P) points, and the contribution function which requires mixing profiles for each species. These atmospheric details are content in different classes, reported in the Table (4.1). In particular, the `ForwardModel` class is responsible for connecting all of these objects and interconnecting them properly to resolve their dependencies. It defines the atmospheric structure through the `model` between transmission and emission, and the contributions to the optical depth at each layer.

Table 4.1: The base classes of the Forward model framework in TauREx3.

Class	Main Purpose
`Chemistry`	Computes chemistry model
`Gas`	Computes single-species mixing ratio for free-type chemistry model
`Global`	Compute the cross sections and opacities
`TemperatureProfile`	Computes temperature profile
`PressureProfile`	Computes pressure profile
`Planet`	Computes planet properties
`Star`	Computes stellar properties and flux
`Contribution`	Performs a calculation on optical depth
`ForwardModel`	Builds and computes a forward model

The `Chemistry` class contains free chemistry models that can define different vertical mixing profiles for each molecule. There are three models: a constant mixing profile along the atmosphere, a two-layer profile, and a profile read from a file. This class can also describe equilibrium chemistry using plugins, for example, the taurex-fastchem plugins and others. The taurex-fastchem plugin, in particular, uses the functionality of the FastChem package[2] [81], a semi-analytical code, that implements a method to minimize the Gibbs energy. The method is based on the solution of a system of coupled non-linear algebraic equations in many variables, that can be decomposed into a set of coupled nonlinear equations in one variable each, which are solved analytically. The theoretical explanation of the equilibrium chemistry in the atmosphere is reported in Section (3.3.2). The possible options of the `Chemistry` class are `fill_gases`, which defines the type of atmosphere such as H_2-He atmosphere, and `ratio` which represents the volume mixing ratio of the gases filled in the atmosphere.

[2] https://pypi.org/project/taurex-fastchem/

In equilibrium chemistry it is possible to add the `metallicity` of the planet, considering a list of `selected_elements` and the presence of ions. To better constrain the metallicity it is possible to consider the ratio of some elements to the oxygen O, as the C/O ratio. For all chemistry models, each molecule is considered either active or inactive. To establish the state of a molecule it is necessary to consider the cross sections: an active molecule is defined by the inclusion of its cross section and has a direct influence on the final spectrum and molecular weight of the atmosphere; otherwise, it is defined as inactive and it only affects the molecular weight.

The `Gas` class describes the molecules included in the free chemistry models. It is possible to choose the type of gas using the `gas_type` option, such as constant gas, and to add the value of the mixing ratio of each species using the `mix_ratio` option.

The `Global` class contains information about the absorption cross-sections and the collision-induced absorption (CIA) of molecules, obtained from databases such as HITRAN (high-resolution transmission molecular absorption database) [38] [48]. It is possible to load absorption cross-sections and CIA using a path that can be set through the `xsec_path` option for cross-sections and through the `cia_path` for collision-induced absorption.

The `TemperatureProfile` class is based on the layer-by-layer approach that permits the modelling of the temperature of each plane-parallel atmospheric layer independently [89]. The temperature profiles included in this class are a basic isothermal profile, a Guillot profile, a profile loaded from a file, and a multipoint temperature profile, chosen using the option `profile_type`. The first and the last profiles are described in Section (3.3.7).

The `PressureProfile` class computes an equally spaced logarithmic grid of $N_l + 1$ pressures at each layer boundary, for a given maximum pressure P_{max}, minimum pressure P_{min}, and number of layers N_l. The central geometric pressure P_l for each layer l is given by:

$$P_l = p_l \sqrt{\frac{p_{l+1}}{p_l}} \tag{4.1}$$

where p_l is the pressure at each layer boundary. The number of layers in which the atmosphere and the pressure profile can be divided is inserted using the `nlayers` option. Typically, the number of layers is set to 100 to consider all the variations in temperature and pressure and to obtain a more detailed T-P profile.

The `Planet` class contains the properties of the planet considered, such as the `planet_radius` and the `planet_mass`, and the `Star` class contains the properties of the star around which the planet orbits, such as `mass`, `radius` and `temperature`. The star's intensity can be approximated using a black body at the stellar temperature through the `star_type=blackbody` or interpolated using the Phoenix library through the `star_type=phoenix`. This library allows obtaining high-resolution synthetic spectra of the star, starting from the stellar atmosphere [46].

In the `Contribution` class are included all the contribution functions C_i necessary to generalize the optical depth τ, that depends on the altitude z and on the wavelength λ:

$$\tau(\lambda, z) = \sum_i C_i(\lambda, z) \tag{4.2}$$

The most basic contribution function is pure `Absorption`, but it is possible to add `Mie` scattering, `Rayleigh` scattering and the collision-induced absorption `CIA`. The theoretical description is reported in Section (3.3.4).

Another possible contribution function regards the clouds, whose presence is described by the `SimpleClouds` option. This contribution considers grey clouds, theoretically described in Section (3.3.6).

4.1.2 Retrieval framework

The retrieval framework aims to fit a `ForwardModel` against an observation (a `Spectrum`). It supports any arbitrary forward model with the only requirement being that the forward model output and the observation fit against match in their overall shape.

Table 4.2: The base classes of the Retrieval framework in TauREx3.

Class	Main Purpose
`Spectrum`	Provides some form of spectral data to fit against
`Optimizer`	Performs retrievals
`Binner`	Bins spectra to given grid
`Output`	Handles file writes
`Instrument`	Bins and assigns noise

The `Spectrum` is a base class where spectra are loaded (or later created). This is used to either plot against the forward model or passed into the optimizer to be used to fit the forward model. A possible option is a `observed_spectrum` to insert the path of the spectrum loaded.

The `Binner` class bins both flux and uncertainties to a given grid with the corresponding widths defined at class creation. The implementation considers the relation between the spectral grid, defined by the central bin λ_i and bin widths $\Delta\lambda_i$, and the resampling target grid, defined by λ_j and its corresponding bin widths $\Delta\lambda_j$.

The `Instrument` class consists of a model simulator that passes the result of the forward model into an instrument noise model and generates a new binned spectrum with instrument noise and systematics. This class is used to bypass loading in an observation and to perform a retrieval directly on the simulated spectrum instead.

The `Optimizer` class performs retrievals starting from the fitting parameters of the forward model. It updates the model using the fitting parameters from the sampler, and computing the likelihood. By default, TauREx uses uniform priors for all free parameters, but all prior bounds can be manually specified considering the `Fitting` section. For the retrievals, TauREx3 can use different nested sampling methods such as Nestle [4], MultiNest, PolyChord [41], and dyPolyChord [44]. The Bayesian statistics, the nested sampling and the MultiNest algorithm, used in the analysis of K2-18b's atmosphere, are described in Section (4.2). Some possible options to insert in the `Optimizer` class are `num_live_points`, which corresponds to the number of live points used in nested sampling and `search_multi_modes`, used to search for multiple modes of the posterior distribution.

The `Output` module generates the best-fit model, plots of all marginalized and conditional posteriors, and releases the statistics on individual parameters and models. Fits can also be defined in the input file under the `Fitting` section. The input file is dynamic and can set user-defined fitting parameters as well.

4.2 Bayesian Analysis and nested sampling

The Bayesian method sets a probability to a hypothesis. A prior probability distribution is used to encode prior information about the hypothesis, and this prior is edited using observations to obtain a posterior probability distribution. The Bayesian statistics can be divided into two analyses: the Bayesian parameter estimation aims to infer the posterior distribution for model parameters given a model, prior information, and observations; the model comparison aims to find which of the possible models is best suited to describe the observations.

4.2.1 Bayesian parameter estimation

The Bayesian parameter estimation is based on the probability distribution for the model parameters, called posterior distribution. It is obtained by prior distribution with a sampling distribution, also known as the likelihood. The role of a prior distribution is to encapsulate the information and assumptions about a model parameter, following, for example, previous research. The likelihood represents the probability that the observations follow from a given model evaluated at a given point in the model parameter space. In parameter estimation, the model consists of a deterministic part that models the signals using a parametric model and a stochastic part that is used to explain the uncertainties [69]. The posterior distribution encodes the information about the model parameters and it is based on Bayes' theorem:

$$P(\boldsymbol{\theta}|\mathbf{x}, M) = \frac{P(\mathbf{x}|\boldsymbol{\theta}, M)P(\boldsymbol{\theta}, M)}{P(\mathbf{x}|M)} \tag{4.3}$$

where $P(\boldsymbol{\theta}|\mathbf{x}, M)$ is the posterior probability of the model parameters $\boldsymbol{\theta}$ given the data $\mathbf{x}$, assuming the forward model M. $P(\boldsymbol{\theta}, M)$ is the Bayesian prior that depends on the model M and its parameters $\boldsymbol{\theta}$. $P(\mathbf{x}|\boldsymbol{\theta}, M)$ indicates the likelihood. The likelihood of a single observation is given by the Gaussian formula:

$$P(x_i|\boldsymbol{\theta}, M_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i - M_i}{\sigma_i}\right)^2\right] \tag{4.4}$$

where σ_i is the error on the observed spectral point. The log-likelihood of a dataset consisting of N observations is given by:

$$\ln P(x_i|\boldsymbol{\theta}, M_i) = -\frac{1}{2}\left(N \ln 2\pi + \sum_{i=1}^{N} \ln \sigma_i^2 + \sum_{i=1}^{N} \frac{(x_i - M_i)^2}{2\sigma_i^2}\right) \tag{4.5}$$

$P(\mathbf{x}|M)$ is the Bayesian evidence, also called marginal likelihood, that permits testing of the adequacy of the model itself and performing the model selection. The marginal likelihood can be ignored in the parameter estimation and for this reason, it is possible to work with an unnormalised joint posterior density, defined as:

$$P(\boldsymbol{\theta}|\mathbf{x}, M) = P(\mathbf{x}|\boldsymbol{\theta}, M)P(\boldsymbol{\theta}, M) \tag{4.6}$$

The posterior distribution for a single parameter θ_i is called marginalisation and it is obtained by integrating the equation (4.6) over all other parameters:

$$P(\theta_i|\mathbf{x}, M) = \int P(\boldsymbol{\theta}|\mathbf{x}, M) d\boldsymbol{\theta} \tag{4.7}$$

The result is a marginal posterior distribution for the parameter [69].

4.2.2 Bayesian model comparison

The Bayesian model comparison is based on the Bayesian evidence $E = P(\mathbf{x}|M)$, which is given by the integral required to normalise equation (4.3):

$$E = P(\mathbf{x}|M) = \int P(\mathbf{x}|\boldsymbol{\theta}, M)P(\boldsymbol{\theta}, M)d\boldsymbol{\theta} \tag{4.8}$$

where $P(\mathbf{x}|\boldsymbol{\theta}, M)$ indicates the likelihood and $P(\boldsymbol{\theta}, M)$ is the Bayesian prior.
The model selection between two models M_1 and M_2 can be determined by comparing their respective posterior probabilities, as follows:

$$\frac{P(\boldsymbol{\theta}|\mathbf{x}, M_1)}{P(\boldsymbol{\theta}|\mathbf{x}, M_2)} = \frac{P(\mathbf{x}|M_1)}{P(\mathbf{x}|M_2)}\frac{P(\boldsymbol{\theta}, M_1)}{P(\boldsymbol{\theta}, M_2)} \tag{4.9}$$

When two models are equally probable a priori, the ratio $P(\boldsymbol{\theta}, M_1)/P(\boldsymbol{\theta}, M_2)$ is unity and the parameter used to establish the best model is the Bayes Factor BF, defined as the ratio of the Bayesian evidence of the two models:

$$BF = \frac{E_1}{E_2} = \frac{P(\mathbf{x}|M_1)}{P(\mathbf{x}|M_2)} \tag{4.10}$$

where $E_1 = P(\mathbf{x}|M_1)$ and $E_2 = P(\mathbf{x}|M_2)$ are the evidences defined by the equation (4.8) of the first and the second model, respectively [90].

The Bayes Factor is a consistent selector of the ratio (4.10) increasingly prefers the correct model in the limit of large data and prefers simpler models if the fits are similar, acts as Occam's razors [49]. To quantify the detection significance of an atmosphere, it is possible to consider the Atmospheric Detectability Index (ADI) [87]. The ADI is the positively defined Bayes Factor between the nominal atmospheric model (M_N) with active trace molecules and a flat-line model (M_F) defined as a pure cloudy model. The following calculation gives it:

$$ADI = \begin{cases} \log(E_N) - \log(E_F), & \text{if } \log(E_N) > \log(E_F) \\ 0, & \text{otherwise} \end{cases} \tag{4.11}$$

where E_N is the Bayesian evidence of the nominal model and E_F is the Bayesian evidence of the flat-line model. The ADI is based on the Jeffreys' scale for the strength of evidence [47] [86], an empirically calibrated scale. In particular, an ADI's value greater than 3 means a significant detection of an atmosphere at 3σ, while a value greater than 11 means a significant detection at 5σ. ADIs below 3 are unable to favour the more complex atmospheric model, M_N over the lower dimensional model M_F because the spectral feature amplitudes are insufficient to define the best one [87].

4.2.3 Nested sampling and MULTINEST

The main problem related to the Bayesian parameter estimation and the Bayesian model comparison is the estimation of the Bayesian evidence, defined by the equation (4.8). The integration is rarely analytically tractable and generally requires specialised numerical methods.

Methods and tools for evidence estimation start from simple techniques considering low-dimensional models, such as direct Monte Carlo integration; to slightly more elaborate ones, such as different importance sampling approaches; and too complicated ones, such as the more advanced nested sampling methods [69].

Nested sampling is a Monte Carlo technique used to evaluate the Bayesian evidence efficiently, and also produces posterior inferences as a by-product. This method manipulates the relation between the likelihood and prior volume to transform the multidimensional evidence integral (4.8) into a one-dimensional integral as follows:

$$E = \int_0^1 \mathcal{L}(X) dX \tag{4.12}$$

where $\mathcal{L}(X)$ is the likelihood in function of the prior volume X, illustrated in the Figure (4.2).

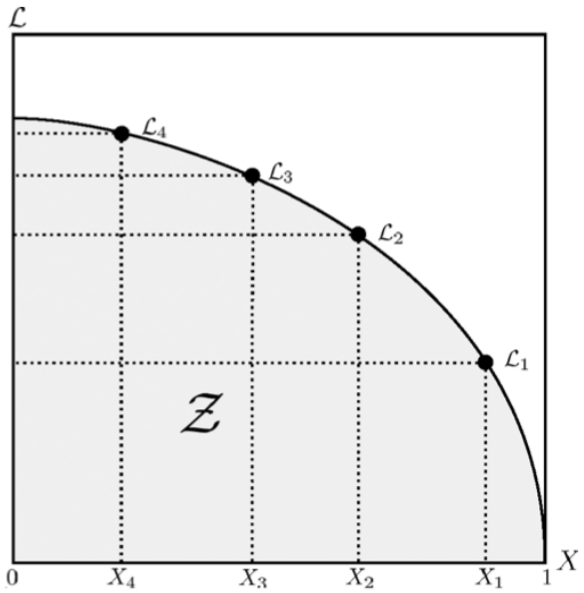

Figure 4.2: Schematic plot illustrating the $\mathcal{L}(X)$ function where the prior volumes X_i are associated with each likelihood $\mathcal{L}(X_i)$. The $\mathcal{Z}$ indicates the evidence E. Credits: [32]

The evidence can be approximated numerically using standard quadrature methods as a weighted sum:

$$E = \sum_{i=1}^{M} \mathcal{L}(X_i) w_i \tag{4.13}$$

where M is the total number of samples, X_i is a sequence of decreasing values and w_i are the weights, defined by the simple trapezium rule:

$$w_i = \frac{1}{2}(X_{i-1} - X_{i+1}) \tag{4.14}$$

The equation (4.13) is obtained considering an iterative process. The iteration counter is first set to $i = 0$ and N live (or active) samples are pulled from the uniform distribution over the prior range, so the initial prior volume is $X_0 = 1$. At each subsequent iteration i, the live samples remove the lowest-likelihood point formula.

The drawing of a replacement with $\mathcal{L}(X) > \mathcal{L}(X_i)$ and the reduction of the corresponding prior volume $X_i = t_i X_{i-1}$ are repeated until the entire prior volume has been crossed. The term t_i follows the formula:

$$P(t) = Nt^{N-1} \tag{4.15}$$

which represents the probability distribution for the largest N samples extracted uniformly from the interval [0, 1]. The process passes through nested shells of likelihood as the prior volume is reduced and becomes equal to $X_i = e^{-i/N}$. The nested sampling algorithm is terminated on determining the evidence at the specified precision of 0.5 in log evidence [32]. The determination of the upper limit on the evidence can be done from the remaining set of current active points. The largest evidence contribution that can be made by the remaining portion of the posterior is:

$$\Delta E_i = \mathcal{L}(X_{max}) X_i \tag{4.16}$$

That is defined as the product of the remaining prior volume X_i and maximum-likelihood value $\mathcal{L}(X_{max})$. The uncertainties on the calculated log-evidence $\ln E$ are:

$$\ln E = \ln \left(\sum_{i=1}^{M} \mathcal{L}(X_i) w_i \right) \pm \sqrt{\frac{H}{N}} \tag{4.17}$$

They can be estimated from a single run of the nested sampling algorithm by calculating the relative entropy H of the full sequence of samples. The relative entropy is given by:

$$H \approx \sum_{i=1}^{M} \frac{\mathcal{L}(X_i) w_i}{E} \ln \left(\frac{\mathcal{L}(X_i)}{E} \right) \tag{4.18}$$

The posterior inferences can be obtained once the evidence E is found. It can be generated using the full sequence of discarded points from the nested sampling method which are the points with the lowest likelihood value at each iteration i of the algorithm. For each point, is assigned the weight p_i defined as:

$$p_i = \frac{\mathcal{L}(X_i) w_i}{E} \tag{4.19}$$

where the sample index i runs from 1 to the total number of sampled points M [31].

The difficult part in implementing the nested sampling algorithm is removing samples from the prior within the constraint $\mathcal{L}(X) > \mathcal{L}(X_i)$ at each iteration i. A possible solution is the Ellipsoidal nested sampling [67], which approximates the iso-likelihood contour $\mathcal{L}(X) > \mathcal{L}(X_i)$ by a D-dimensional ellipsoid determined from the covariance matrix of the set of live points. An algorithm implemented for the simultaneous ellipsoidal nested sampling method is MULTINEST [32]. It permits partitioning into clusters the full set of N active points, at each iteration i of the nested sampling process, to construct the ellipsoidal bounds for each cluster, without degeneracies and considering the expectation-maximization approach. At the ith iteration of the nested sampling process, this approach is used to find the optimal ellipsoidal decomposition of N live points distributed uniformly in the prior volume X_i. The lowest-likelihood point from the set of N active points is removed becoming inactive and replaced by a new point drawn from the set of ellipsoids, considering any overlaps. Once a point becomes inactive, it plays no further part in the nested sampling process.

One of the advantages of the MULTINEST algorithm is the implementation of multi-modal posterior distributions, solving the curving degeneracy problem. To solve the multimodal problem it is required a large number N of active points to guarantee that all the modes are detected. The large number is responsible for a very slow convergence of the algorithm. For this reason, it is necessary to decrease the number of active points as the algorithm proceeds to higher likelihood levels. The number of isolated regions in the iso-likelihood surface is expected to decrease with increasing likelihood. It is possible to reduce the number of active points assuming that the change in the largest evidence contribution ΔE, that can be made by the remaining portion of the posterior, is linear locally. At the ith iteration the number of active points N_i is set to:

$$N_i = N_{i-1} - N_{min}\frac{\Delta E_{i-1} - \Delta E_i}{\Delta E_i - \text{tol}} \tag{4.20}$$

where $N_{min} \leq N_i \leq N_{i-1}$ with N_{min} the minimum number of active points allowed and tol is the precision on the final evidence.

