

# Toward a multidimensional analysis of transmission spectroscopy

## II. Day-night-induced biases in retrievals from hot to ultrahot Jupiters

William Pluriel<sup>1,4</sup>, Jérémy Leconte<sup>1</sup>, Vivien Parmentier<sup>2</sup>, Tiziano Zingales<sup>1,3</sup>, Aurélien Falco<sup>1</sup>,  
Franck Selsis<sup>1</sup>, and Pascal Bordé<sup>1</sup>

<sup>1</sup> Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France

<sup>2</sup> Department of Physics, Oxford University, OX1 2JD, UK  
e-mail: [vivien.parmentier@physics.ox.ac.uk](mailto:vivien.parmentier@physics.ox.ac.uk)

<sup>3</sup> Università di Padova, Dipartimento di Astronomia, vicolo dell'Osservatorio 3, 35122 Padova, Italy

<sup>4</sup> Observatoire astronomique de l'Université de Genève, chemin Pegasi 51, 1290 Versoix, Switzerland

Received 3 August 2021 / Accepted 14 October 2021

### ABSTRACT

Hot Jupiters (HJs) are very good targets for transmission spectroscopy analysis. Their atmospheres have a large scale height, implying a high signal-to-noise ratio. As these planets orbit close to their stars, they often present strong thermal and chemical heterogeneities between the day- and nightside of their atmosphere. For the hottest of these planets, the thermal dissociation of several species occurs in their atmospheres, which leads to a stronger chemical dichotomy between the two hemispheres. It has already been shown that the current retrieval algorithms, which are using 1D forward models, find biased molecular abundances in ultrahot Jupiters. Here, we quantify the effective temperature domain over which these biases are present. We used a set of 12 simulations of typical HJs from  $T_{\text{eq}} = 1000$  K to  $T_{\text{eq}} = 2100$  K performed with the substellar and planetary atmospheric radiation and circulation global climate model and generate transmission spectra that fully account for the 3D structure of the atmosphere with Pytmosph3R. These spectra were then analyzed using the 1D TauREx retrieval code. We find that for *James Webb* Space Telescope like data, accounting for nonisothermal vertical temperature profiles is required over the whole temperature range. We further find that 1D retrieval codes start to estimate incorrect parameter values for planets with equilibrium temperatures greater than 1400 K if there are absorbers in the visible (such as TiO and VO, e.g.) that are able to create a hot stratosphere. The high temperatures at low pressures indeed entail a thermal dissociation of species that creates a strong chemical day-night dichotomy. As a byproduct, we demonstrate that when synthetic observations are used to assess the detectability of a given feature or process using a Bayesian framework (e.g., by comparing the Bayesian evidence of retrievals with different model assumptions), it is valid to use nonrandomized input data as long as the anticipated observational uncertainties are correctly taken into account.

**Key words.** planets and satellites: atmospheres – radiative transfer – methods: numerical – techniques: spectroscopic

### 1. Introduction

Transmission spectroscopy enables detecting molecules in exoplanetary atmospheres, measuring molecular abundances, or setting upper limits on them. In this regard, several studies have pointed out that the 3D nature of real atmospheres for thermal and compositional effects often needs to be taken into account in the modeling for the inference process to yield accurate estimates (Caldas et al. 2019; Changeat et al. 2019; MacDonald et al. 2020; Lacy & Burrows 2020). Because the atmospheric region probed by transmission spectroscopy is not as thin as often assumed, measured spectra may be affected by variations through and along the terminator (day-to-night and pole-to-equator gradients).

For ultrahot Jupiters (UHJs), Pluriel et al. (2020b) demonstrated that  $[\text{CO}]/[\text{H}_2\text{O}]$  estimated with 1D spherically symmetric models can be different by several orders of magnitude because of strong day-to-night heterogeneities:  $\text{H}_2\text{O}$  is thermally dissociated on the hot dayside, while it is not on the cool nightside. Conversely, the CO abundance remains constant everywhere because the temperature is not high enough in any

location for CO to be thermally dissociated. For these planets, transmission spectra are the combined result of hot regions in the CO bands and cold regions in the  $\text{H}_2\text{O}$  bands. To capture this complexity, it is necessary to resort to models with more than one dimension. This is especially important in the context of the *James Webb* Space Telescope (JWST) and Ariel (ESA), which aims at measuring atmospheric elemental abundances with high accuracies (Tinetti et al. 2021).