To summarize, it is possible to notice that the MULTINEST algorithm is controlled by the number of active points N. To calculate the evidence accurately, N has to be large enough and must always exceed the dimensionality D of the parameter space so that all the regions of the parameter space are sampled adequately [32].

Chapter 5

Atmospheric Analysis of K2-18b

In this chapter, I describe the observations and data reduction processes employed by Madhusudhan et al. (2023) [63] to obtain the transmission spectrum of the planet K2-18b. I also explain the atmospheric characterization of K2-18b describing the simulations computed first to investigate the presence of a primary atmosphere (light and with volatile elements) and only after assuming the possibility of having a secondary atmosphere (heavy with no volatile elements anymore).

5.1 Observations and data reduction

The atmospheric characterization of the planet K2-18b, reported in this work, was possible using the transmission spectrum obtained by Madhusudhan et al. (2023) [63]. They observed the planet's transit using the James Webb Space Telescope (JWST). This telescope, launched on December 25, 2021 and positioned at the Sun-Earth L2 point, is an international collaboration led by NASA, with significant contributions from the European Space Agency (ESA) and the Canadian Space Agency (CSA). It features a 6.5-meter diameter cold mirror and sophisticated instruments covering wavelengths 0.6-28 μm [36].

The telescope's Integrated Science Instrument Module (ISIM) includes several components, which can be summarized into four main elements: NIRSpec, NIRISS, NIRCam, and MIRI, as illustrated in the Figure (5.1).

- NIRSpec, Near-Infrared Spectrograph, covers wavelengths from 0.6 to 5.3 μm with different resolution spectroscopy, $R \sim 100$, $R \sim 1000$ and $R \sim 3000$. This ability permits it to resolve spectral features of molecules in exoplanet atmospheres, outperforming HST/WFC3 and Spitzer in resolution and wavelength coverage.
- NIRISS, Near Infrared Imager and Slitless Spectrograph has a variety of imaging and spectroscopy modes, covering wavelengths from 0.6 to 2.5 μm with a spectral resolution of $R \sim 300 - 800$.
- NIRCam, Near-Infrared Camera, has different filters capable of covering a wavelength range from 0.7 to 5 μm. The filters can be narrow ($R = 100$), medium ($R = 10$), wide ($R = 4$), and double wide ($R = 2$). NIRCam has grism with resolution spectroscopy of $R \sim 1700$ for the wavelength range 2.4-5 μm and they are immune to slit losses and any signal modulation during transit observations.

- MIRI, Mid-Infrared Instrument, presents four different observing modes, including imaging with 9 photometric bands, coronagraphic imaging with 4 different filters, low-resolution spectroscopy $R \sim 100$ with a slit or slitless configuration, and medium-resolution spectroscopy $R \sim 1550$–3250 [7].

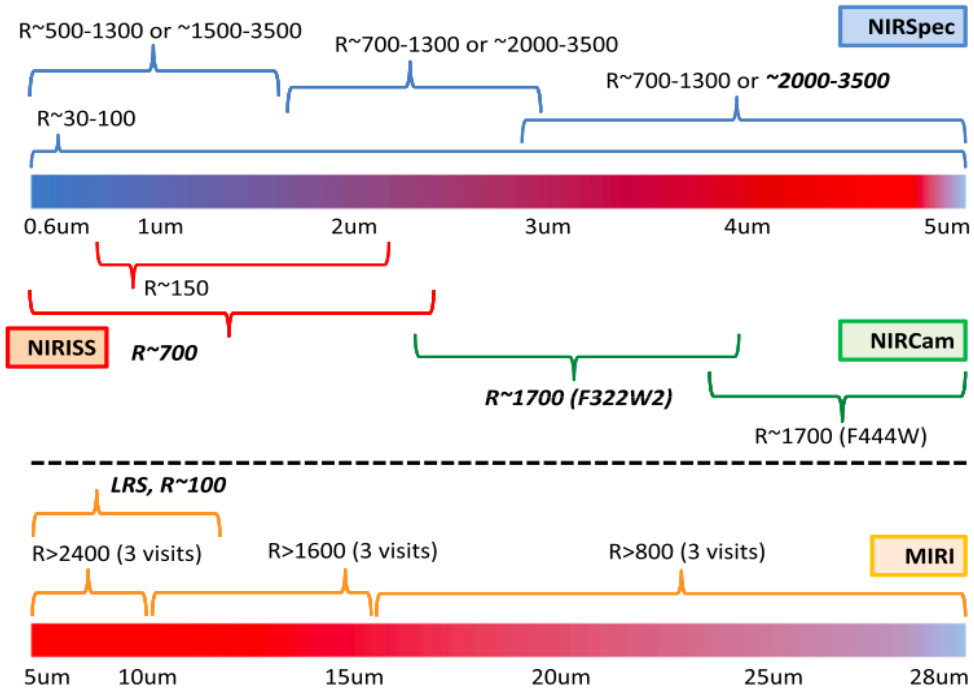

Figure 5.1: A schematic overview of the JWST instrument modes available for transit observations. Credits: [7]

Madhusudhan et al. (2023) [63] observed two primary transits of the planet K2-18b in front of its host star: the first between Jan 20, 2023, 18:37:38 UTC and Jan 21, 2023, 01:11:32 UTC for a total exposure time of 5.3 hours; the second between Jun 1, 2023, 13:49:20 UTC and Jun 1, 2023, 19:36:05 UTC, with an exposure time of 4.9 hours.

The first transit was observed using the NIRSpec G395H grating. This instrument permits the highest-resolution mode with $R \sim 2700$, considering the F290LP filter, the SUB2048 subarray and the NRSRAPID readout pattern. There were used two different detectors, NRS1 and NRS2, with span wavelength ranges of 2.73- 3.72 μm and 3.82-5.17 μm, respectively, with a gap between 3.72-3.82 μm. The data reduction was made in three steps, using the JWST Science Calibration Pipeline [15] for the first two and Madhusudhan et al. (2023) [63] custom-built pipeline for the spectral extraction in the third step. The first stage was characterized by the detector-level corrections with 2D images of the spectroscopic time-series data of the count rate for each integration as outputs. In the second stage, they applied the wavelength calibration for the spectral trace. In the last stage, they extracted the 1D spectrum from the 2D image corrected and calibrated, for each of the two detectors (NRS1 and NRS2) separately.

The second transit was observed using the NIRISS Single Object Slitless Spectroscopy (SOSS) instrument mode for the medium-resolution spectroscopy in a wavelength range 0.85–2.85 μm. The observation considered the GR700XD grism, with a resolution $R \sim 700$, the CLEAR filter, the SUB-STRIP256 subarray and the NISRAPID readout pattern.

Also in this case, the data reduction provides three steps, the first two using the JWST Science Calibration Pipeline [15] and the third using the JExoRES pipeline [45] for the 1D spectral extraction.

After the extraction, the 1D spectral time series from both observations were used for light-curve fitting to derive the transit depths in three steps. In the first stage the white light curves, from both observations, were used to derive the wavelength-independent system parameters at high precision. For NIRISS the white light curve was from the first order and for NIRSpec they used the combined white light curve from both NRS1 and NRS2. In the second stage, the light curves were binned into $R \sim 20$ and the wavelength-dependent limb-darkening coefficients (LDCs) were fitted, using the two-parameter quadratic limb-darkening law. Finally, in the third stage, these LDCs were fixed in the respective $R \sim 20$ bins to fit the transit depths at native resolution to obtain the final transmission spectrum of the planet.

The atmospheric simulations explained in the following paragraphs are made using the planet's transmission spectrum at low resolution, as observed in the Figure (5.2). For this reason, the spectrum of NIRISS is binned to $R \approx 25$ while the spectrum of NIRSpec is binned to $R \approx 55$.

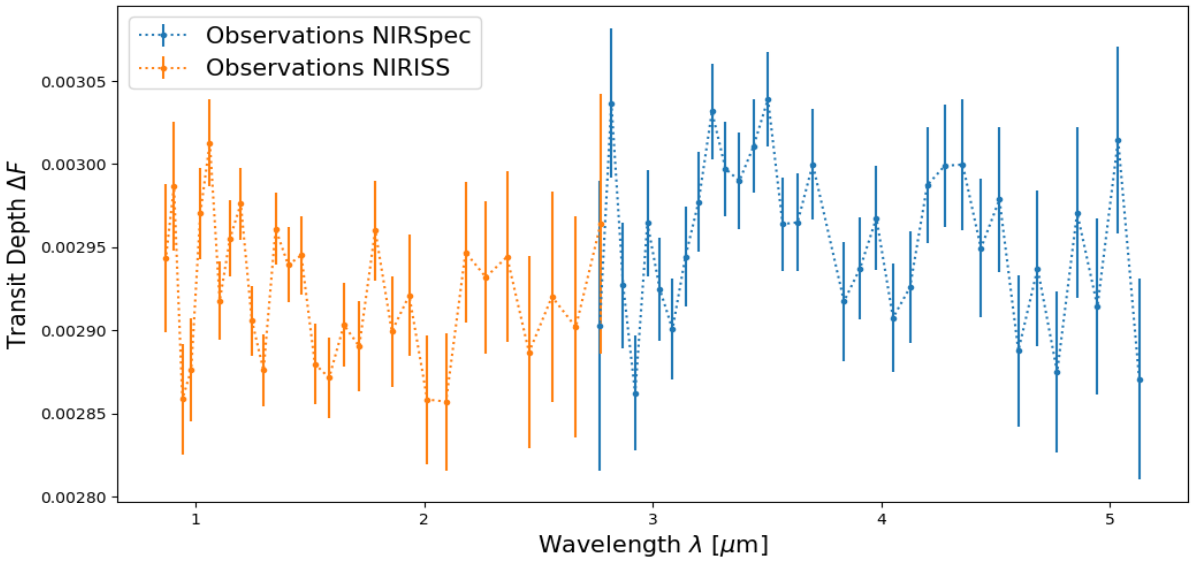

Figure 5.2: The transmission spectrum of K2-18b. The data in orange show the NIRISS spectrum between 0.9-2.8 μm and those in blue show the NIRSpec spectrum between 2.8-5.2 μm. The spectra are binned to $R \approx 25$ and $R \approx 55$ for NIRISS and NIRSpec, respectively.

5.2 Atmospheric retrievals

To compute the atmospheric analysis of the planet K2-18b I utilized the Bayesian atmospheric retrieval framework TauREx3, described in Section (4.1), which maps the correlations between the fitted atmospheric parameters through nested sampling. I considered different retrievals with different parameters and profiles, summarized in the following list and described in detail in the following sections. Initially, I simulated a primary atmosphere considering hydrogen and helium and thereafter I added nitrogen, a heavier element to simulate a secondary atmosphere and verify which type is best suited to describe the spectrum of K2-18b.

1. *CA-iso*: retrieval of an atmosphere full of He and H_2 with constant molecular abundances (*CA*) and an isothermal temperature-pressure profile (*iso*).
2. *CA-iso-clouds*: retrieval of an atmosphere full of He and H_2 with constant molecular abundances, an isothermal temperature-pressure profile and a clouds layer (*clouds*).
3. *CA-4p*: retrieval of an atmosphere full of He and H_2 with constant molecular abundances and a 4-point temperature-pressure profile (*4p*).
4. *CA-4p-N2*: retrieval of an atmosphere full of He, H_2 and N_2 (*N2*) with constant molecular abundances and a 4-point temperature-pressure profile.
5. *CA-iso-N2*: retrieval of an atmosphere full of He, H_2 and N_2 with constant molecular abundances and an isothermal temperature-pressure profile.
6. *EC-4p*: retrieval of an atmosphere full of He and H_2 with equilibrium chemistry (*EC*) and a 4-point temperature-pressure profile.

In all retrievals, the atmosphere of K2-18b was simulated as a plane-parallel atmosphere, considering the transmission spectrum shown in the Figure (5.2), which is based on the data reduction by Madhusudhan et al. (2023) [63]. I included the absorption and Rayleigh scattering contributions from the ExoMol database [82], which provides molecular line lists for atmospheric models; and also the CIA contributions, assuming collision-induced absorption between H_2-He and H_2-H_2, from the HITRAN database [38].

The retrievals use two types of temperature-pressure (T-P) profiles: the N-point profile and the isothermal profile. The N-point profile assumes the atmosphere is divided into 4 layers so that it can be referred to as a 4-point profile. Both profiles cover a pressure range from 10^{-4} Pa to 10^{7} Pa, divided into 100 atmospheric layers that are evenly spaced on a logarithmic scale, and have a temperature (constant in the isothermal and at the top of the atmosphere in the 4-points) obtained by Tsiaras et al. (2019) [88] analysing the transmission spectrum of K2-18b from data of the Hubble Space Telescope (HST). In the following tables are reported the values of temperature and pressure set for the T-P profiles.

Table 5.1: Temperature and pressures set for the isothermal profile, used in the various retrievals.

Isothermal	T [K]	P surface [Pa]	P top [Pa]
	286	10^{7}	10^{-4}

Table 5.2: Temperatures and pressures set for the N-point profile, used in the various retrievals.

N-points	T surface [K]	P surface [Pa]	T top [K]	P top [Pa]
	950	10^{7}	286	10^{-4}
	T point 1 [K]	P point 1 [Pa]	T point 2 [K]	P point 2 [Pa]
	600	10^{3}	400	10^{0}

The parameters of the planet K2-18b and the star K2-18 were set following the ones used by Tsiaras et al. (2019) [88] and are reported in the following tables. In particular, I considered the Phoenix library [46] to synthesize the star's spectrum.

Table 5.3: Planet parameters, used in all retrievals.

Planet	Radius [R_J]	Mass [M_J]
	0.246	0.025

Table 5.4: Star parameters, used in all retrievals.

Star	Radius [$R_\odot$]	Mass [$M_\odot$]	Temperature [K]
	0.411	0.359	3457

In the retrievals with constant molecular abundances, I assumed an atmosphere with hydrogen and helium as inactive molecules, in a He/H_2 ratio of 0.041975 [13] and, in some retrievals, I added the presence of the inactive molecule N_2 in a N_2/H_2 ratio of 0.75. The choice to consider N_2 in the atmosphere is explained by the fact of wanting to simulate a heavier atmosphere to observe whether the planet has a primary atmosphere (with only hydrogen and helium) or a secondary atmosphere (considering heavy elements such as nitrogen). I chose to consider H_2O, NH_3, CO, CO_2, CH_4, as active molecules based on the works of Madhusudhan et al. (2023) [63] and Tsiaras et al. (2019) [88]. In particular, the mixing ratio values, that I utilised, follow the results obtained by Madhusudhan et al. (2023) [63] considering the no offsets case which represents the data without modification.

Table 5.5: Inactive molecules, used in the retrievals with constant molecular abundances.

Inactive molecules	He/H_2	N_2/H_2
	0.041975	0.75

Table 5.6: Mixing ratios, used in the retrievals with constant molecular abundances.

Mixing ratios	X_{H_2O}	X_{NH_3}	X_{CO}	X_{CO_2}	X_{CH_4}
	6.17×10^{-4}	3.47×10^{-5}	1.00×10^{-3}	1.78×10^{-2}	9.12×10^{-3}

In the retrieval with equilibrium chemistry, I assumed an atmosphere with hydrogen and helium in a H/He ratio corresponding to 0.2% of the planet's total mass, which is set at 7.96 $M_\oplus$ [88]. In the study by Naho Fujita et al. (2022) [35], the atmosphere accounts for 1% of the total mass for a H_2/He atmosphere on a planet that is 1 Gyr old. However, since the retrieval requires the H/He ratio, not H_2/He, the percentage is reduced by half. Additionally, because the planet's age is around 2.4 Gyr [39], the percentage is further reduced to 0.2%, which corresponds to a H/He ratio of 0.01592. I also considered the planet's metallicity with a value of 175, based on Blain et al. (2020) [13], assuming the presence of H, He, C, O, N, and e^-. To better constrain the planet's metallicity, I set the C/O ratio to 0.5, which is the same as the solar C/O ratio, following Wogan et al. (2024) [91] and Tsiaras et al. (2019) [88].

Table 5.7: Parameters used in the retrievals with equilibrium chemistry.

Equilibrium chemistry	Metallicity [$(Z/H)_\odot$]	H/He	C/O
	175	0.01592	0.5

The optimizer used to perform all the retrievals is MULTINEST, a nested sampling based on the Bayesian statistics, described in Section (4.2.3). The live points assumed are 1000, to have a number larger enough concerning the quantity of free parameters of the various retrievals. The optimizer also permitted the search of the multi-modes in the posterior distributions.

After setting the parameters for each retrieval it is necessary to impose the free parameters and the corresponding Bayesian priors to compute the simulations of the atmosphere of K2-18b. In the following sections, I report each simulation's fitting parameters and the results of the retrievals compared to the results of Madhusudhan et al. (2023) [63] and Tsiaras et al. (2019) [88], reported in the following table.

Table 5.8: Parameters obtained by Madhusudhan et al. (2023) [63] and Tsiaras et al. (2019) [88], used as a comparison for the retrievals described in the successive sections.

Parameter	Madhusudhan et al. (2023)	Tsiaras et al. (2019)	Description
R_p (R_J)	-	0.2190 ± 0.0007	Planet's radius
T (K)	257^{+127}_{-74}	$286.28^{+21.69}_{-18.12}$	Isothermal temperature
$\log(X_{H_2O})$	< -3.21	-0.50 ± 0.21	Mixing ratio of H_2O
$\log(X_{NH_3})$	< -4.46	-	Mixing ratio of NH_3
$\log(X_{CO})$	< -3.00	-	Mixing ratio of CO
$\log(X_{CO_2})$	$-1.75^{+0.45}_{-1.03}$	-	Mixing ratio of CO_2
$\log(X_{CH_4})$	$-2.04^{+0.61}_{-0.72}$	-	Mixing ratio of CH_4
μ (amu)	28.47	$8.05^{+3.49}_{-2.19}$	Mean molecular weight

5.2.1 Simulation *CA-iso*

The retrieval *CA-iso* refers to an atmosphere with constant molecular abundances and an isothermal temperature-pressure profile.

Table 5.9: Fitting parameters of the model *CA-iso* with the uniform Bayesian priors and the posterior results. The molecular abundances obtained from the posteriors are shown as $\log_{10}$ of volume mixing ratios.

Parameter	Bayesian prior	Posterior $\pm 1\sigma$	Description
R_p (R_J)	$\mathcal{U}(0.123, 0.984)$	0.21	Planet's radius
T (K)	$\mathcal{U}(50, 3500)$	$410.18^{+100.17}_{-93.00}$	Isothermal temperature
X_{H_2O}	$\mathcal{U}(10^{-10}, 1)$	$-6.22^{+2.61}_{-2.31}$	Mixing ratio of H_2O
X_{NH_3}	$\mathcal{U}(10^{-10}, 1)$	$-7.29^{+1.99}_{-1.75}$	Mixing ratio of NH_3
X_{CO}	$\mathcal{U}(10^{-10}, 1)$	$-5.65^{+3.63}_{-2.92}$	Mixing ratio of CO
X_{CO_2}	$\mathcal{U}(10^{-10}, 1)$	$-1.10^{+0.24}_{-0.57}$	Mixing ratio of CO_2
X_{CH_4}	$\mathcal{U}(10^{-10}, 1)$	$-0.97^{+0.50}_{-0.71}$	Mixing ratio of CH_4
He/H_2	$\mathcal{U}(10^{-3}, 10^{3})$	$0.59^{+1.48}_{-2.12}$	Helium-hydrogen ratio
μ (amu)	-	$9.13^{+3.39}_{-2.81}$	Mean molecular weight

To simulate this model I assumed as free parameters the radius of the planet reported in the Table (5.3), the temperature of the isothermal T-P profile in the Table (5.1), the ratio He/H_2 from the Table (5.5) and the mixing ratios of the molecules in the Table (5.6).