HJs and UHJs have been studied with global climate models (GCMs) (Showman & Guillot 2002; Showman et al. 2008b; Menou & Rauscher 2009; Wordsworth et al. 2011; Heng et al. 2011; Charnay et al. 2015; Kataria et al. 2016; Drummond et al. 2016; Tan & Komacek 2019). GCMs are designed to describe planetary atmospheres with their full three-dimensional structures and dynamics, hence leading to deeper insights into their rich physics and chemistry (Showman et al. 2008b; Leconte et al. 2013; Guerlet et al. 2014; Venot et al. 2014; Parmentier et al. 2018). We know from GCMs that the cooler the planet, the weaker its atmospheric day-night dichotomy with corresponding 3D biases. Somewhere between UHJs and cold planets lies a boundary. It is the aim of this work to assess the equilibrium tem-

perature regions where 1D spherically symmetric models can be safely used and where they should not be used for the purpose of estimating physical and chemical parameters. To investigate this limit, we have designed and carried out a numerical experiment to identify the origin of the biases and to quantify them.

In the following, we produce synthetic data with a realistic 3D forward model chaining the substellar and planetary atmospheric radiation and circulation global climate model (SPARC/MIT GCM) developed by Showman et al. (2008b, 2013) with our 3D transmission spectrum computation module, Pytmosph3R (Caldas et al. 2019; Falco et al. 2022). Then, we use TauREx (Waldmann et al. 2015) to solve the inverse problem with a 1D forward model and estimate the parameters of interest. We refer to this last part as the retrieval process or retrieval for short because we investigate whether correct parameter values are retrieved from the synthetic data. In this last part, we adopt the point of view of an observer.

In Sect. 2 we explain how we built our numerical models and describe the parameterization we used. In Sect. 2.4 we discuss the use of nonrandomized input data and their validity for comparing various retrieval hypotheses. In Sect. 3 we describe the GCM outputs and the transmission spectra generated with Pytmosph3R, which are used as input for the retrieval analysis with TauREx (see Sect. 4). Finally, we discuss in Sect. 5 how 3D effects (vertical, across and along the limb) alter transmission spectra and vary as a function of equilibrium temperature and the nature of optical absorbers.