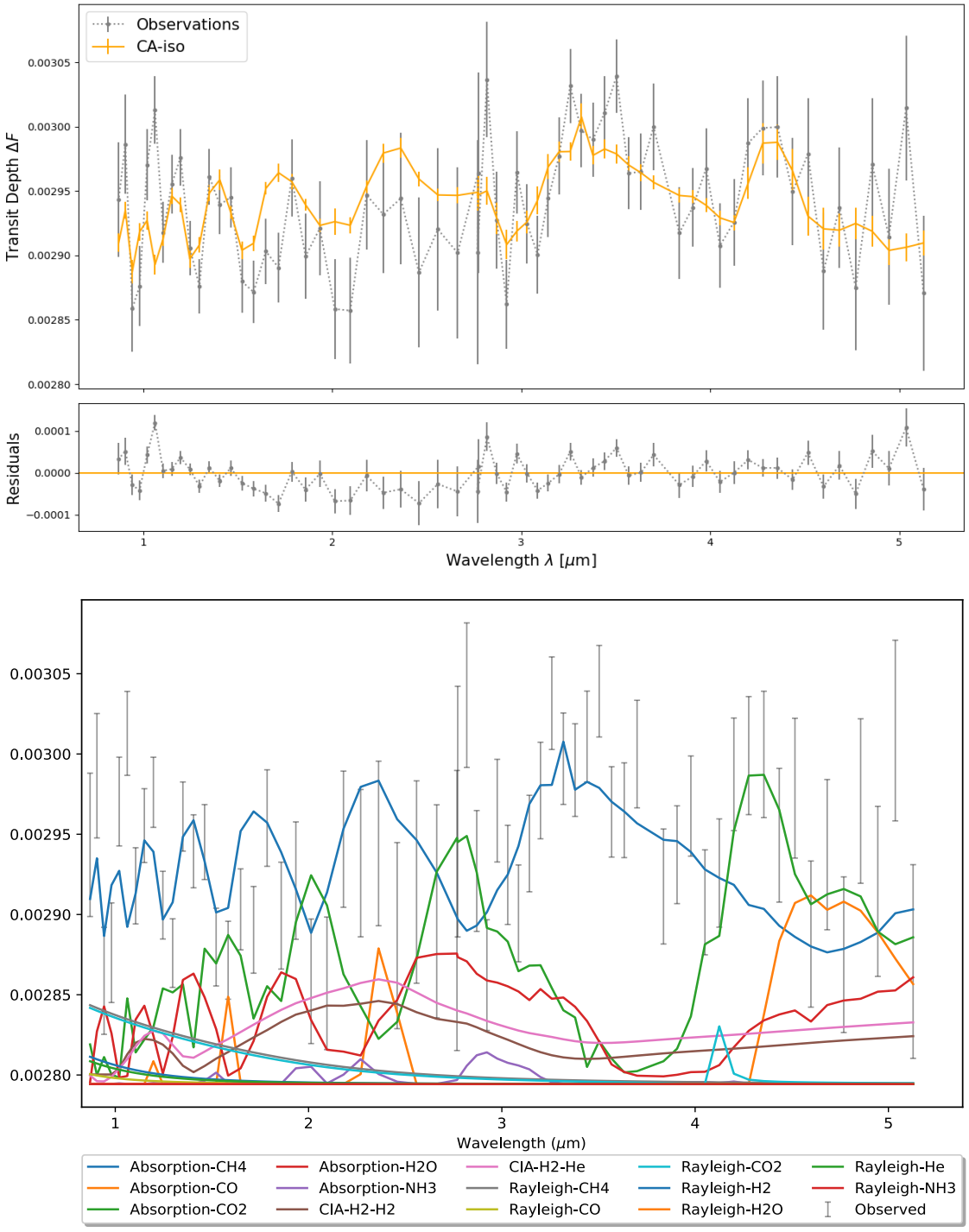

Figure 5.3: Top: fitted spectrum and residuals using the model with constant molecular abundances and an isothermal T-P profile. Bottom: spectrum showing absorption, Rayleigh scattering, and CIA contributions.

The fitted spectrum is shown at the top of the Figure (5.3). As a first observation, the model fits well at wavelengths above 3 μm, but the fit is less accurate at shorter wavelengths.

This result might be due to a too strict constraint on the isothermal temperature-pressure profile. I chose to consider this profile based on the work of Tsiaras et al. (2019) [88], but it probably is not the best choice. Another possible reason could be Rayleigh scattering, which is stronger at shorter wavelengths. It is possible that I did not include certain molecules in the atmosphere that contribute to scattering, such as those found in clouds. The spectrum showing different contributions from absorption, Rayleigh scattering, and CIA is displayed at the bottom of the Figure (5.3). CH_4 absorption features dominate at wavelengths below 4 μm and above 5 μm, while CO_2 absorption is more significant between 4 and 5 μm. In some parts of the spectrum, different absorption features overlap, particularly with H_2O and CO absorption. The NH_3 molecule is negligible in this model. H_2-He CIA is dominant for the CIA contribution, indicating an atmosphere rich in hydrogen and helium. However, the consistent presence of H_2-H_2 CIA suggests a higher amount of hydrogen. Rayleigh scattering, which is more prominent at lower wavelengths, is only observed for CH_4 and CO_2 molecules, confirming the dominance of these two molecules.

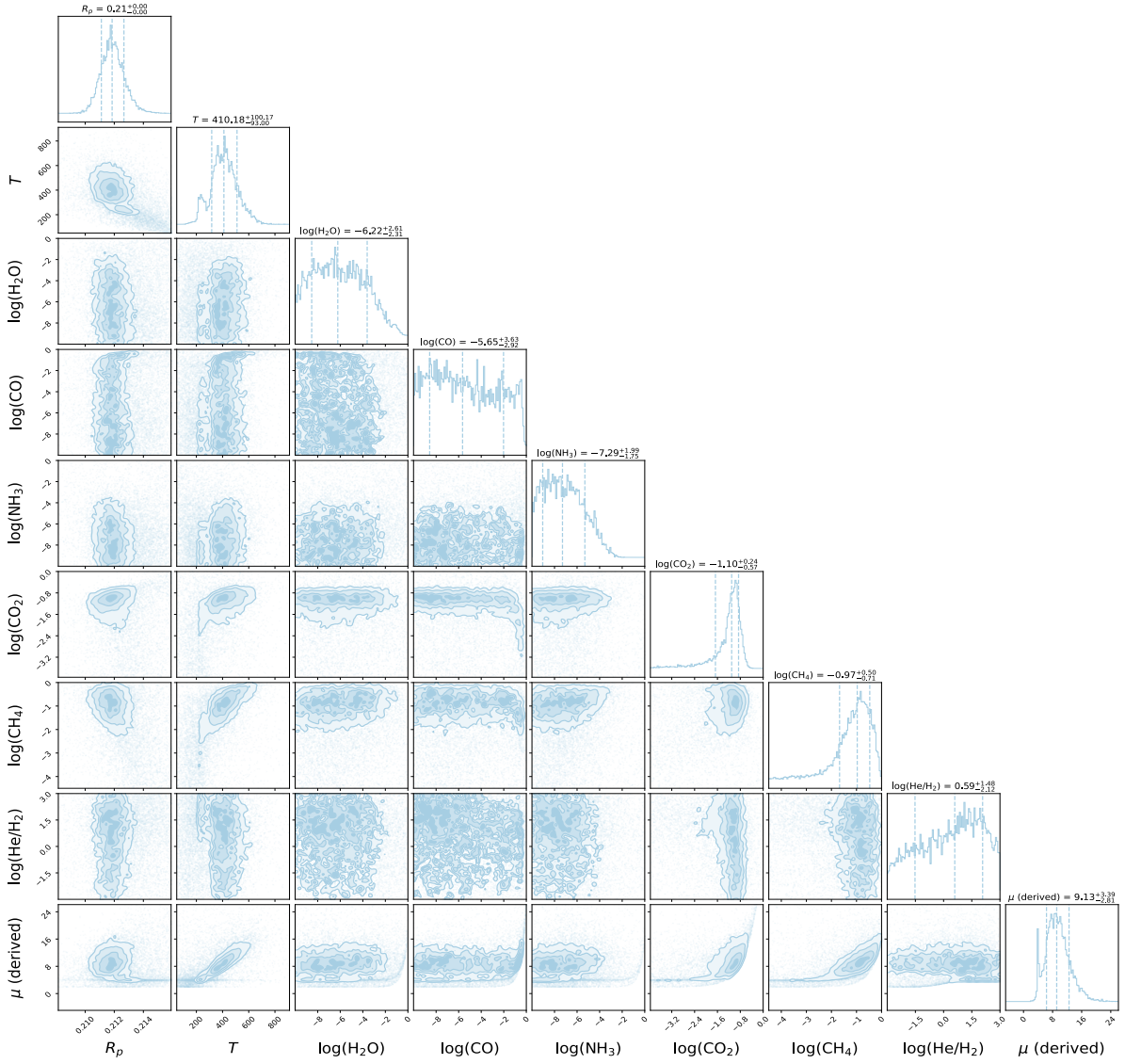

Figure 5.4: Atmospheric retrieval posterior distributions of *CA-iso* model, with planetary radius R_p (R_J), temperature T (K), logarithmic volume mixing ratios of H_2O, NH_3, CO, CO_2, CH_4, logarithmic He/H_2 ratio and the derived mean molecular weight μ (amu).

The posterior distributions shown in the Figure (5.4) support the qualitative statements mentioned earlier. Among the main CNO molecules, the logarithmic mixing ratios of CH_4 and CO_2 are well defined and show higher values than those reported by Madhusudhan et al. (2023) [63] in the Table (5.8). This indicates that the *CA-iso* model assumes a higher amount of CH_4 and CO_2 molecules in the atmosphere. H_2O, NH_3, and CO contribute little, despite having strong spectral features that should be detectable in the 0.9-5.2 μm range. The near-absence of H_2O detection contradicts earlier findings by Tsiaras et al. (2019) [88] using the HST WFC3 spectrum in the 1.1-1.7 μm range, which might be due to overlaps with CH_4. The lack of CO detection is not surprising given the low-temperature, H_2-rich atmosphere of the sub-Neptune K2-18b. The logarithmic mixing ratios align with Madhusudhan et al. (2023) [63] within 2σ. The posterior distribution for temperature from the isothermal profile is well defined, but its average value is higher than those reported by Madhusudhan et al. (2023) [63] and Tsiaras et al. (2019) [88]. A small second peak at temperatures closer to the values found in those studies can also be observed. The derived mean molecular weight μ suggests that a significant portion of the atmosphere consists of H_2 and He molecules, which is confirmed by the consistent logarithmic value of the He/H_2 ratio. Additionally, traces of other gases like CH_4 and CO_2 have been identified in this analysis and are included in the mean molecular weight μ of the atmosphere, which could explain why the value found with the *CA-iso* model is within 1σ of the value of Tsiaras et al. (2019) [88] reported in the Table (5.8), but slightly higher. The posterior distribution of μ also shows a small second peak at a lower value, possibly due to the second peak in the temperature distribution.

From this model, I observed two main problems: the spectrum fit at low wavelengths and the second peak in the temperature posterior distribution. To solve the first one I tried to add the presence of grey clouds analysing the model *CA-iso-clouds*, and for the second I assumed a complex temperature-pressure profile in the *CA-4p* model to break the degeneracy suggested by the bi-modal distribution of the temperature.

5.2.2 Simulation *CA-iso-clouds*

The retrieval *CA-iso-clouds* has the same characteristic as the previous *CA-iso* model.

Table 5.10: Fitting parameters of the model *CA-iso-clouds* with the uniform Bayesian priors and the posterior results.

Parameter	Bayesian prior	Posterior $\pm 1\sigma$	Description
R_p (R_J)	$\mathcal{U}(0.123, 0.984)$	0.21	Planet's radius
T (K)	$\mathcal{U}(50, 3500)$	$393.07^{+88.08}_{-61.20}$	Isothermal temperature
X_{H_2O}	$\mathcal{U}(10^{-10}, 1)$	$-6.11^{+2.69}_{-2.49}$	Mixing ratio of H_2O
X_{NH_3}	$\mathcal{U}(10^{-10}, 1)$	$-7.22^{+1.97}_{-1.78}$	Mixing ratio of NH_3
X_{CO}	$\mathcal{U}(10^{-10}, 1)$	$-5.72^{+3.12}_{-2.89}$	Mixing ratio of CO
X_{CO_2}	$\mathcal{U}(10^{-10}, 1)$	$-1.23^{+0.31}_{-1.20}$	Mixing ratio of CO_2
X_{CH_4}	$\mathcal{U}(10^{-10}, 1)$	$-1.36^{+0.67}_{-0.90}$	Mixing ratio of CH_4
He/H_2	$\mathcal{U}(10^{-3}, 10^{3})$	$1.21^{+1.18}_{-1.98}$	Helium-hydrogen ratio
μ (amu)	-	$7.20^{+3.60}_{-2.96}$	Mean molecular weight

It always considers an atmosphere with constant molecular abundances and an isothermal temperature-pressure profile but includes the effect of clouds to improve the fit of the spectrum. To do that I utilized the same free parameters of the *CA-iso* retrieval, reported in the Table (5.9), and a grey/flat-line clouds model.

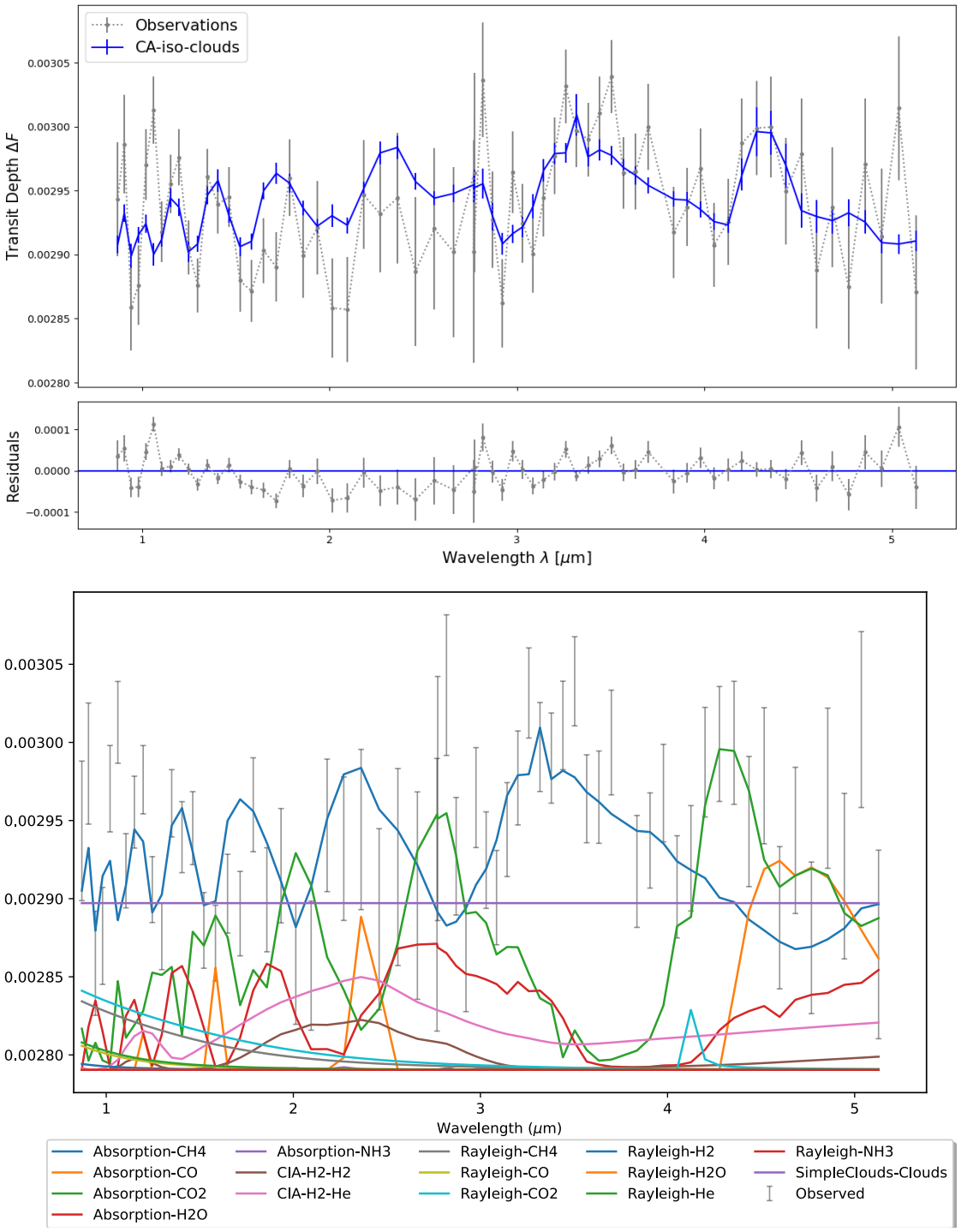

Figure 5.5: Top: fitted spectrum and residuals using the model with constant molecular abundances, an isothermal T-P profile and adding clouds. Bottom: spectrum showing a cloud layer and absorption, Rayleigh scattering, and CIA contributions.

The fitted spectrum shown at the top of the Figure (5.5) does not show clear differences compared to the spectrum in the Figure (5.3). In the contributions spectrum at the bottom of the Figure (5.5), the uniform cloud layer is overlapped by the absorption bands of CH_4, CO_2, and CO. This indicates that the presence of clouds is not enough to improve the fit of the spectrum at low wavelengths. However, the observations about the contributions of absorption, Rayleigh scattering, and CIA are the same as in the previous analysis.

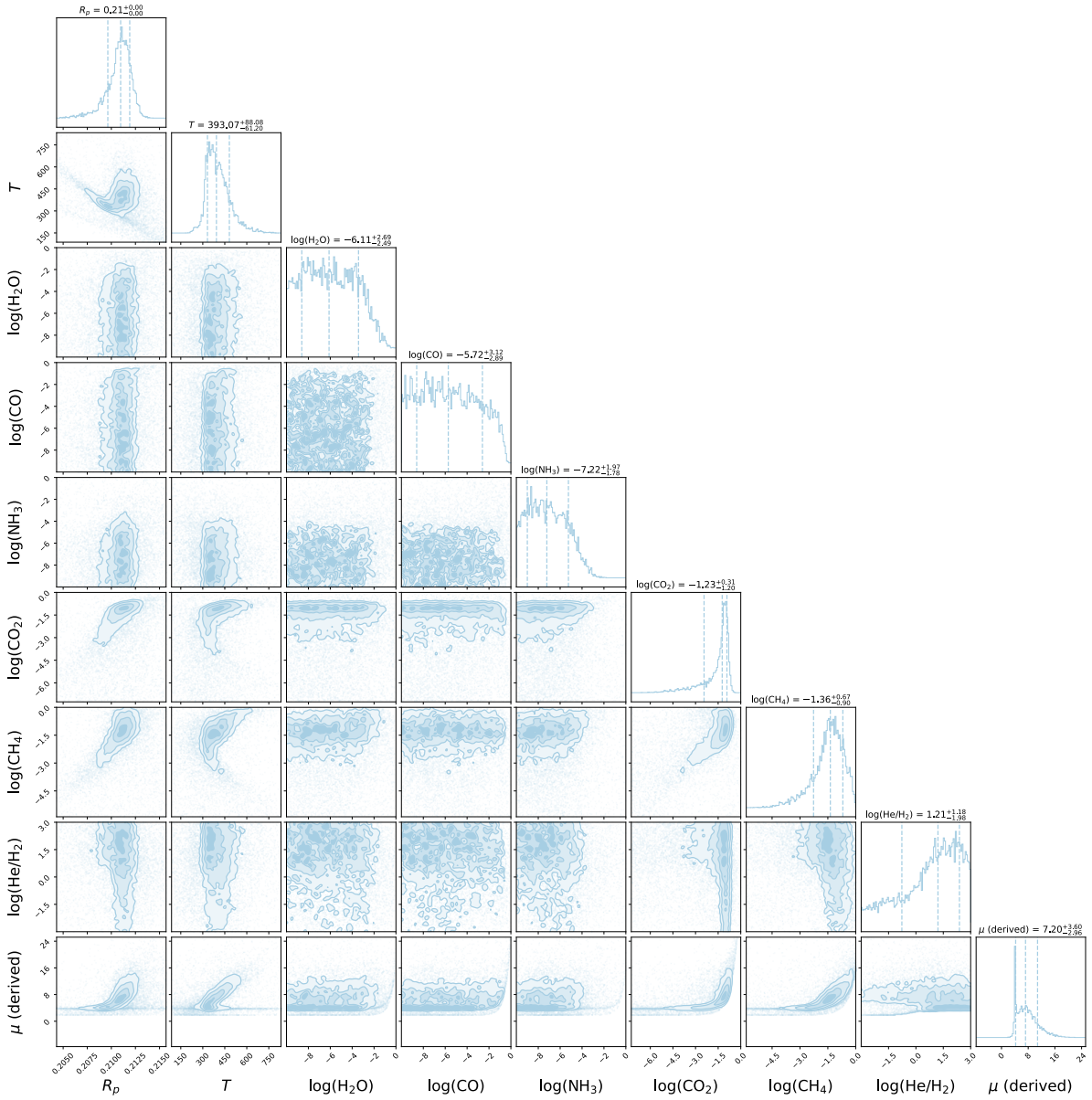

Figure 5.6: Atmospheric retrieval posterior distributions of *CA-iso-clouds* model, with planetary radius R_p (R_J), temperature T (K), logarithmic volume mixing ratios of H_2O, NH_3, CO, CO_2, CH_4, logarithmic He/H_2 ratio and the derived mean molecular weight μ (amu).

The posterior distributions of the *CA-iso-clouds* model are shown in the Figure (5.6). The distributions and logarithmic values of the mixing ratios are the same as those in the *CA-iso* model. Hence, the comparison with the values in the Table (5.8) remains unchanged.

However, the H_2/He distribution is more defined and shows a higher value than in the previous model, indicating a greater presence of hydrogen and helium molecules, which are concentrated in the clouds. The temperature distribution no longer shows the second peak visible in the Figure (5.4), but the mean molecular weight distribution still has a second, thin, higher peak. This could be due to the cloud layer containing a large amount of hydrogen and helium, which have a lower mean molecular weight between 2 and 4 amu, close to the value of the higher peak in the distribution.
Since adding a cloud layer does not improve the fit of the spectrum or the distributions, I decided to proceed with the analysis without including the cloud model. In the previous retrieval, an extra peak in the temperature distribution was visible, so I decided to use a more detailed temperature-pressure profile in the next analysis. This helps verify if the temperature (T) and molecular weight (μ) distributions change and become single-peaked.

5.2.3 Simulation *CA-4p*

The retrieval *CA-4p* refers to an atmosphere with constant molecular abundances and a 4-point temperature-pressure profile. To simulate this model I assumed as free parameters always the radius of the planet reported in the Table (5.3) and the mixing ratios of the molecules in the Table (5.6). For the temperatures and pressures, I assumed the 4-point T-P profile parameters in the Table (5.2). I decided to impose a strict upper bound for the temperature and pressure values and the planet's radius.

Table 5.11: Fitting parameters of the model *CA-4p* with the uniform Bayesian priors and the posterior results. The molecular abundances obtained from the posteriors are shown as $\log_{10}$ of volume mixing ratios.

Parameter	Bayesian prior	Posterior $\pm 1\sigma$	Description
R_p (R_J)	$\mathcal{U}(0.123, 0.984)$	0.21	Planet's radius
$T_{surface}$ (K)	$\mathcal{U}(50, 2000)$	$623.17^{+624.02}_{-377.17}$	Surface temperature
T_{top} (K)	$\mathcal{U}(50, 2000)$	$1024.59^{+485.86}_{-472.40}$	Top temperature
T_{point1} (K)	$\mathcal{U}(50, 2000)$	$294.12^{+268.71}_{-150.85}$	Temperature at point 1
T_{point2} (K)	$\mathcal{U}(50, 2000)$	$573.50^{+421.97}_{-337.33}$	Temperature at point 2
P_{point1} (bar)	$\mathcal{U}(10^7, 10^{-4})$	$2.90^{+2.53}_{-2.80}$	Pressure at point 1
P_{point2} (bar)	$\mathcal{U}(10^7, 10^{-4})$	$-1.27^{+2.65}_{-1.80}$	Pressure at point 2
X_{H_2O}	$\mathcal{U}(10^{-10}, 1)$	$-5.74^{+2.69}_{-2.62}$	Mixing ratio of H_2O
X_{NH_3}	$\mathcal{U}(10^{-10}, 1)$	$-6.75^{+2.12}_{-2.00}$	Mixing ratio of NH_3
X_{CO}	$\mathcal{U}(10^{-10}, 1)$	$-5.24^{+3.71}_{-3.01}$	Mixing ratio of CO
X_{CO_2}	$\mathcal{U}(10^{-10}, 1)$	$-0.98^{+0.23}_{-0.37}$	Mixing ratio of CO_2
X_{CH_4}	$\mathcal{U}(10^{-10}, 1)$	$-0.49^{+0.24}_{-0.35}$	Mixing ratio of CH_4
μ (amu)	-	$12.50^{+3.51}_{-3.00}$	Mean molecular weight

The fitted spectrum, shown in the upper Figure (5.7), follows the same pattern as the fitted spectrum in the upper Figure (5.3): the model fits well at wavelengths above 3 μm, but the fit is less accurate at shorter wavelengths. The justification related to the strict temperature-pressure profile is not sufficient, considering the passage from an isothermal to a 4-point profile for this analysis. The other case related to the Rayleigh scattering of not-included molecules remains a possible explanation.

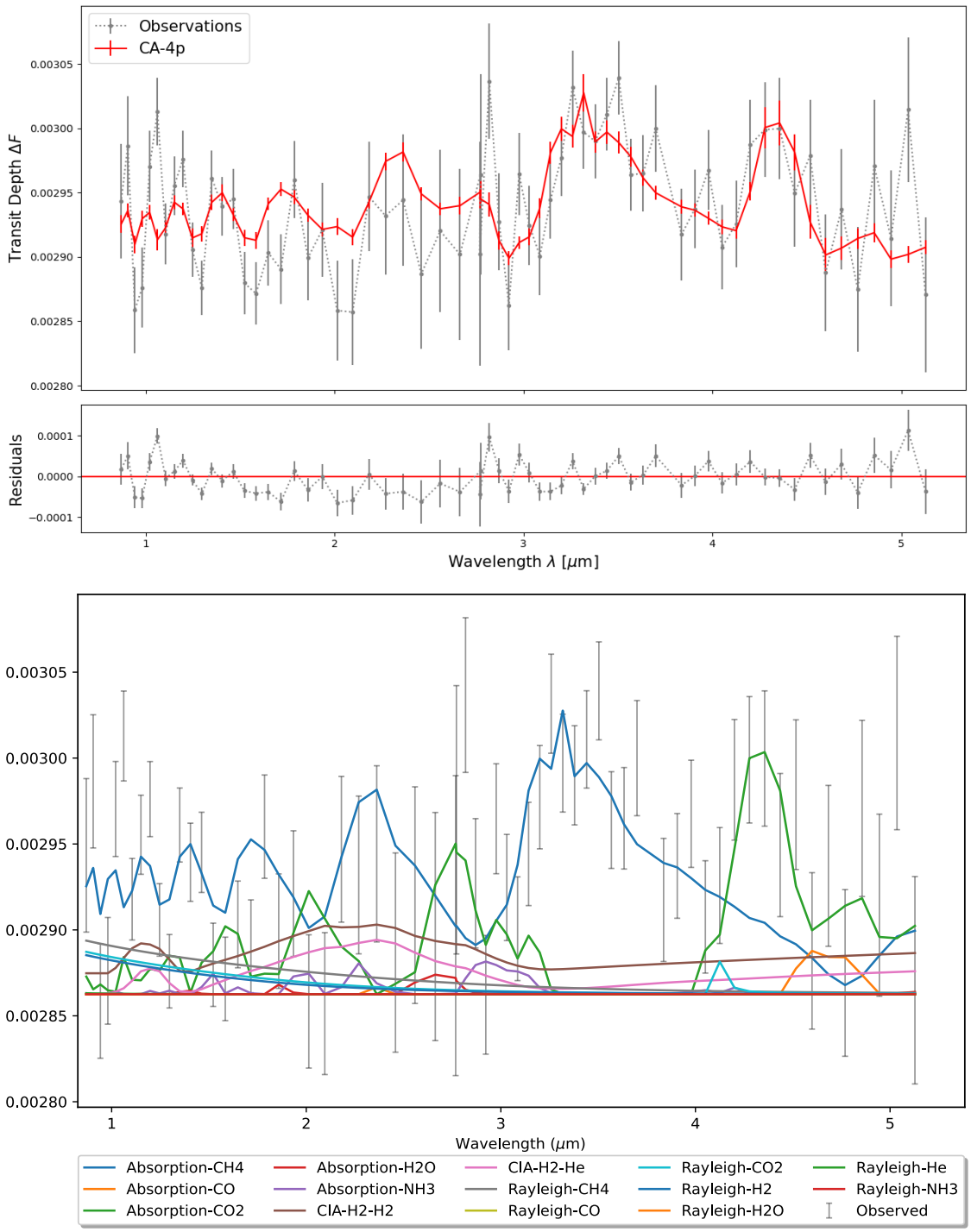

Figure 5.7: Top: fitted spectrum and residuals using the model with constant molecular abundances and a 4-point T-P profile. Bottom: spectrum showing absorption, Rayleigh scattering, and CIA contributions.

The spectrum with different contributions is shown at the bottom of the Figure (5.7). As in the previous retrieval, the absorption features of CH_4 and CO_2 dominate the spectrum. The molecular bands alternate in the central part of the spectrum, with CH_4 being dominant at lower wavelengths and CO_2 at higher wavelengths. The overlapping effects are less evident than at the bottom of the Figure (5.3), and NH_3, CO, and H_2O molecules are almost absent compared to CH_4 and CO_2 bands.

The dominance of CIA contributions has reversed compared to the previous analysis, with H_2-H_2 CIA being stronger than H_2-He CIA. Rayleigh scattering contributions are present for CH_4, CO_2, and H_2, confirming the dominance of these molecules, but are not strong enough to improve the fit at lower wavelengths.

The posterior distributions of the *CA-4p* model are shown in the Figure (5.9). The observations about the mixing ratios are similar to those in the previous analysis: CH_4 and CO_2 distributions are well-defined, while H_2O, NH_3, and CO are almost negligible. The H_2O mixing ratio is lower than the value reported by Tsiaras et al. (2019) [88], likely due to overlap with CH_4 bands, which obscure H_2O features. The mixing ratios of CH_4 and CO_2 are much higher than those found by Madhusudhan et al. (2023) [63], with lower uncertainties, indicating more precise constraints. This suggests that the *CA-4p* model assumes a higher concentration of CH_4 and CO_2 in the atmosphere, not only compared to the simulation by Madhusudhan et al. (2023) but also relative to the *CA-iso* and *CA-iso-clouds* models.

The 4-point temperature-pressure profile shown in the Figure (5.8) reveals two main temperature trends. Starting from the planet's surface, the temperature decreases until it reaches a pressure of 10^{-2} bar, then increases as it approaches the top of the atmosphere. The temperatures near the surface and the top of the atmosphere have higher uncertainties compared to the middle of the profile, as also reflected in the posterior distributions in the Figure (5.9).

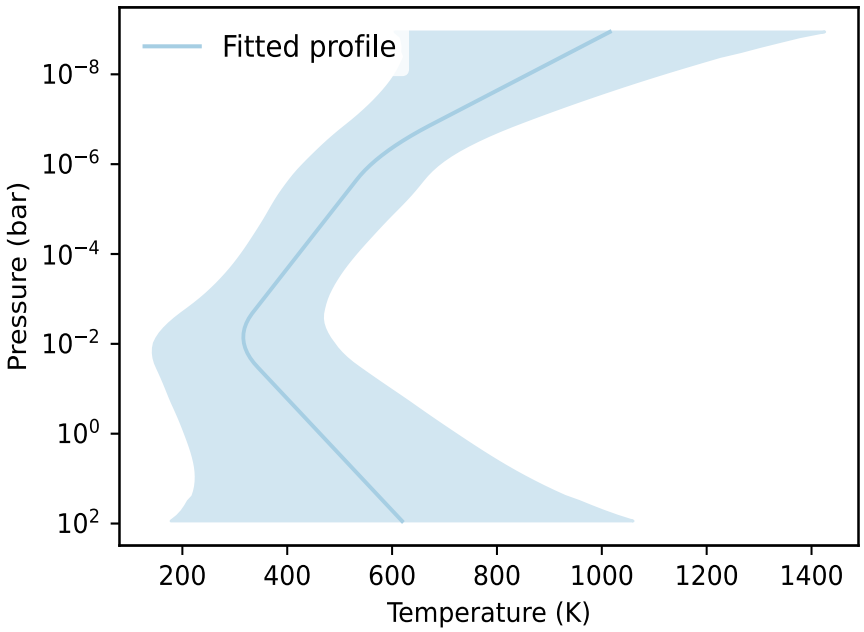

Figure 5.8: Fitted 4-point temperature-pressure profile of *CA-4p* model. The blue curve shows the median retrieved profile, while the lighter-blue contour denotes the 1σ interval.

The derived mean molecular weight μ is higher than the previous result but still within 1σ, and also higher than the value found by Tsiaras et al. (2019) [88], though within 2σ. This suggests that part of the atmosphere consists of H_2–He, along with significant amounts of other gases like CH_4 and CO_2, as confirmed by the higher mixing ratios of these molecules. The posterior distribution of μ shows only one peak, thanks to the change from an isothermal to a 4-point temperature-pressure profile.

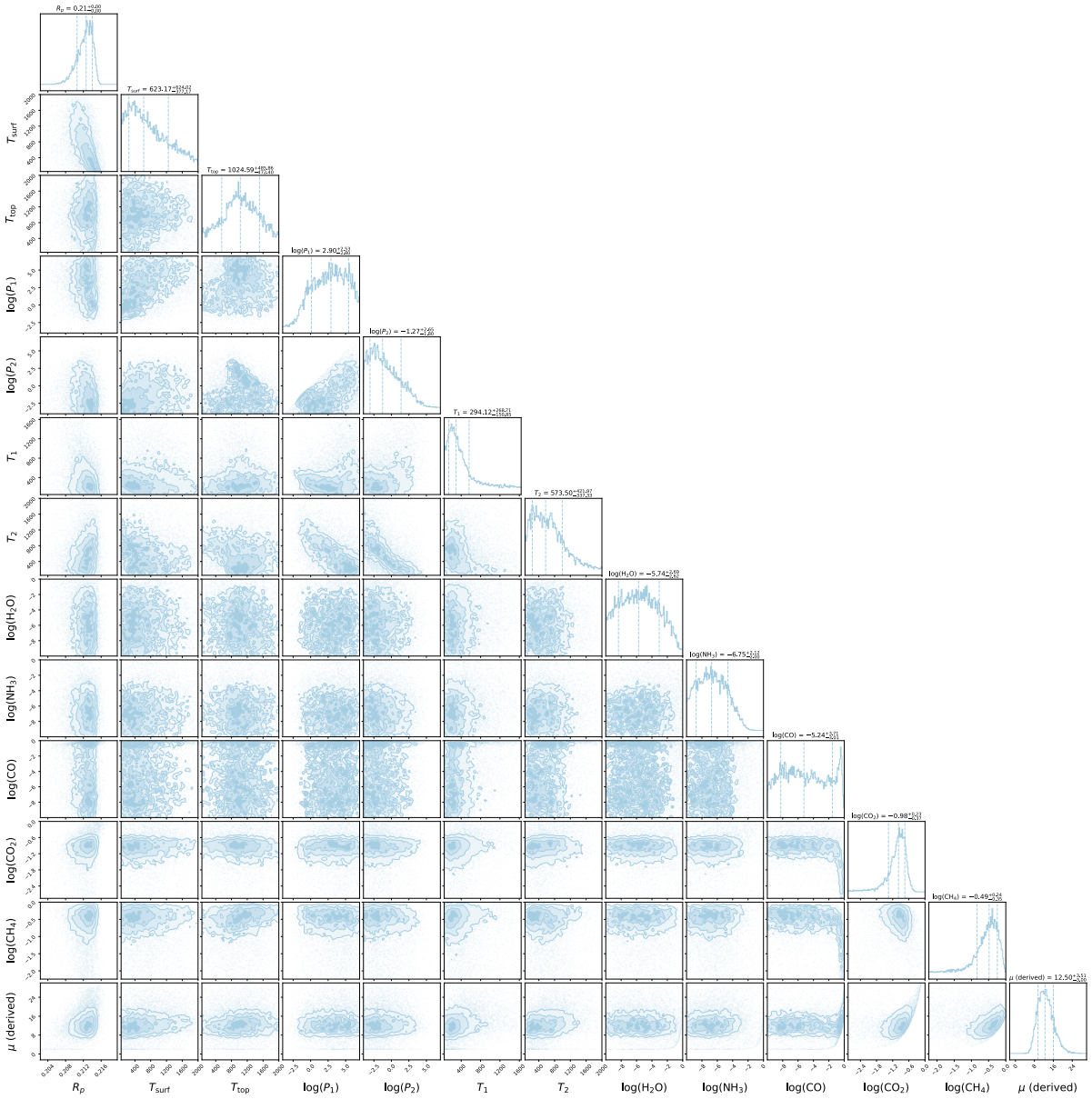

Figure 5.9: Atmospheric retrieval posterior distributions of *CA-4p* model, with planetary radius R_p (R_J), temperatures (K) and pressures (bar), logarithmic volume mixing ratios of H_2O, NH_3, CO, CO_2, CH_4, and the derived mean molecular weight μ (amu).

5.2.4 Simulation *CA-4p-N2*

To address the problem with fitting the spectrum at lower wavelengths and the weaker constraints on the 4-point temperature-pressure parameters, I decided to include the N_2 molecule, as suggested by Tsiaras et al. (2019) [88]. The addition of the N_2 molecule aims to investigate the presence of a secondary atmosphere, which is heavier and without volatile elements. The retrieval model *CA-4p-N2* has the same free parameters as the previous model *CA-4p* but with the addition of nitrogen N_2 in a ratio with hydrogen H_2. I also considered the He/H_2 ratio as in the *CA-iso* and *CA-iso-clouds* models and I chose to use their same priors, which have larger temperature parameters, to try and achieve better results.