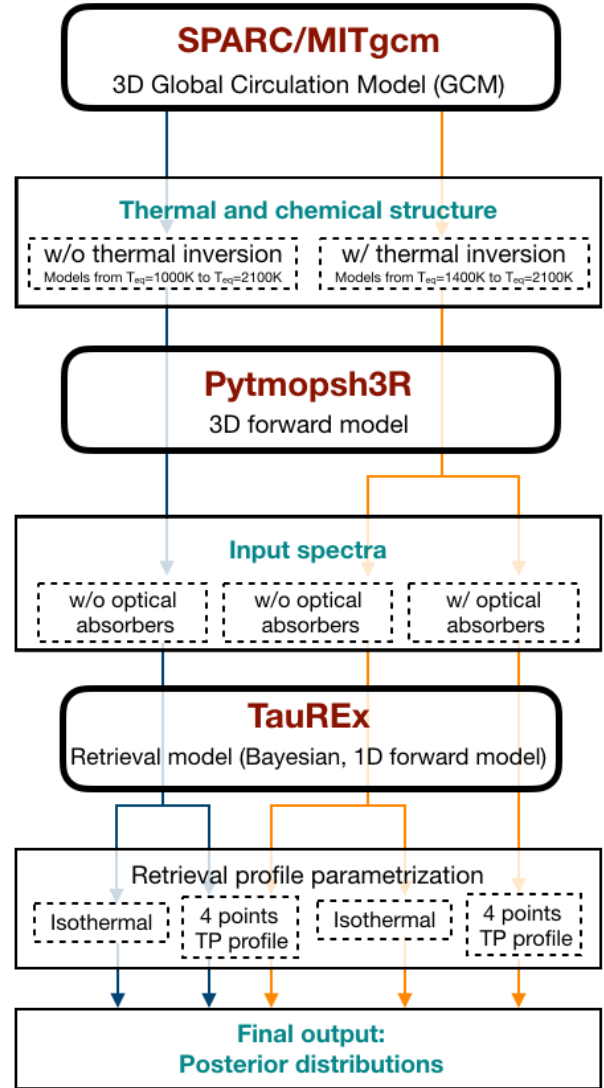
## 2. Numerical experiment

### 2.1. GCMs

To implement our numerical experiment (Fig. 1), we followed the method described by Pluriel et al. (2020b). In a first step, we ran SPARC/MIT simulations of HJs with equilibrium temperatures between 1000 and 2100 K as in Parmentier et al. (2021). We decided to truncate at 2100 K because we already showed in Pluriel et al. (2020b) the limitation of 1D retrieval model on a 2350 K equilibrium temperature planet (WASP-121 b). As we already demonstrated the limitation for hottest planet, we are interested here in colder planets. In each simulation, the pressure ranged from  $2.0 \cdot 10^7$  to 0.2 Pa over 53 levels. We used a horizontal resolution of C32, meaning that each of the six cube faces has a resolution of  $32 \times 32$  finite volume elements (Showman et al. 2009), equivalent to an approximated resolution of 128 cells in longitude and 64 in latitude. Radiative transfer was handled with a two-stream radiation scheme (Marley & McKay 1999). Opacities were treated using eight correlated-k coefficients (Goody & Yung 1989) within each of 11 wavelength bins (Kataria et al. 2013), assuming chemical equilibrium. To facilitate comparisons, we used the following planetary and stellar parameters in all simulations:

- $T_* = T_S = 5778$  K,
- $R_* = R_S = 6.957 \times 10^8$  m,
- $R_p = R_J = 7.1492 \times 10^7$  m,
- surface gravity:  $g = 10$  m s<sup>-2</sup>,
- $\frac{H_e}{H_2} = 0.25774$ ,
- solar metallicity,
- chemical equilibrium.

Thus, from one simulation to the next, only the chemical composition and the equilibrium temperature varied. This approach makes it possible to isolate different effects and facilitates interpretation. We thus hypothesize that the parameters chosen fit for different planet types from HJs to UHJs, which is consistent with



**Fig. 1.** Flowchart of the numerical experiment. The three models are shown in red, the outputs are indicated in green-blue, and the parameters are presented in black. There are two main sets of simulations from the GCM SPARC/MIT, with and without thermal inversion. This is indicated with the orange and the blue arrows, respectively. The final output consists of five sets of posterior distributions.

the current observations of this range of exoplanets. Moreover, we assumed a G-type star emitting like a blackbody. A different spectral type of the host star would affect the transmission spectra even if the G-star blackbody hypothesis is a fair assumption. For instance, we do not observe HJs and UHJs around M dwarfs, which have indeed a more complex spectral type with strong lines in the UV.

The strong atmospheric dichotomy in UHJs is not only due to extreme irradiation, but also to absorbers in the near-UV and optical domains, such as TiO and VO (Fortney et al. 2008; Parmentier et al. 2015, 2018), causing both a strong heating and a stratospheric thermal inversion. Although these molecules can be depleted by cold traps (Spiegel et al. 2009) in cool planets, this will not occur on the daysides of UHJs. Therefore thermal inversion is very likely in HJ atmospheres. It is possible, however, that TiO and VO rain out. Thus, the upper atmosphere will be depleted in TiO and VO according to the equilibrium (Parmentier et al. 2016). To consider this possibility, we chose to simulate planets with and without TiO and VO for this study.

Thus, we ran two sets of simulations: 12 simulations without thermal inversion without TiO and VO (called “no thermal inversion”), and with equilibrium temperatures ranging from 1000 to 2100 K in steps of 100 K, and 8 “thermal inversion” simulations with TiO and VO, and with equilibrium temperatures ranging from 1400 (minimum value ensuring gaseous forms for TiO and VO) to 2100 K in steps of 100 K. Although we did not include these species, we note that metals such as Fe or Mg and ionized hydrogen might also contribute significantly to thermal inversion in UHJs, as pointed out by [Lothringer et al. \(2018\)](#).