Table 5.12: Fitting parameters of the model *CA-4p-N2* with the uniform Bayesian priors and the posterior results. The molecular abundances obtained from the posteriors are shown as $\log_{10}$ of volume mixing ratios.

Parameter	Bayesian prior	Posterior $\pm 1\sigma$	Description
R_p (R_J)	$\mathcal{U}(0.123, 0.984)$	0.21	Planet's radius
$T_{surface}$ (K)	$\mathcal{U}(50, 2000)$	$682.36^{+1055.60}_{-450.91}$	Surface temperature
T_{top} (K)	$\mathcal{U}(50, 3500)$	$1673.71^{+1012.14}_{-855.05}$	Top temperature
T_{point1} (K)	$\mathcal{U}(50, 3500)$	$492.24^{+573.05}_{-276.92}$	Temperature at point 1
T_{point2} (K)	$\mathcal{U}(50, 3500)$	$1046.13^{+1018.05}_{-627.38}$	Temperature at point 2
P_{point1} (bar)	$\mathcal{U}(10^7, 10^{-4})$	$3.20^{+2.32}_{-2.82}$	Pressure at point 1
P_{point2} (bar)	$\mathcal{U}(10^7, 10^{-4})$	$-1.31^{+2.93}_{-1.74}$	Pressure at point 2
X_{H_2O}	$\mathcal{U}(10^{-10}, 1)$	$-6.89^{+2.35}_{-1.95}$	Mixing ratio of H_2O
X_{NH_3}	$\mathcal{U}(10^{-10}, 1)$	$-7.74^{+1.91}_{-1.40}$	Mixing ratio of NH_3
X_{CO}	$\mathcal{U}(10^{-10}, 1)$	$-6.31^{+3.04}_{-2.32}$	Mixing ratio of CO
X_{CO_2}	$\mathcal{U}(10^{-10}, 1)$	$-3.42^{+1.65}_{-2.65}$	Mixing ratio of CO_2
X_{CH_4}	$\mathcal{U}(10^{-10}, 1)$	$-2.31^{+1.47}_{-2.31}$	Mixing ratio of CH_4
He/H_2	$\mathcal{U}(10^{-3}, 10^3)$	$-0.24^{+2.04}_{-1.69}$	Helium-hydrogen ratio
N_2/H_2	$\mathcal{U}(10^{-3}, 10^3)$	$1.24^{+1.09}_{-1.19}$	Nitrogen-hydrogen ratio
μ (amu)	-	$19.13^{+7.66}_{-7.11}$	Mean molecular weight

Observing the fitted spectrum in the Figure (5.10), the addition of the N_2 molecule does not seem to completely solve the problem of fitting at low wavelengths. The justification for the Rayleigh scattering of not-included molecules remains a possible explanation.

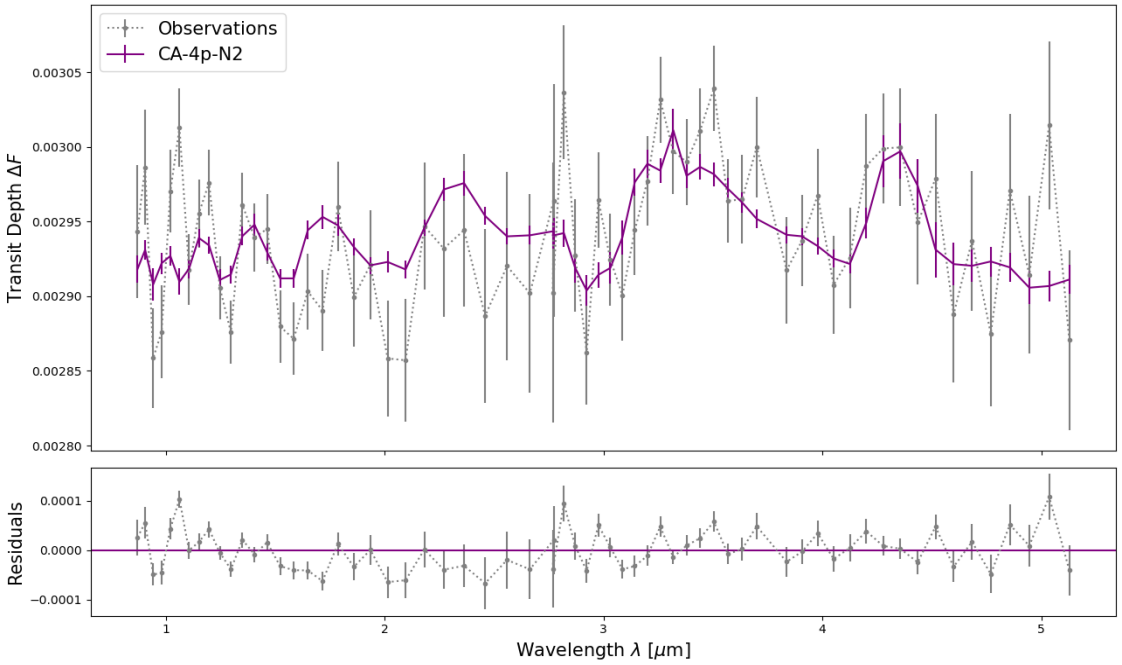

Figure 5.10: Fitted spectrum and residuals using the model with constant molecular abundances, adding the N_2 molecule and a 4-point T-P profile.

Observing the spectrum with different contributions at the top of the Figure (5.11), the Rayleigh scattering does not contribute to the features at shorter wavelengths. As before, the absorptions of CH_4 and CO_2 dominate and alternate in the spectrum. The overlapping effects are evident in the right part of the spectrum, between CO and CO_2 bands, and in the middle and left regions, between H_2O, CH_4 and CO_2 absorption bands and CIA contributions. The absorption of NH_3 is still almost absent compared to CH_4 and CO_2.

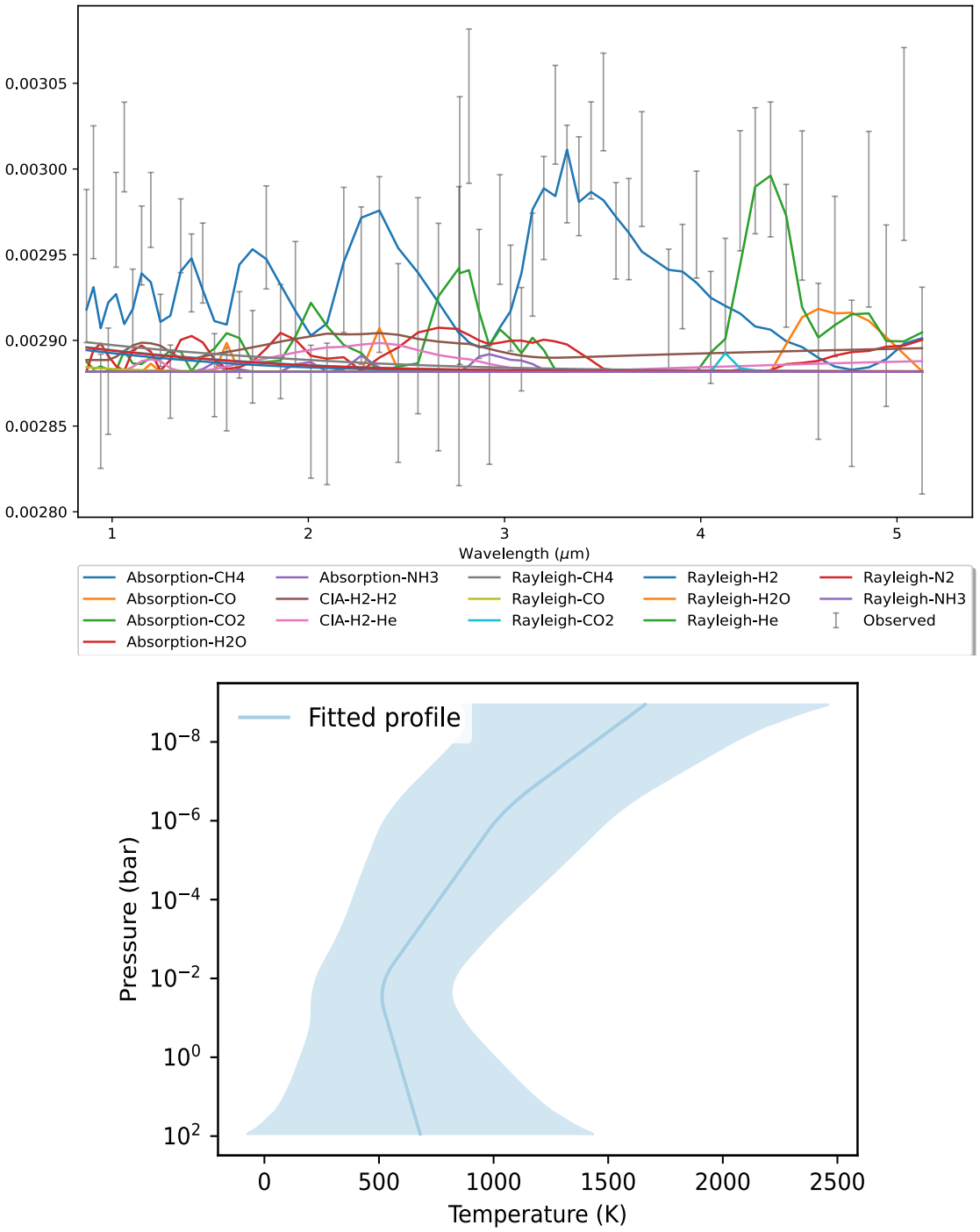

Figure 5.11: Top: spectrum showing absorption, Rayleigh scattering, and CIA contributions, using the model with constant molecular abundances, adding the N_2 molecule and a 4-point T-P profile. Bottom: fitted 4-point temperature-pressure profile of *CA-4p-N2* model. The blue curve shows the median retrieved profile, while the lighter-blue contour denotes the 1σ interval.

The temperature-pressure profile at the bottom of the Figure (5.11) shows a more constant trend compared to the profile in the Figure (5.8), but it has larger uncertainties. These uncertainties are also visible in the posterior distributions in the Figure (5.12), which are not well-defined and mostly show only the upper limits. This suggests that a simple isothermal profile might be more appropriate.

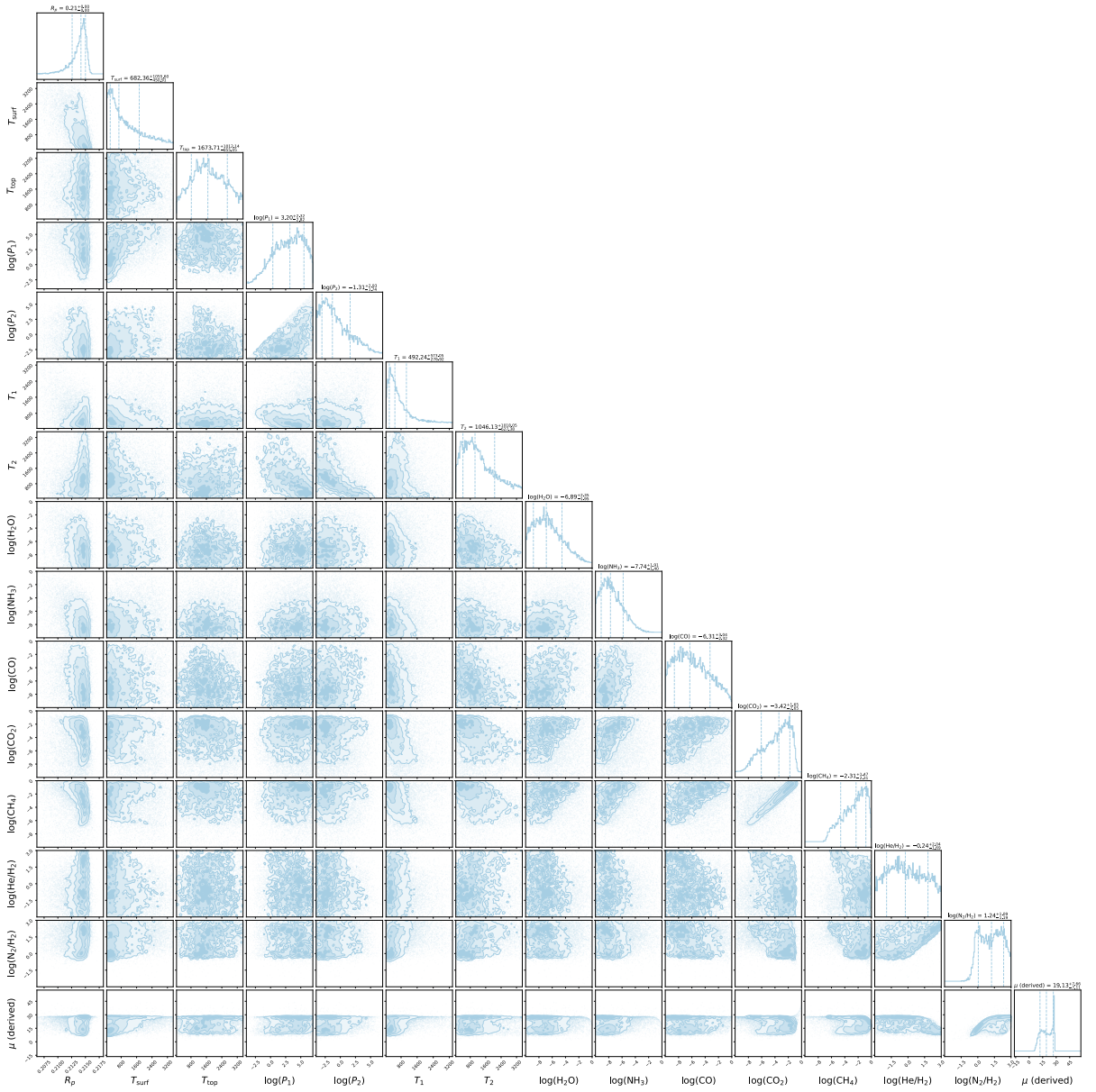

Figure 5.12: Atmospheric retrieval posterior distributions of *CA-4p-N2* model, with planetary radius R_p (R_J), temperatures (K) and pressures (bar), logarithmic volume mixing ratios of H_2O, NH_3, CO, CO_2, CH_4, logarithmic ratios of N_2 and He with H_2, and the derived mean molecular weight μ (amu).

This solution suggests that H_2O, NH_3 and CO are negligible in the atmosphere, while CH_4 and CO_2 distributions are better defined and compatible within 1σ with the values in the Table (5.8), relatively to the Madhusudhan et al. (2023) [63] results. A particular result regards the N_2/H_2 ratio distribution, which presents an irregular shape, assuming an initial bimodal trend.

This bimodal effect is visible also in the mean molecular weight μ distribution, as a consequence of the N_2/H_2 ratio distribution. The bimodal trend might occur due to degeneracies in the atmospheric model. Different N_2/H_2 ratios could produce similar spectral features when combined with temperature parameters with imprecise distributions and large uncertainties.

5.2.5 Simulation *CA-iso-N2*

Since the imprecise distributions and large uncertainties of the temperature-pressure profile in the Figure (5.12) and the visible constant trend at the bottom of the Figure (5.11), I decided to simulate the atmosphere using the same free parameters as the *CA-4p-N2* model, but with an isothermal T-P profile.

Table 5.13: Fitting parameters of the model *CA-iso-N2* with the uniform Bayesian priors and the posterior results. The molecular abundances obtained from the posteriors are shown as $\log_{10}$ of volume mixing ratios.

Parameter	Bayesian prior	Posterior $\pm 1\sigma$	Description
R_p (R_J)	$\mathcal{U}(0.123, 0.984)$	0.21	Planet's radius
T (K)	$\mathcal{U}(50, 3500)$	$570.26^{+256.83}_{-152.63}$	Isothermal temperature
X_{H_2O}	$\mathcal{U}(10^{-10}, 1)$	$-7.36^{+2.09}_{-1.71}$	Mixing ratio of H_2O
X_{NH_3}	$\mathcal{U}(10^{-10}, 1)$	$-8.15^{+1.76}_{-1.21}$	Mixing ratio of NH_3
X_{CO}	$\mathcal{U}(10^{-10}, 1)$	$-6.99^{+3.05}_{-1.99}$	Mixing ratio of CO
X_{CO_2}	$\mathcal{U}(10^{-10}, 1)$	$-4.29^{+2.11}_{-2.33}$	Mixing ratio of CO_2
X_{CH_4}	$\mathcal{U}(10^{-10}, 1)$	$-3.22^{+1.81}_{-1.97}$	Mixing ratio of CH_4
He/H_2	$\mathcal{U}(10^{-3}, 10^3)$	$0.94^{+1.53}_{-2.62}$	Helium-hydrogen ratio
N_2/H_2	$\mathcal{U}(10^{-3}, 10^3)$	$1.33^{+0.98}_{-1.46}$	Nitrogen-hydrogen ratio
μ (amu)	-	$13.02^{+10.74}_{-4.48}$	Mean molecular weight

The fitted spectrum shown in the Figure (5.13) provides a more accurate fit at shorter wavelengths than previous results. This improvement might be due to a better match between the isothermal profile and the significant presence of N_2 in the atmosphere.

Looking at the spectrum with different contributions in the Figure (5.14), CH_4 absorption is still the most prominent feature, but other contributions are also observable. The absorption bands of H_2O are more consistent than in the previous analysis, allowing for significant overlap at wavelengths below 3.5 μm. Another strong indication is the H_2/He CIA in the same part of the spectrum as the H_2O molecules, suggesting an atmosphere rich in hydrogen and helium. NH_3 absorption is still nearly absent, while CO appears only on the right part of the spectrum. A key difference from the previous analysis is the dominance of Rayleigh scattering from N_2 at shorter wavelengths, indicating the presence of nitrogen in the atmosphere. Since the presence of nitrogen does not allow such a notable improvement in the spectrum, the presence of other molecules, also capable of overlapping with the contribution of nitrogen, cannot be excluded. Rayleigh scattering is also observed for He molecules, confirming a high amount of helium in the atmosphere.

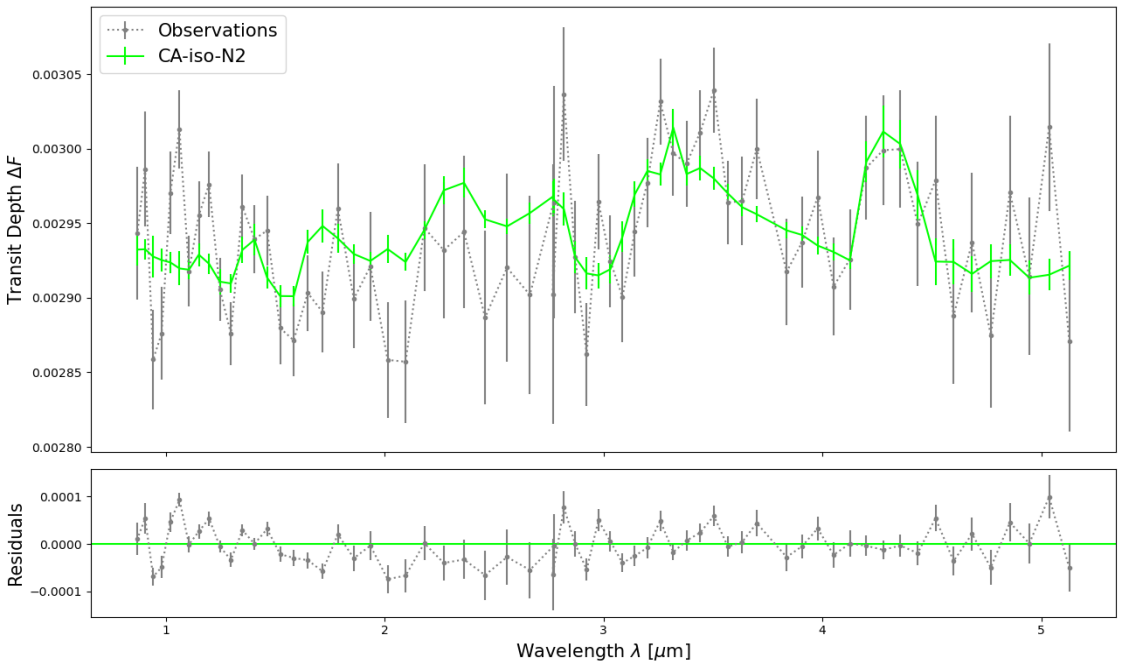

Figure 5.13: Fitted spectrum and residuals using the model with constant molecular abundances, adding the N_2 molecule and an isothermal profile.