## 2.2. Transmission spectra

In a second step, we took for each simulated planet the temperature maps and chemical abundance tables produced by our GCM to generate 3D transmission spectra with the new version of Pytmosph3R ([Falco et al. 2022](#)) and using monochromatic cross-sections calculated by ExoMol ([Yurchenko et al. 2011, 2014; Tennyson & Yurchenko 2012; Barton et al. 2013, 2014](#)). The spectra were computed according to the two envisioned atmospheric compositions: (i) an atmosphere mainly composed of H<sub>2</sub> and He containing only H<sub>2</sub>O and CO as active gases. To account for thermal dissociation of H<sub>2</sub>O and H<sub>2</sub>, we used the following equation from [Parmentier et al. \(2018\)](#),

$$\frac{1}{A} = \left( \frac{1}{A_0^{0.5}} + \frac{1}{A_d^{0.5}} \right)^2, \quad (1)$$

where  $A_0$  is the deep abundance that is not affected by dissociation and  $A_d$  is the abundance in the region dominated by dissociation.

In the second composition (ii), we used same atmosphere as in (i), but added TiO and VO. Condensation of these species was accounted for.

In this work, we assumed that observations were carried out with the JWST, and we assumed that measurement noise is dominated by quantum detection noise following a Poisson statistics with equal mean and variance between  $\sim 2$  and  $10 \mu\text{m}$ . We estimated the mean number of collected photoelectrons to be

$$N_{\text{ph}} = \frac{\pi\tau\Delta t}{hc} \left( \frac{R_\star D}{2d} \right)^2 \int_{\lambda_1}^{\lambda_2} B(\lambda, T_\star) \lambda d\lambda, \quad (2)$$

where  $\lambda_1$  and  $\lambda_2$  are the limiting wavelengths of the spectral bin,  $d$  is the distance of the star (we took here 270 pc as if it were WASP-121), and  $R_\star$  and  $T_\star$  are the stellar radius and temperature, respectively. The parameters  $D$ ,  $\tau$ , and  $\Delta t$  are the telescope diameter, the system throughput, and the integration time, respectively, whose values were fixed for the JWST according to [Cowan et al. \(2015\)](#). Because systematics may prevent us from reaching a 10 ppm precision with the JWST, we assumed a floor noise of 30 ppm throughout the whole spectral domain with a normal distribution ([Greene et al. 2016](#)) from 0.6 to  $\sim 2 \mu\text{m}$ , wherever quantum noise was lower than 30 ppm. The noise depends on the wavelength above 2 microns according to Eq. (2). It reaches about 100 ppm at 10 microns. We simulated JWST observations from 0.6 to  $10 \mu\text{m}$  using the low-resolution prism mode provided by the Near-Infrared Spectrograph (NIR-Spec) and the Mid Infra-Red Instrument (MIRI) ([Stevenson et al. 2016](#)). We are aware that this observational setup is somewhat ideal, but we seek here to uncover biases due to atmospheric modeling, not due to instrumental effects. We specify that we used the standard deviation of the quantum noise as the estimate

**Table 1.** Retrieval parameters.

Parameters	W/ thermal inversion	W/O thermal inversion
[H <sub>2</sub> O]	10 <sup>-12</sup> –10 <sup>-1</sup>	10 <sup>-12</sup> –10 <sup>-1</sup>
[CO]	10 <sup>-12</sup> –10 <sup>-1</sup>	10 <sup>-12</sup> –10 <sup>-1</sup>
[TiO]	10 <sup>-12</sup> –10 <sup>-1</sup>	–
[VO]	10 <sup>-12</sup> –10 <sup>-1</sup>	–
$R_p$ ( $R_{\text{jup}}$ )	$\pm 50\%$	$\pm 50\%$
$P_{\text{clouds}}$ [Pa]	10 <sup>-2</sup> –10 <sup>6</sup>	10 <sup>-2</sup> –10 <sup>6</sup>
Isothermal profile		
$T_p$ [K]	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$
4-points TP profile		
$T^{\text{bot}}$ [K]	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$
$T_1$ [K]	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$
$P_1$ [Pa]	10 <sup>6</sup> –510 <sup>2</sup>	10 <sup>6</sup> –10 <sup>2</sup>
$T_2$ [K]	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$
$P_2$ [Pa]	10 <sup>6</sup> –10 <sup>0</sup>	10 <sup>6</sup> –10 <sup>0</sup>
$T_{\text{top}}$ [K]	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$	$0.3 \times T_{\text{eq}} - 2 \times T_{\text{eq}}$

**Notes.** List of retrieval parameters with corresponding fitting ranges. Priors are assumed to be flat in log space for abundances, in linear space otherwise.  $T_1$  and  $T_2$  are the temperature points corresponding to pressure points  $P_1$  and  $P_2$ , respectively.

of the uncertainty affecting the spectrum, but we did not add random noise to the spectra we computed. In Sect. 2.4 we present a study that confirms that using nonrandomized spectra does not affect our conclusions.