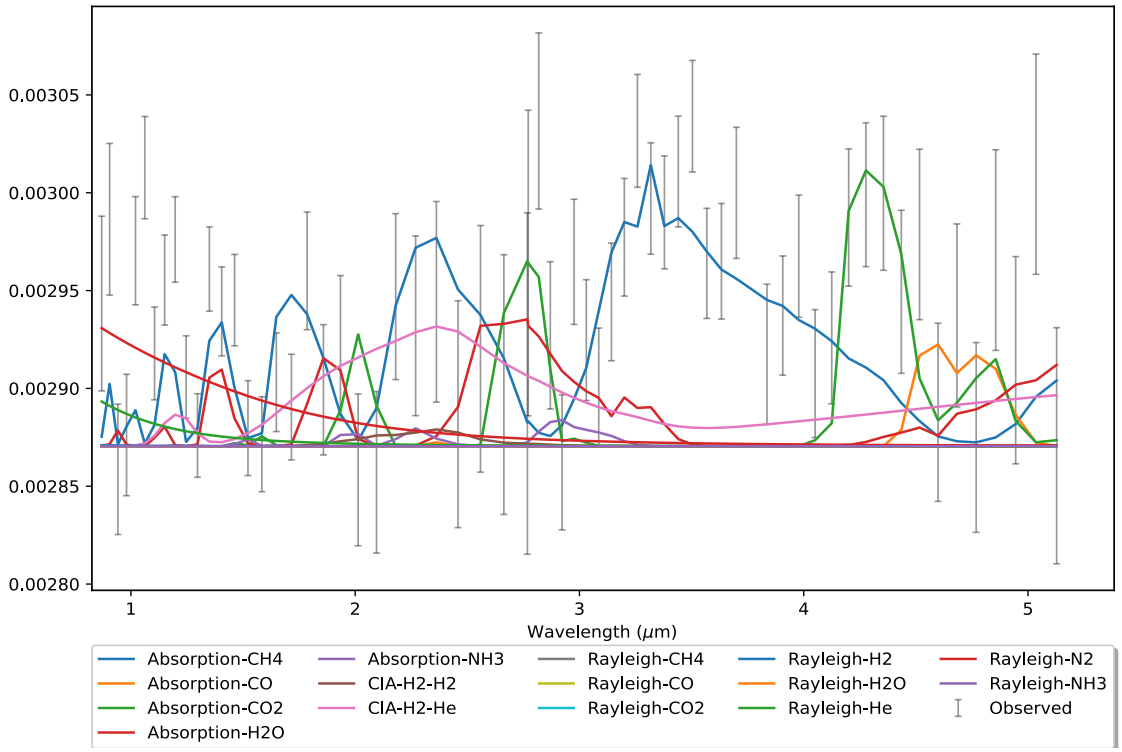

Figure 5.14: Spectrum of the planet K2-18b, using the model with constant molecular abundances, adding the N_2 molecule and an isothermal T-P profile, showing absorption, Rayleigh scattering, and CIA contributions.

The posterior distributions of the *CA-iso-N2* model are shown in the Figure (5.15). The temperature distribution is well-defined and is consistent within 1σ with the results of the *CA-iso* model. However, its value is almost double that of Madhusudhan et al. (2023) [63] and Tsiaras et al. (2019) [88], as reported in table (5.8), within 3σ, indicating higher uncertainties.

The mixing ratios of H_2O, NH_3, and CO remain almost negligible, showing only upper limits. Meanwhile, the CH_4 and CO_2 distributions are better defined and consistent within 1σ with the values from Madhusudhan et al. (2023) [63]. However, they have a broader shape, and larger uncertainties, and their mixing ratios are lower compared to the previous analysis.

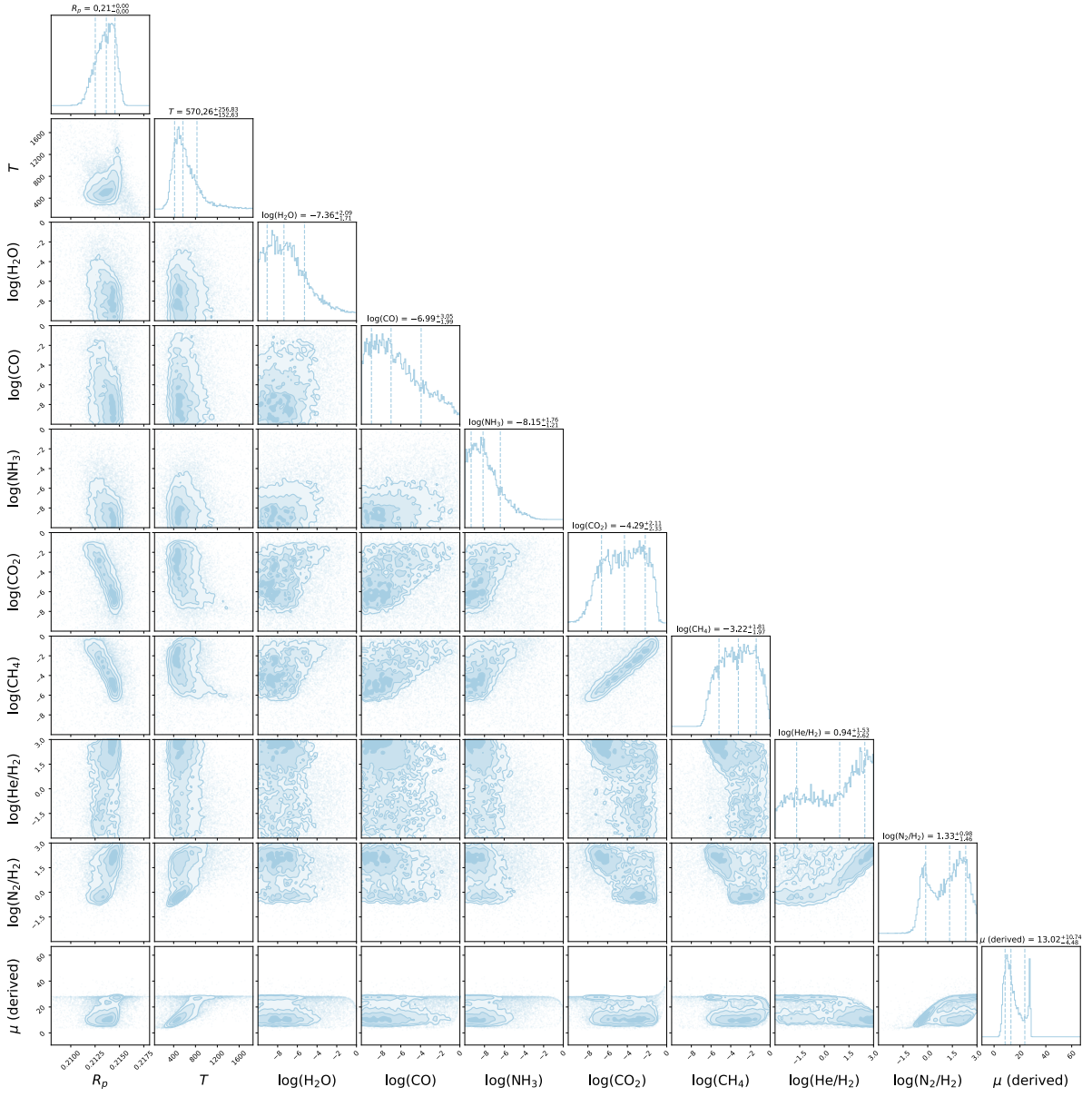

Figure 5.15: Atmospheric retrieval posterior distributions of *CA-iso-N2* model, with planetary radius R_p (R_J), temperatures (K) and pressures (bar), logarithmic volume mixing ratios of H_2O, NH_3, CO, CO_2, CH_4, logarithmic ratios of N_2 and He with H_2, and the derived mean molecular weight μ (amu).

As before, the N_2/H_2 ratio distribution shows a bimodal trend, which is also visible, though less pronounced, in the He/H_2 ratio distribution. Consequently, this bimodal effect is reflected in the mean molecular weight (μ) distribution. Since the isothermal profile was used and has a well-constrained distribution without a second peak, it can be excluded as the cause of the bimodal effect in the distributions.
One possible statistical explanation is that the two peaks in the N_2/H_2 distribution represent different amounts of N_2 in the atmosphere. This suggests the presence of two distinct solutions that are both consistent with the model. The bimodality may suggest a weak degeneration of the solution for an atmospheric model, where different combinations of atmospheric parameters and gases with overlapping spectral features cause the retrieval process to be more complex, resulting in a bimodal outcome [57]. In the Figure (5.14), various overlaps are visible and are more prominent than in other retrievals, which is why the bimodal distributions are more observable in the current retrieval.
Another possible explanation involves a transition between atmospheric conditions. The exoplanet's atmosphere might be in a transitional state or capable of existing in two different stable states, such as a hybrid atmosphere. For instance, the peak at 27 amu, compatible with the result of Madhusudhan et al. (2023) [63] within 1σ, might correspond to an atmosphere with a higher proportion of nitrogen as a secondary atmosphere, while the other peak at 11 amu, compatible with the result of Tsiaras et al (2019) [88] within 2σ, corresponds to a hydrogen-dominated primary atmosphere [84].

5.2.6 Simulation *EC-4p*

From the previous simulations, the fit of the spectrum is not very accurate at low wavelengths. This may be due to not-considered molecules, especially their Rayleigh scattering. To improve the results, I tried to simulate an atmospheric model of the planet using chemical equilibrium, through the `FastChem` plug-in of TauREx3. Specifically, I considered an atmosphere rich in hydrogen and helium, with the planet's metallicity and carbon-oxygen ratio as free parameters, reported in the Table (5.7). The temperature-pressure profile used is based on 4 points to maintain physical accuracy and to evaluate whether the isothermal approximation is still suitable in this case.

Table 5.14: Fitting parameters of the model *EC-4p* with the uniform Bayesian priors and the posterior results. The molecular abundances obtained from the posteriors are shown as $\log_{10}$ of volume mixing ratios.

Parameter	Bayesian prior	Posterior $\pm 1\sigma$	Description
R_p (R_J)	$\mathcal{U}(0.123, 0.492)$	0.21	Planet's radius
$T_{surface}$ (K)	$\mathcal{U}(50, 2000)$	$628.05^{+668.44}_{-385.74}$	Surface temperature
T_{top} (K)	$\mathcal{U}(50, 2000)$	$644.81^{+495.82}_{-341.94}$	Top temperature
T_{point1} (K)	$\mathcal{U}(50, 2000)$	$454.17^{+268.33}_{-254.86}$	Temperature at point 1
T_{point2} (K)	$\mathcal{U}(50, 2000)$	$550.68^{+154.00}_{-202.19}$	Temperature at point 2
P_{point1} (bar)	$\mathcal{U}(10^7, 10^{-4})$	$4.47^{+1.67}_{-4.10}$	Pressure at point 1
P_{point2} (bar)	$\mathcal{U}(10^7, 10^{-4})$	$-1.14^{+4.44}_{-2.06}$	Pressure at point 2
Z ($(Z/H)_\odot$)	$\mathcal{U}(0.1, 10^3)$	$284.13^{+143.17}_{-68.36}$	Metallicity
C/O	$\mathcal{U}(0.01, 2)$	0.07 ± 0.10	Carbon-oxygen ratio
μ (amu)	-	$13.91^{+7.27}_{-4.18}$	Mean molecular weight

The fit obtained from the simulation of the atmosphere in chemical equilibrium, shown at the top of the Figure (5.16), has the same bad issues as the previous fits at low wavelengths, appearing even worse at fitting the spectrum in the wavelength range between 1 and 2 μm. Regarding the temperature-pressure profile, the 4-point profile can be easily approximated to an isothermal profile, as seen at the bottom of the Figure (5.16) and as deduced in the previous retrievals. This is also supported by the posterior distributions of temperatures, which are imprecise and have large uncertainties, particularly for the top and surface temperatures, as reported in Table (5.14) and visible in the Figure (5.17).

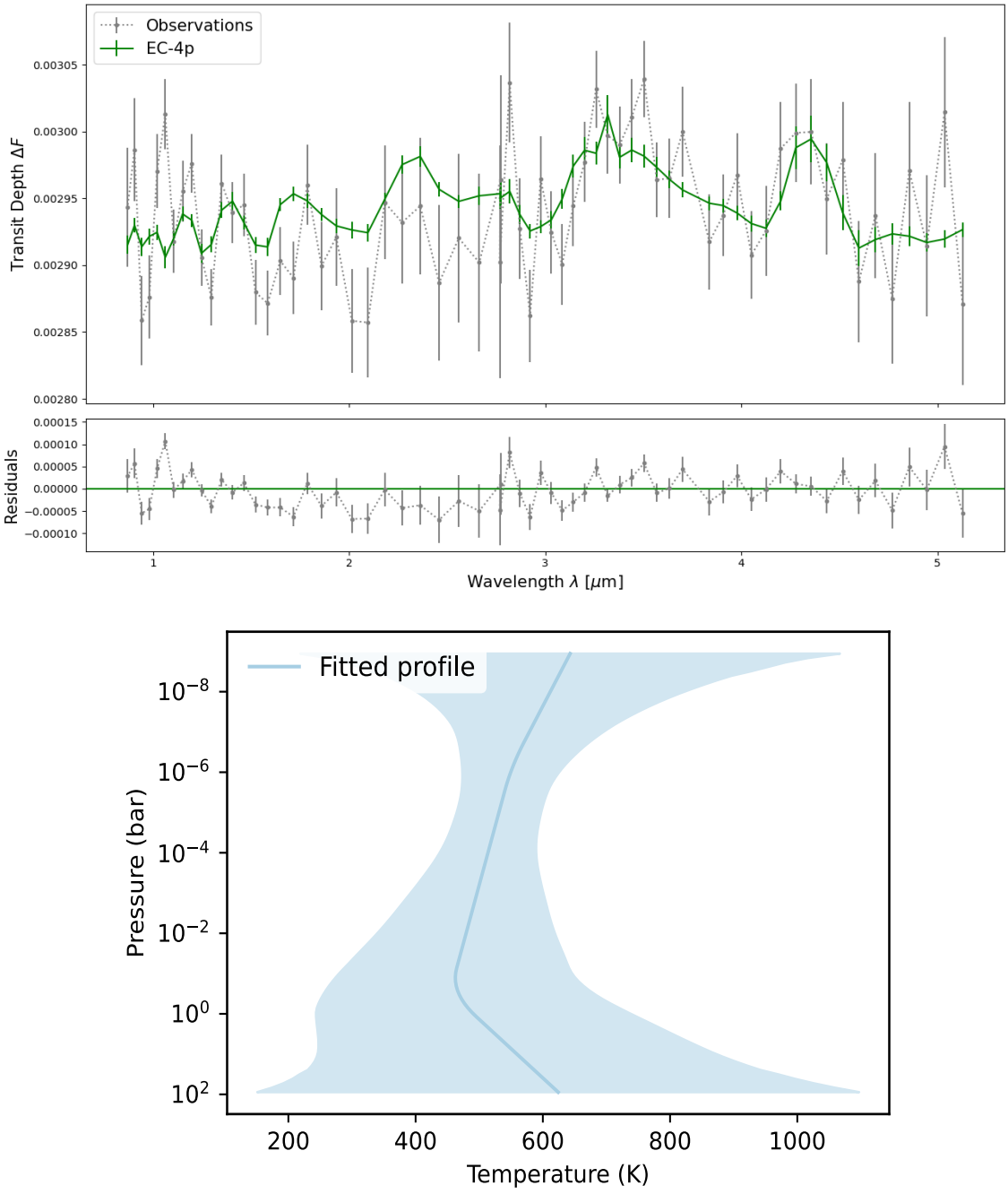

Figure 5.16: Top: fitted spectrum and residuals using the model with equilibrium chemistry. Bottom: 4-point T-P profile. The blue curve shows the median retrieved model spectrum, while the lighter-blue contour denotes the 1σ interval.

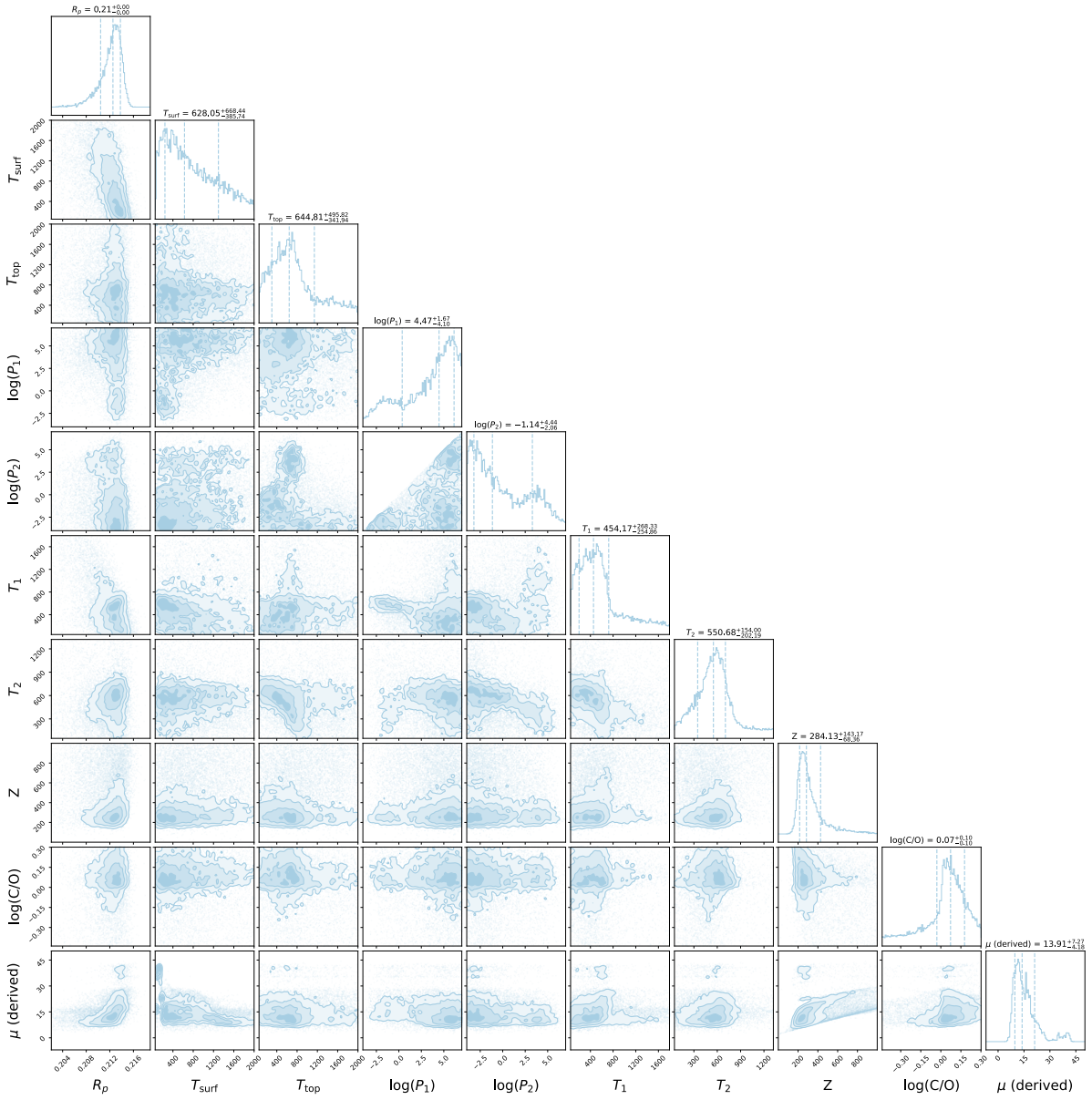

Figure 5.17: Atmospheric retrieval posterior distributions with planetary radius R_p (R_J), temperatures (K) and pressures (bar), metallicity ($Z/H_\odot$), C/O ratio and the derived mean molecular weight μ (amu).

The distributions of metallicity and the C/O ratio are well-defined but have significant uncertainties. The average metallicity value is higher than what is reported in the Table (5.8) within 2σ, while the C/O ratio is much lower compared to the value in the same table, which was assumed to be equal to the solar ratio. At very low C/O ratios and high metallicity, the primary component of an atmosphere is O_2 [59]. However, the presence of O_2 contradicts the previous retrievals and the expected result due to the high carbon content, especially in the form of CH_4 molecules. One possible explanation could be related to the limitations of the equilibrium chemistry model used through the FastChem package [80]. FastChem simplifies calculations by assuming local thermodynamic equilibrium (LTE) and does not account for non-equilibrium processes such as photochemistry or vertical mixing, which can significantly affect the abundances of molecules like O_2 and CH_4. This leads to an overestimation of metallicity and an underestimation of C/O ratio.

Another factor influencing the metallicity and C/O distributions could be the large temperature uncertainties from the 4-point profile. These uncertainties might also explain why the average molecular weight of the atmosphere aligns with previous values but presents greater uncertainties.

Since the posterior distributions obtained under the assumption of equilibrium chemistry, as shown in the Figure (5.17), do not align with the results from previous retrievals, and the fitted spectrum has not improved, as seen in the Figure (5.16), I did not explore more complex models with chemical equilibrium.

5.3 Atmospheric model comparison

To simulate the atmosphere of K2-18b I decided to utilize different models, such as the molecular constant abundances and the chemical equilibrium, with different temperature-pressure profiles, such as the isothermal and the 4-point, also adding molecules and clouds to fit better the spectrum at lower wavelengths. In the following figure, I report an overlap of the various models used to fit the transmission spectrum of the sub-Neptune K2-18b.

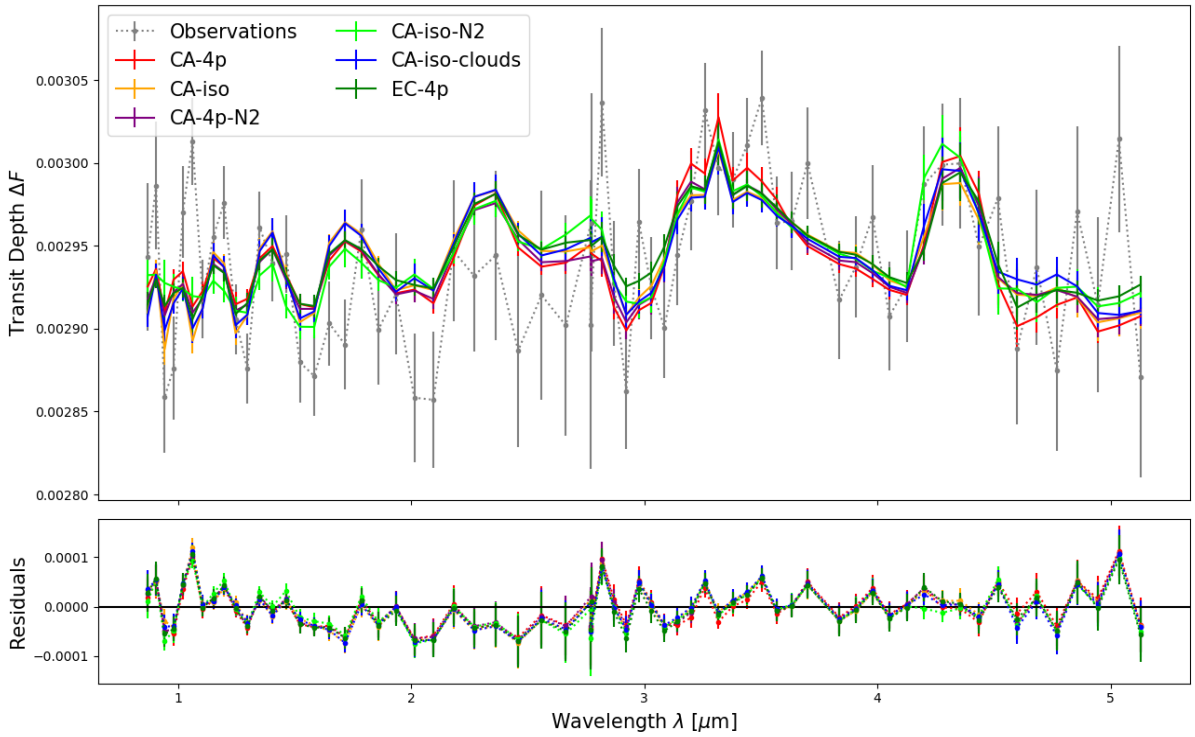

Figure 5.18: Overlap of the various models used to fit the transmission spectrum of K2-18b, with the residuals.

From a first observation and a qualitative comparison, it is impossible to establish the best model compared to the others. To define the best model which describes the atmosphere it is necessary to do a statistical comparison. I utilized MULTINEST [32], a nested sampling, as an optimizer of the various retrievals. One output of the optimizer is the Bayesian evidence of each model, which can be used to compute the ADI (Atmospheric Detectability Index) [87], the positively defined Bayes Factor, theoretically described in Section (4.2.2). The logarithmic value of the Bayesian evidence of the models is reported in the Table (5.15).

Table 5.15: Logarithmic Bayesian evidence of all retrievals used to simulate the atmosphere of K2-18b. Each value has the uncertainties of $\pm 1\sigma$.

log(Bayesian evidence $\pm 1\sigma$)		
CA-iso	*CA-iso-clouds*	*CA-4p*
535.92 ± 0.11	536.99 ± 0.11	538.25 ± 0.11
CA-4p-N2	*CA-iso-N2*	*EC-4p*
535.72 ± 0.12	535.87 ± 0.11	536.48 ± 0.12

ADI's value greater than 3 means a significant detection of an atmosphere, while ADIs below 3 are unable to favour one model over the other. To compute the ADIs it is necessary to subtract the logarithmic Bayesian evidence of two models, following the expression (4.11). In the Table (5.16) are reported ADIs from all the possible combinations capable of giving positive results from the subtraction of the logarithmic value of the Bayesian evidence.

Table 5.16: Atmospheric Detectability Index (ADI) from all possible combinations of the various models used to simulate the atmosphere of K2-18b.

Atmospheric Detectability Index (ADI)		
(*CA-iso-clouds*) - (*CA-iso*)	(*EC-4p*) - (*CA-iso-N2*)	(*CA-4p*) - (*EC-4p*)
1.07 ± 0.22	0.61 ± 0.23	1.77 ± 0.23
(*CA-iso-clouds*) - (*CA-4p-N2*)	(*CA-4p*) - (*CA-4p-N2*)	(*EC-4p*) - (*CA-iso*)
1.27 ± 0.23	2.53 ± 0.23	0.56 ± 0.23
(*CA-iso-clouds*) - (*CA-iso-N2*)	(*CA-4p*) - (*CA-iso-N2*)	(*CA-4p*) - (*CA-iso*)
1.12 ± 0.22	2.38 ± 0.22	2.33 ± 0.22
(*CA-iso-clouds*) - (*EC-4p*)	(*CA-iso-N2*) - (*CA-4p-N2*)	(*EC-4p*) - (*CA-4p-N2*)
0.51 ± 0.23	0.15 ± 0.23	0.76 ± 0.24
(*CA-4p*) - (*CA-iso-clouds*)	(*CA-iso*) - (*CA-iso-N2*)	(*CA-iso*) - (*CA-4p-N2*)
1.77 ± 0.22	0.05 ± 0.22	0.20 ± 0.23

Observing the ADIs in the Table (5.16) the results are lower than 3 in all cases. For this reason, it is impossible to favour one model over another because the statistical comparison seems to be inconclusive. Since there are several competing models that can be possible and probable to describe the atmosphere of K2-18b, I identify the best model following the principle of simplicity and economy of an explanation through Occam's razor: the simplest model compatible with the observation evidence can be preferred [86]. Assuming Occam's razor principle, the simplest model is *CA-iso*, which describes the atmosphere of the sub-Neptune K2-18b considering constant molecular abundances and an isothermal T-P profile.

The comparison of this model with the ones of Madhusudhan et al. (2023) [63] and Tsiaras et al. (2019) [88] is found in Section (5.2.1). The results indicate that the *CA-iso* model suggests higher amounts of CH_4 and CO_2 in the atmosphere compared to earlier findings by Madhusudhan et al. (2023) [63], while H_2O, NH_3, and CO contribute less, despite their strong spectral signatures. The absence of H_2O detection contrasts with previous results from Tsiaras et al. (2019) [88] and could be due to overlap with CH_4. CO is not expected due to the planet's low temperature and H_2-rich atmosphere. The temperature profile is well defined, though slightly higher than previous studies, with a secondary peak closer to their values.

The goal of my work is to characterize the atmosphere of K2-18b to determine whether it has a primary or secondary atmosphere. This is estimated by examining the mean molecular weight (μ). A primary atmosphere, composed mainly of hydrogen and helium, is expected to have a molecular weight between 2 and 4 amu, while a secondary atmosphere would have a higher μ.

Tsiaras et al. (2019) [88] reported a value of $8.05^{+3.49}_{-2.19}$ amu, and Madhusudhan et al. (2023) [63] found a value of 28.47 amu. Neither result indicates a pure primary atmosphere, as both suggest the presence of heavier elements. In particular, Madhusudhan et al. (2023) [63] point to a secondary atmosphere rich in methane and carbon dioxide.

The result from the *CA-iso* model simulation is $9.13^{+3.39}_{-2.81}$ amu, confirming a mostly hydrogen and helium atmosphere with traces of CH_4 and CO_2, consistent with Tsiaras et al. (2019) [88] within 1σ. Although this value is higher than expected for a pure primary atmosphere, the abundance of hydrogen and helium suggests it is not a fully secondary atmosphere dominated by heavier elements. A possible explanation is that the planet has a transition, or hybrid, atmosphere between primary and secondary phases.

This can be attributed to the fact that K2-18b orbits an M-dwarf. Planets around M-dwarfs may experience long-term atmospheric transitions due to the star's long life and extended stellar activity, which can drive atmospheric loss over billions of years [77]. This process can cause a transition from a hydrogen-rich atmosphere to one containing heavier molecules such as water, CO_2, or nitrogen [83].

However, with all this work and the current JWST observations, it is not yet possible to definitively distinguish between atmospheric models for K2-18b. The data we have suggests that the planet might have a transitional or hybrid atmosphere, but we need more observations to fully understand the complex physics of this sub-Neptune. Future JWST observations, particularly with expanded wavelength coverage and higher precision, will be essential to find more details of K2-18b's atmosphere and refine our understanding of its evolutionary stage. These observations could help define whether the atmosphere is primarily, composed of lighter elements like hydrogen and helium, or if it is in transition toward a secondary atmosphere rich in heavier molecules, such as methane and carbon dioxide. With more data, we will be able to better understand the planet's temperature, molecular weight, and overall atmospheric composition, bringing us closer to fully understanding the atmosphere of K2-18b.

Chapter 6

Conclusion

In this final chapter, I summarize the fundamental aspects analyzed and the results obtained in this thesis to understand the atmosphere of K2-18b better. The main goal of this work was to characterize K2-18b's atmosphere and determine whether it is a primary atmosphere, mainly composed of hydrogen and helium, or if it has transitioned to a secondary atmosphere rich in heavier molecules.

First, I introduced a class of exoplanets called small planets, which can be divided into super-Earths and sub-Neptunes. I explained the difference between these two classes, highlighting the gap in radius distribution; super-Earths are bare cores, while sub-Neptunes still have an atmospheric envelope. I discussed possible explanations for this gap, focusing on atmospheric loss mechanisms like photoevaporation and core-powered mass loss, which could cause smaller planets to lose their primary atmospheres and develop secondary ones. I also explored hybrid atmospheres, which helped justify some of the results obtained in this study. An important aspect of small planets is habitability, influenced by factors such as a planet's position in its star's habitable zone, climate, and the type of star it orbits. I introduced the field of astrobiology, illustrating the conditions a planet needs to support life, especially on planets orbiting M-dwarf stars.

K2-18b, the focus of this thesis, is one of the most interesting M-dwarf habitable zone small exoplanets for atmospheric characterization. I reviewed K2-18b's general properties and existing studies on its atmosphere, concluding that it is a sub-Neptune with a hydrogen-rich atmosphere. However, whether its atmosphere is primary or secondary remains unclear from the literature. For this reason, the goal of my work was to characterize the atmosphere of K2-18b to determine whether it is a primary or secondary one.

To better constrain the atmosphere of a specific planet I first introduced the key atmospheric physics concepts, focusing on the radiative transfer process, on the various chemical models, on the temperature-pressure profiles and the presence of clouds.

To study the atmosphere of K2-18b I used the planet's transmission spectrum reduced by Madhusudhan et al. (2023) [63] from the transit data of the planet in front of its star obtained by the James Webb Space Telescope (JWST). I simulated K2-18b's atmosphere assuming different atmospheric models to better interpret the input spectrum. I used the TauREx3 framework [1], a line-by-line radiative transfer fully Bayesian retrieval framework. It consists of the forward model framework, which builds and computes a forward model, and the retrieval framework, which fits a forward model against the observations.

To test whether K2-18b has a primary atmosphere, I first simulated three constant molecular abundance models with hydrogen, helium, and molecules like H_2O, CH_4, NH_3, CO, and CO_2, considering (1) an isothermal temperature-pressure (T-P) profile, (2) a similar model with a cloud layer and (3) a 4-point T-P profile. I then simulated other two constant molecular abundance models adding nitrogen (N_2) at (1) and (3) to simulate a heavier secondary atmosphere. Finally, I tested an equilibrium chemical model to check for other molecules, considering a 4-point profile. In total, I ran six models that displayed significant amounts of methane (CH_4) and carbon dioxide (CO_2) in K2-18b's atmosphere, while H_2O, NH_3, and CO contributed less.

To find the best model, I used the comparison of the models' Bayesian evidence, obtained by the nested sampling MULTINEST [32]. In particular, I employed the Atmospheric Detectability Index (ADI) [87], a positively defined Bayes Factor obtained by subtracting the logarithmic Bayesian evidence of two models. ADI's value greater than 3 means a significant detection of an atmosphere, while ADIs below 3 are unable to favour one model over the others. Since the ADI values were all below 3, the statistical comparison was inconclusive. However, following Occam's razor, I identified the simplest model, *CA-iso*, as the best fit. This model describes K2-18b's atmosphere with constant molecular abundances and an isothermal temperature-pressure profile.

One way to determine if K2-18b's atmosphere is primarily, composed of hydrogen and helium, or if it has heavier molecules is possible by examining the mean molecular weight, μ, of the atmosphere. The *CA-iso* model presents a value μ of $9.13^{+3.39}_{-2.81}$ amu, which aligns with the result from Tsiaras et al. (2019) [88] within 1σ, indicating a mostly hydrogen and helium atmosphere with traces of other molecules dominated by CH_4 and CO_2. This value, higher than expected for a pure primary atmosphere and lower for a full secondary atmosphere, suggests that K2-18b may possess a hybrid atmosphere, combining elements of both primary and secondary atmospheres.

Despite the progress made with current JWST observations and atmospheric modelling, it remains challenging to definitively distinguish between primary and secondary atmospheric models for K2-18b. The data suggests that the planet may be in a transitional or hybrid state, but more detailed observations are necessary to fully understand the atmospheric dynamics and composition. Several upcoming JWST observations of K2-18b will be able to verify the atmospheric composition, in particular more observations with NIRSpec G395H (JWST GO 2372), which could confirm the results with higher precision, and MIRI LRS in the range 5-10 μm (JWST GO 2722), which specifically confirm the presence of biomarkers, such as the dimethyl sulfide which is expected to have a strong spectral feature around 7 μm [63].

Future missions like ARIEL [85] and Twinkle [25], particularly with wider wavelength coverage and higher precision, will have a main goal to reveal the profound nature and the way planets in the low mass regime, like K2-18b, form. They will be crucial to improve the knowledge of K2-18b's atmosphere. Additionally, the development of more sophisticated models to account for complex atmospheric processes, such as photochemistry and cloud formation, will be essential for refining the understanding of K2-18b and in general the small exoplanets.

Bibliography

[1] A. F. Al-Refaie, Q. Changeat, O. Venot, I. P. Waldmann, and G. Tinetti. A comparison of chemical models of exoplanet atmospheres enabled by taurex 3.1. *The Astrophysical Journal*, 932(2):123, jun 2022.

[2] Ahmed F Al-Refaie, Quentin Changeat, Ingo P Waldmann, and Giovanna Tinetti. Taurex 3: A fast, dynamic, and extendable framework for retrievals. *The Astrophysical Journal*, 917(1):37, 2021.

[3] David G. Andrews. *An Introduction to Atmospheric Physics*. Cambridge University Press, 2 edition, 2010.

[4] Kyle Barbary. nestle: Nested sampling algorithms for evaluating bayesian evidence. *Astrophysics Source Code Library*, pages ascl–2103, 2021.

[5] Natalie M. Batalha. Exploring exoplanet populations with nasa's kepler mission. *Proceedings of the National Academy of Sciences*, 111(35):12647–12654, 2014.

[6] Jacob L Bean, Sean N Raymond, and James E Owen. The nature and origins of sub-neptune size planets. *Journal of Geophysical Research: Planets*, 126(1):e2020JE006639, 2021.

[7] Charles Beichman, Bjoern Benneke, Heather Knutson, Roger Smith, Pierre-Olivier Lagage, Courtney Dressing, David Latham, Jonathan Lunine, Stephan Birkmann, et al. Observations of transiting exoplanets with the james webb space telescope (jwst). *Publications of the Astronomical Society of the Pacific*, 126(946):1134, dec 2014.

[8] Björn Benneke, Heather A Knutson, Joshua Lothringer, Ian JM Crossfield, Julianne I Moses, Caroline Morley, Laura Kreidberg, Benjamin J Fulton, Diana Dragomir, Andrew W Howard, et al. A sub-neptune exoplanet with a low-metallicity methane-depleted atmosphere and mie-scattering clouds. *Nature Astronomy*, 3(9):813–821, 2019.

[9] Björn Benneke and Sara Seager. How to distinguish between cloudy mini-neptunes and water/volatile-dominated super-earths. *The Astrophysical Journal*, 778(2):153, 2013.