## 2.3. Parameter retrieval

In a third step, we used TauREx to retrieve the atmospheric parameters listed in Table 1, assuming uninformative flat priors with broad ranges, exactly as an observer would do. We ran the retrieval process twice, first with an isothermal profile, then with a four-point (six-parameter) thermal profile. We did not include clouds in the first or second step of the numerical experiment. It would be interesting to add clouds in the simulations, however. As the clouds would mostly affect the short wavelengths, the biases in the CO/H<sub>2</sub>O ratio we observe would probably still occur because the CO bands are at a longer wavelength. The presence of clouds might create degeneracies in the retrieval, however. Clouds are nevertheless part of the retrieval parameters in the TauREx model. This can break up possible degeneracies, and we verified that the model works correctly by not retrieving a cloud layer when we knew that none was implemented. When performing retrievals, we imposed a limiting condition to maintain physical scenarios: we set the cloud pressure range to be between the bottom and the top of the atmosphere. Finally, we compared the parameter inferred values with their actual values.

## 2.4. Nonrandomized spectra

As mentioned in Sect. 2.2, we computed nonrandomized spectra, that is, spectra to which no random noise was added. The reason is that a particular instance of added random noise would randomly affect the results of our retrieval studies in a particular way, whereas we seek to uniquely identify biases of 1D model-based retrievals. An approach could have been to perform a series of retrievals on a series of instances of added random

noise. However, as shown by Feng et al. (2018), this would have been unnecessarily computationally expensive as the posterior probability distributions of all these combined retrievals would have converged to the posterior distributions obtained from the nonrandomized spectrum.

In Sect. 2.3 we take the point of view of an observer. As observers would usually do, we therefore introduce the reduced chi-square, denoted by  $\tilde{\chi}^2$ , a common statistical goodness-of-fit metric that is defined as follows:

$$\tilde{\chi}^2 = \frac{1}{N-p} \sum_{i=1}^N \left[ \frac{O_i - C_i}{\sigma_i} \right]^2, \quad (3)$$

where  $O$  and  $C$  are the observed and calculated spectra, respectively,  $\sigma$  are the uncertainties,  $p$  is the number of parameters, and  $N$  is the total number of measurements. In the case of nonrandomized spectra with uncertainty  $\sigma_i$ , but without added white Gaussian noise with standard deviation  $\sigma_i$ ,  $O_i - C_i$  can very well be much smaller than  $\sigma_i$ . Hence,  $\tilde{\chi}^2$  will not have a statistical mean value of 1. In our experiment,  $\tilde{\chi}^2 \ll 1$  is therefore to be expected and should not be seen as the sign of noise fitting.

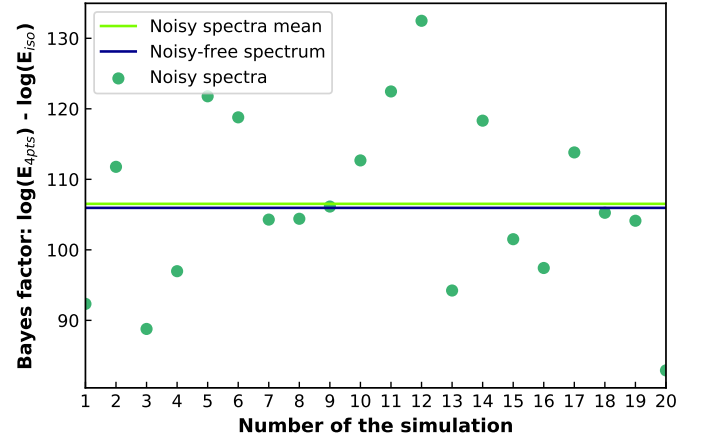
Furthermore, we wished to confirm that we can use logarithmic Bayes factors to compare different forward models (assumptions) using nonrandomized input spectra. For this purpose, we used Pytmosph3R to generate a transmission spectrum of a simple homogeneous atmosphere with a vertical temperature gradient. This nonrandomized spectrum serves as a reference. Then, from this spectrum, we generated 20 noisy spectra by adding to each data point a random value drawn from a normal distribution with a 30 ppm standard deviation. We then performed retrievals with TauREx on each of the 21 spectra using two different assumptions on the temperature-pressure (TP) profile: (i) an isothermal atmosphere, and (ii) a four-point TP profile atmosphere, as described in Table 1. For all retrievals, TauREx was provided with the same  $1\sigma$  uncertainty (in this simple case, 30 ppm), whether the input spectrum was noisy or not.