[10] Björn Benneke, Michael Werner, Erik Petigura, Heather Knutson, Courtney Dressing, Ian JM Crossfield, Joshua E Schlieder, John Livingston, Charles Beichman, Jessie Christiansen, et al. Spitzer observations confirm and rescue the habitable-zone super-earth k2-18b for future characterization. *The Astrophysical Journal*, 834(2):187, 2017.

[11] Björn Benneke, Ian Wong, Caroline Piaulet, Heather A Knutson, Joshua Lothringer, Caroline V Morley, Ian JM Crossfield, Peter Gao, Thomas P Greene, Courtney Dressing, et al. Water vapor and clouds on the habitable-zone sub-neptune exoplanet k2-18b. *The Astrophysical Journal*, 887(1):L14, 2019.

[12] Bruno Bézard, Benjamin Charnay, and Doriann Blain. Methane as a dominant absorber in the habitable-zone sub-neptune k2-18 b. *Nature Astronomy*, 6(5):537–540, 2022.

[13] Doriann Blain, Benjamin Charnay, and Bruno Bézard. 1d atmospheric study of the temperate sub-neptune k2-18b. *Astronomy & Astrophysics*, 646:A15, 2021.

[14] Giordano Bruno. *On the Infinite Universe and Worlds*, volume 3. 1584.

[15] Howard Bushouse. The jwst science calibration pipeline. *Astronomical Data Analysis Software and Systems XXIX*, 527:583, 2020.

[16] Quentin Changeat, Billy Edwards, Ahmed F Al-Refaie, Angelos Tsiaras, Ingo P Waldmann, and Giovanna Tinetti. Disentangling atmospheric compositions of k2-18 b with next generation facilities. *Experimental Astronomy*, pages 1–26, 2021.

[17] David Charbonneau, Timothy M. Brown, Robert W. Noyes, and Ronald L. Gilliland. Detection of an extrasolar planet atmosphere*. *The Astrophysical Journal*, 568(1):377, mar 2002.

[18] R Cloutier, N Astudillo-Defru, R Doyon, X Bonfils, J-M Almenara, B Benneke, F Bouchy, X Delfosse, D Ehrenreich, T Forveille, et al. Characterization of the k2-18 multi-planetary system with harps-a habitable zone super-earth and discovery of a second, warm super-earth on a non-coplanar orbit. *Astronomy & Astrophysics*, 608:A35, 2017.

[19] R Cloutier, N Astudillo-Defru, R Doyon, Xavier Bonfils, J-M Almenara, F Bouchy, X Delfosse, T Forveille, C Lovis, M Mayor, et al. Confirmation of the radial velocity super-earth k2-18c with harps and carmenes. *Astronomy & Astrophysics*, 621:A49, 2019.

[20] Ziqi Dai, Dong Ni, Lizhuang Pan, and Yiheng Zhu. Five methods of exoplanet detection. *Journal of Physics: Conference Series*, 2012(1):012135, sep 2021.

[21] Gennaro D'Angelo and Jack J Lissauer. Formation of giant planets. *arXiv preprint arXiv:1806.05649*, 2018.

[22] David J Des Marais, Louis J Allamandola, Steven A Benner, Alan P Boss, David Deamer, Paul G Falkowski, Jack D Farmer, S Blair Hedges, Bruce M Jakosky, Andrew H Knoll, et al. The nasa astrobiology roadmap. *Astrobiology*, 3(2):219–235, 2003.

[23] Peter Deuflhard. *Newton methods for nonlinear problems: affine invariance and adaptive algorithms*, volume 35. Springer Science & Business Media, 2011.

[24] Courtney D Dressing and David Charbonneau. The occurrence rate of small planets around small stars. *The Astrophysical Journal*, 767(1):95, 2013.

[25] Billy Edwards, Malena Rice, Tiziano Zingales, Marcell Tessenyi, Ingo Waldmann, Giovanna Tinetti, Enzo Pascale, Giorgio Savini, and Subhajit Sarkar. Exoplanet

spectroscopy and photometry with the twinkle space telescope. *Experimental Astronomy*, 47:29–63, 2019.

[26] David Ehrenreich, Giovanna Tinetti, A Lecavelier Des Etangs, Alfred Vidal-Madjar, and Franck Selsis. The transmission spectrum of earth-size transiting planets. *Astronomy & Astrophysics*, 448(1):379–393, 2006.

[27] Ehrenreich, D., Tinetti, G., Lecavelier des Etangs, A., Vidal-Madjar, A., and Selsis, F. The transmission spectrum of earth-size transiting planets. *AA*, 448(1):379–393, 2006.

[28] Linda T Elkins-Tanton and Sara Seager. Ranges of atmospheric mass and composition of super-earth exoplanets. *The Astrophysical Journal*, 685(2):1237, 2008.

[29] Epicuro. *Letter to Herodotus*.

[30] Raissa Estrela, Mark R Swain, Akash Gupta, Christophe Sotin, and Adriana Valio. The evolutionary track of h/he envelopes of the observed population of sub-neptunes and super-earths. *The Astrophysical Journal*, 898(2):104, 2020.

[31] F. Feroz and M. P. Hobson. Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. *Monthly Notices of the Royal Astronomical Society*, 384(2):449–463, 01 2008.

[32] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, 398(4):1601–1614, 09 2009.

[33] Bradford J. Foley. The role of plate tectonic–climate coupling and exposed land area in the development of habitable climates on rocky planets. *The Astrophysical Journal*, 812(1):36, oct 2015.

[34] Daniel Foreman-Mackey, Benjamin T Montet, David W Hogg, Timothy D Morton, Dun Wang, and Bernhard Schölkopf. A systematic search for transiting planets in the k2 data. *The Astrophysical Journal*, 806(2):215, 2015.

[35] Naho Fujita, Yasunori Hori, and Takanori Sasaki. Orbital evolution of close-in superearths driven by atmospheric escape. *The Astrophysical Journal*, 928(2):105, 2022.

[36] Jonathan P Gardner, John C Mather, Randy Abbott, James S Abell, Mark Abernathy, Faith E Abney, John G Abraham, Roberto Abraham, Yasin M Abul-Huda, Scott Acton, et al. The james webb space telescope mission. *Publications of the Astronomical Society of the Pacific*, 135(1048):068001, 2023.

[37] Gebhard, Timothy D., Angerhausen, Daniel, Konrad, Björn S., Alei, Eleonora, Quanz, Sascha P., and Schölkopf, Bernhard. Parameterizing pressure-temperature profiles of exoplanet atmospheres with neural networks. *AA*, 681:A3, 2024.

[38] Iouli E Gordon, Laurence S Rothman, ea RJ Hargreaves, R Hashemi, Ekaterina Vladimirovna Karlovets, FM Skinner, Eamon K Conway, Christian Hill, Roman V Kochanov, Y Tan, et al. The hitran2020 molecular spectroscopic database. *Journal of quantitative spectroscopy and radiative transfer*, 277:107949, 2022.

[39] Edward F Guinan and Scott G Engle. The k2-18b planetary system: Estimates of the age and x-uv irradiances of a habitable zone "wet" sub-neptune planet. *Research Notes of the AAS*, 3(12):189, 2019.

[40] Jianpo Guo, Fenghui Zhang, Xuefei Chen, and Zhanwen Han. Probability distribution of terrestrial planets in habitable zones around host stars. *Astrophysics and Space Science*, 323, 03 2010.

[41] WJ Handley, MP Hobson, and AN Lasenby. Polychord: next-generation nested sampling. *Monthly Notices of the Royal Astronomical Society*, 453(4):4384–4398, 2015.

[42] Artie P Hatzes and Heike Rauer. A definition for giant planets based on the mass–density relationship. *The Astrophysical Journal Letters*, 810(2):L25, 2015.

[43] Gregory W. Henry, Geoffrey W. Marcy, R. Paul Butler, and Steven S. Vogt. A transiting "51 peg-like" planet*. *The Astrophysical Journal*, 529(1):L41, dec 1999.

[44] Edward Higson. dypolychord: dynamic nested sampling with polychord. 2018.

[45] Måns Holmberg and Nikku Madhusudhan. Exoplanet spectroscopy with jwst niriss: diagnostics and case studies. *Monthly Notices of the Royal Astronomical Society*, 524(1):377–402, 2023.

[46] T-O Husser, Sebastian Wende-von Berg, Stefan Dreizler, Derek Homeier, Ansgar Reiners, Travis Barman, and Peter H Hauschildt. A new extensive library of phoenix stellar atmospheres and synthetic spectra. *Astronomy & Astrophysics*, 553:A6, 2013.

[47] Harold Jeffreys. *The theory of probability*. OuP Oxford, 1998.

[48] Tijs Karman, Iouli E Gordon, Ad van Der Avoird, Yury I Baranov, Christian Boulet, Brian J Drouin, Gerrit C Groenenboom, Magnus Gustafsson, Jean-Michel Hartmann, Robert L Kurucz, et al. Update of the hitran collision-induced absorption section. *Icarus*, 328:160–175, 2019.

[49] Robert E Kass and Adrian E Raftery. Bayes factors. *Journal of the american statistical association*, 90(430):773–795, 1995.

[50] James F. Kasting, Ravikumar Kopparapu, Ramses M. Ramirez, and Chester E. Harman. Remote life-detection criteria, habitable zone boundaries, and the frequency of earth-like planets around m and late k stars. *Proceedings of the National Academy of Sciences*, 111(35):12641–12646, 2014.

[51] James F Kasting, Daniel P Whitmire, and Ray T Reynolds. Habitable zones around main sequence stars. *Icarus*, 101(1):108–128, 1993.

[52] Heather A. Knutson, David Charbonneau, Robert W. Noyes, Timothy M. Brown, and Ronald L. Gilliland. Using stellar limb-darkening to refine the properties of hd 209458b. *The Astrophysical Journal*, 655(1):564, jan 2007.

[53] Laura Kreidberg. Exoplanet atmosphere measurements from transmission spectroscopy and other planet-star combined light observations. *arXiv preprint arXiv:1709.05941*, 2017.

[54] Alain Léger, Franck Selsis, Christophe Sotin, Tristan Guillot, Didier Despois, D Mawet, Marc Ollivier, A Labèque, C Valette, F Brachet, et al. A new family of planets?"ocean-planets". *Icarus*, 169(2):499–504, 2004.

[55] K.N. Liou. *An Introduction to Atmospheric Radiation*. International Geophysics. Elsevier Science, 2002.

[56] Eric D Lopez and Jonathan J Fortney. Understanding the mass–radius relation for sub-neptunes: radius as a proxy for composition. *The Astrophysical Journal*, 792(1):1, 2014.

[57] Nikku Madhusudhan. Atmospheric retrieval of exoplanets. *arXiv preprint arXiv:1808.04824*, 2018.

[58] Nikku Madhusudhan. Exoplanetary atmospheres: Key insights, challenges, and prospects. *Annual Review of Astronomy and Astrophysics*, 57(1):617–663, August 2019.

[59] Nikku Madhusudhan, Marcelino Agúndez, Julianne I Moses, and Yongyun Hu. Exoplanetary atmospheres—chemistry, formation conditions, and habitability. *Space science reviews*, 205:285–348, 2016.

[60] Nikku Madhusudhan, Heather Knutson, Jonathan Fortney, and Travis Barman. Exoplanetary atmospheres. *arXiv preprint arXiv:1402.1169*, 2014.

[61] Nikku Madhusudhan, Matthew C Nixon, Luis Welbanks, Anjali AA Piette, and Richard A Booth. The interior and atmosphere of the habitable-zone exoplanet k2-18b. *The Astrophysical Journal*, 891(1):L7, 2020.

[62] Nikku Madhusudhan, Anjali AA Piette, and Savvas Constantinou. Habitability and biosignatures of hycean worlds. *The Astrophysical Journal*, 918(1):1, 2021.

[63] Nikku Madhusudhan, Subhajit Sarkar, Savvas Constantinou, Måns Holmberg, Anjali AA Piette, and Julianne I Moses. Carbon-bearing molecules in a possible hycean atmosphere. *The Astrophysical Journal Letters*, 956(1):L13, 2023.

[64] Mark S Marley, Andrew S Ackerman, Jeffrey N Cuzzi, and Daniel Kitzmann. Clouds and hazes in exoplanet atmospheres. *Comparative climatology of terrestrial planets*, 1:367–391, 2013.

[65] M. Mayor and D. Queloz. A jupiter-mass companion to a solar-type star. *Nature*, 378:55–359, 1995.

[66] Benjamin T Montet, Timothy D Morton, Daniel Foreman-Mackey, John Asher Johnson, David W Hogg, Brendan P Bowler, David W Latham, Allyson Bieryla, and Andrew W Mann. Stellar and planetary properties of k2 campaign 1 candidates and validation of 17 planets, including a planet receiving earth-like insolation. *The Astrophysical Journal*, 809(1):25, 2015.

[67] Pia Mukherjee, David Parkinson, and Andrew R Liddle. A nested sampling algorithm for cosmological model selection. *The Astrophysical Journal*, 638(2):L51, 2006.

[68] Léna Parc, François Bouchy, Julia Venturini, Caroline Dorn, and Ravit Helled. From super-earths to sub-neptunes: observational constraints and connections to theoretical models. *arXiv preprint arXiv:2406.04311*, 2024.

[69] Hannu Parviainen. Bayesian methods for exoplanet science. *arXiv preprint arXiv:1711.03329*, 2017.

[70] Michael Perryman. *The Exoplanet Handbook*. Cambridge University Press, 2 edition, 2018.

[71] Arazi Pinhas, Nikku Madhusudhan, Siddharth Gandhi, and Ryan MacDonald. H2o abundances and cloud properties in ten hot giant exoplanets. *Monthly Notices of the Royal Astronomical Society*, 482(2):1485–1498, 2019.

[72] Michael Radica, Étienne Artigau, David Lafreniére, Charles Cadieux, Neil J Cook, René Doyon, Pedro J Amado, José A Caballero, Thomas Henning, Andreas Quirrenbach, et al. Revisiting radial velocity measurements of the k2-18 system with the line-by-line framework. *Monthly Notices of the Royal Astronomical Society*, 517(4):5050–5062, 2022.

[73] H. Rauer, C. Catala, C. Aerts, T. Appourchaux, W. Benz, A. Brandeker, J. Christensen-Dalsgaard, M. Deleuil, L. Gizon, et al. The plato 2.0 mission. *Experimental Astronomy*, 38(1–2):249–330, September 2014.

[74] Cyril Richard, Iouli E Gordon, Laurence S Rothman, Martin Abel, Lothar Frommhold, Magnus Gustafsson, J-M Hartmann, Christian Hermans, Walter J Lafferty, Glenn S Orton, et al. New section of the hitran database: Collision-induced absorption (cia). *Journal of Quantitative Spectroscopy and Radiative Transfer*, 113(11):1276–1285, 2012.

[75] Edward W Schwieterman, Nancy Y Kiang, Mary N Parenteau, Chester E Harman, Shiladitya DasSarma, Theresa M Fisher, Giada N Arney, Hilairy E Hartnett, Christopher T Reinhard, Stephanie L Olson, et al. Exoplanet biosignatures: a review of remotely detectable signs of life. *Astrobiology*, 18(6):663–708, 2018.

[76] Sara Seager. Exoplanet atmospheres: A theoretical outlook. *Proceedings of the International Astronomical Union*, 6(S276):198–207, 2010.

[77] Aomawa L Shields, Sarah Ballard, and John Asher Johnson. The habitability of planets orbiting m-dwarf stars. *Physics Reports*, 663:1–38, 2016.

[78] Oliver Shorttle, Sean Jordan, Harrison Nicholls, Tim Lichtenberg, and Dan J Bower. Distinguishing oceans of water from magma on mini-neptune k2-18b. *The Astrophysical Journal Letters*, 962(1):L8, 2024.

[79] Kristin S Sotzen, Kevin B Stevenson, Erin M May, Natasha E Batalha, Noam R Izenberg, Sarah M Hörst, Calley L Tinsman, Carey M Lisse, Nikole K Lewis, Jayesh M Goyal, et al. On the utility of transmission color analysis i: Differentiating super-earths and sub-neptunes. *The Astronomical Journal*, 162(4):168, 2021.

[80] Joachim W Stock, Daniel Kitzmann, A Beate C Patzer, and Erwin Sedlmayr. Fastchem: A computer program for efficient complex chemical equilibrium calculations in the neutral/ionized gas phase with applications to stellar and planetary atmospheres. *Monthly Notices of the Royal Astronomical Society*, 479(1):865–874, 2018.

[81] Joachim W Stock, Daniel Kitzmann, A Beate C Patzer, and Erwin Sedlmayr. FastChem: A computer program for efficient complex chemical equilibrium calcu-

lations in the neutral/ionized gas phase with applications to stellar and planetary atmospheres . *Monthly Notices of the Royal Astronomical Society*, 479(1):865–874, 06 2018.

[82] Jonathan Tennyson, Sergei N Yurchenko, Ahmed F Al-Refaie, Emma J Barton, Katy L Chubb, Phillip A Coles, S Diamantopoulou, Maire N Gorman, Christian Hill, Aden Z Lam, et al. The exomol database: molecular line lists for exoplanet and other hot atmospheres. *Journal of Molecular Spectroscopy*, 327:73–94, 2016.

[83] Feng Tian. Thermal escape from super earth atmospheres in the habitable zones of m stars. *The Astrophysical Journal*, 703(1):905, 2009.

[84] Meng Tian and Kevin Heng. Atmospheric chemistry of secondary and hybrid atmospheres of super earths and sub-neptunes. *The Astrophysical Journal*, 963(2):157, mar 2024.

[85] Giovanna Tinetti, Paul Eccleston, Carole Haswell, Pierre-Olivier Lagage, Jérémy Leconte, Theresa Lüftinger, Giusi Micela, Michel Min, Göran Pilbratt, Ludovic Puig, et al. Ariel: Enabling planetary science across light-years. *arXiv preprint arXiv:2104.04824*, 2021.

[86] Roberto Trotta. Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, 49(2):71–104, 2008.

[87] A. Tsiaras, I. P. Waldmann, T. Zingales, M. Rocchetto, G. Morello, M. Damiano, K. Karpouzas, G. Tinetti, L. K. McKemmish, J. Tennyson, and S. N. Yurchenko. A population study of gaseous exoplanets. *The Astronomical Journal*, 155(4):156, mar 2018.

[88] Angelos Tsiaras, Ingo P Waldmann, Giovanna Tinetti, Jonathan Tennyson, and Sergey N Yurchenko. Water vapour in the atmosphere of the habitable-zone eight-earth-mass planet k2-18 b. *Nature Astronomy*, 3(12):1086–1091, 2019.

[89] Ingo P Waldmann, Marco Rocchetto, Giovanna Tinetti, Emma J Barton, Sergey N Yurchenko, and Jonathan Tennyson. Tau-rex. ii. retrieval of emission spectra. *The Astrophysical Journal*, 813(1):13, 2015.

[90] Ingo P Waldmann, Giovanna Tinetti, Marco Rocchetto, Emma J Barton, Sergey N Yurchenko, and Jonathan Tennyson. Tau-rex i: A next generation retrieval code for exoplanetary atmospheres. *The Astrophysical Journal*, 802(2):107, 2015.

[91] Nicholas F Wogan, Natasha E Batalha, Kevin J Zahnle, Joshua Krissansen-Totton, Shang-Min Tsai, and Renyu Hu. Jwst observations of k2-18b can be explained by a gas-rich mini-neptune with no habitable surface. *The Astrophysical Journal Letters*, 963(1):L7, 2024.

[92] A. Wolszczan and D. Frail. A planetary system around the millisecond pulsar psr1257+12. *Nature*, 355:145–147, 1992.