We used Bayesian evidence as defined by Waldmann et al. (2015) to compute the logarithmic Bayes factor,

$$\log B = \log \frac{E_{4\text{pts}}}{E_{\text{iso}}} = \log E_{4\text{pts}} - \log E_{\text{iso}}, \quad (4)$$

where  $E_{4\text{pts}}$  and  $E_{\text{iso}}$  is the evidence of the four-point TP profile and the isothermal profile, respectively. Figure 2 shows the result, which strongly favors the four-point TP profile. Although  $\log B$  is widely spread for the randomized spectra, its average value is very close to the nonrandomized reference value. We conclude that it is perfectly valid to carry out comparisons of various retrieval hypotheses based on nonrandomized data as long as uncertainties are correctly accounted for. This approach even alleviates potential biases in model selection due to a particular instance of noise. This works because evidence computation does not rely on the  $\tilde{\chi}^2$  value of the best-fit model, but integrates information over the whole parameter space while always accounting for uncertainty on the data points. Consequently, in the remainder of this article, model selection will always be made by computing logarithmic Bayes factors and never by comparing  $\tilde{\chi}^2$  values. However, we still keep and indicate  $\tilde{\chi}^2$  values as potential warnings for unacceptable fits pleading for individual model rejection ( $\tilde{\chi}^2 \gg 1$ ).

While a spread of  $\log B$  is to be expected when using randomized data, the spread of about 40 visible in Fig. 2 appears relatively large compared to the typical value of 5 that is



**Fig. 2.** Logarithmic Bayes factor for a four-point TP profile vs. an isothermal profile. The figure shows in blue the value obtained with nonrandomized spectra, and green dots represent the 20 values for normally distributed noisy spectra (30 ppm standard deviation). The logarithmic Bayes factor averaged on the 20 simulations is shown in light green. The nonrandomized logarithmic Bayes factor and its noisy average counterpart are almost identical, proving that our study can be based on nonrandomized spectra.

generally used for the model selection. This probably means that selection thresholds for the types of retrievals we are dealing with will have to be properly calibrated in the future.

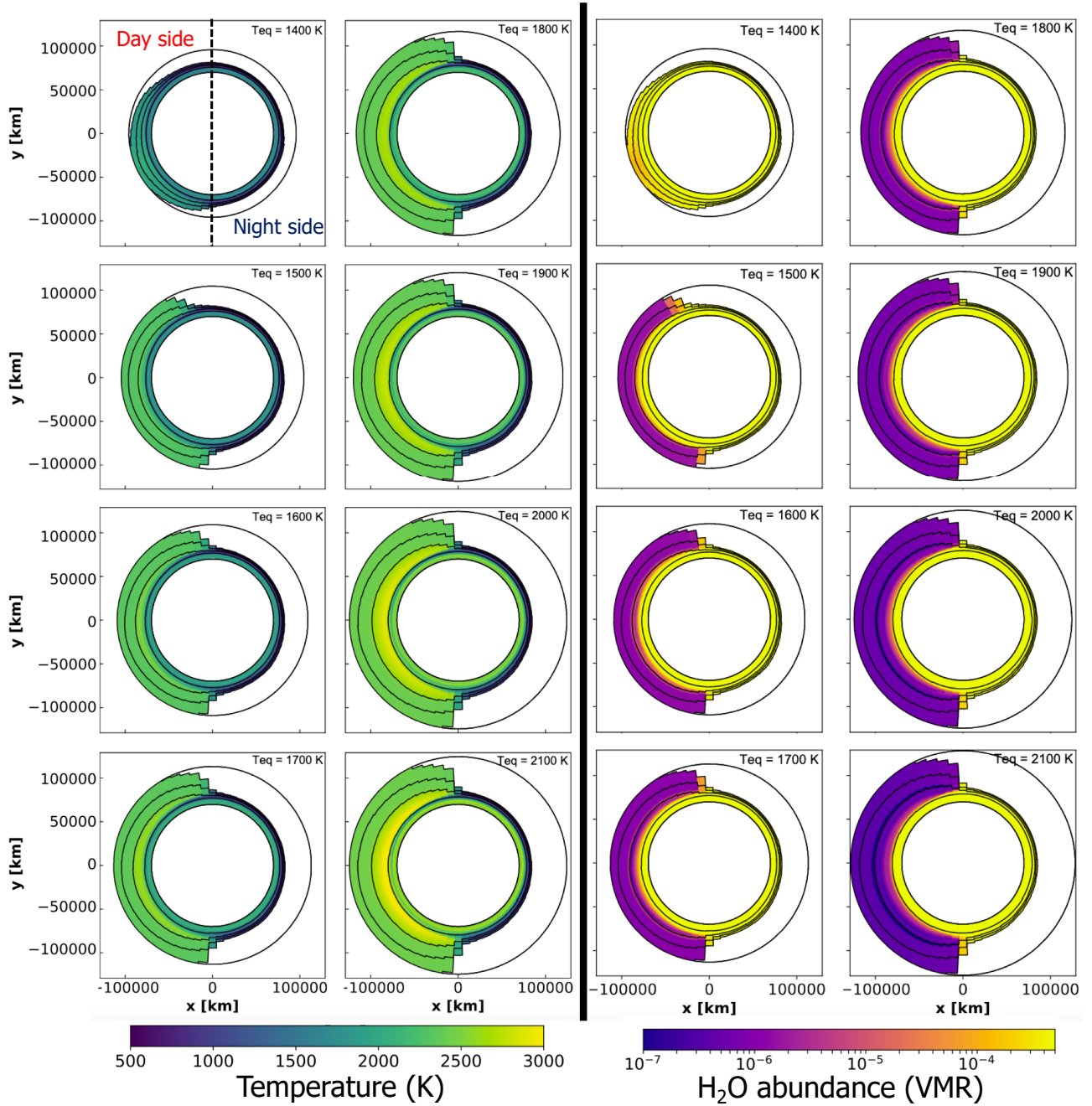
### 3. GCM simulations and transmission spectra

#### 3.1. GCM simulations with optical absorbers

In Fig. 3 we show the temperature and water abundance equatorial maps obtained from GCM simulations with optical absorbers TiO and VO. As in the GCM simulation of WASP-121 b by Pluriel et al. (2020b), we can distinguish three regions in the atmosphere: (i) a quasi-isothermal layer (with only small variations with latitude and longitude) corresponding to the deepest layers of the simulations; above this deep layer, (ii) an overall cold region with a temperature that decreases with increasing altitude on the nightside, and (iii) on the dayside, a hot stratosphere with the highest temperatures in the simulation.

The temperature of the quasi-isothermal layer increases roughly linearly with the equilibrium temperature of the planet, starting from about 1300 K ( $T_{\text{eq}} = 1400$  K) to about 2200 K ( $T_{\text{eq}} = 2100$  K). This layer is compressed to higher pressures as  $T_{\text{eq}}$  increases. It extends to the altitude of the  $4 \times 10^3$  Pa level for  $T_{\text{eq}} = 1400$  K, but not above the altitude of the  $2 \times 10^4$  Pa level for  $T_{\text{eq}} = 2100$  K.

On the dayside, the variation in temperature with latitude and altitude is quite complex. We observe that the location of the hottest region of the atmosphere depends on the equilibrium temperature of the simulation. While it is aligned with the substellar point for the hottest simulations, it is shifted eastward for cooler simulations, up to  $23^\circ$  for  $T_{\text{eq}} = 1400$  K. This well-known shift (Knutson et al. 2007; Showman et al. 2008b) occurs when the energy advection timescale becomes smaller than the radiative timescale. In this case, the hottest point is controlled by both circulation and radiation and is displaced to the east by zonal winds. The more intense and organized the wind dynamics (jets, super-rotation), the greater this shift between the substellar point and the hottest region. However, for the most highly irradiated planets, the radiative timescale becomes shorter than the

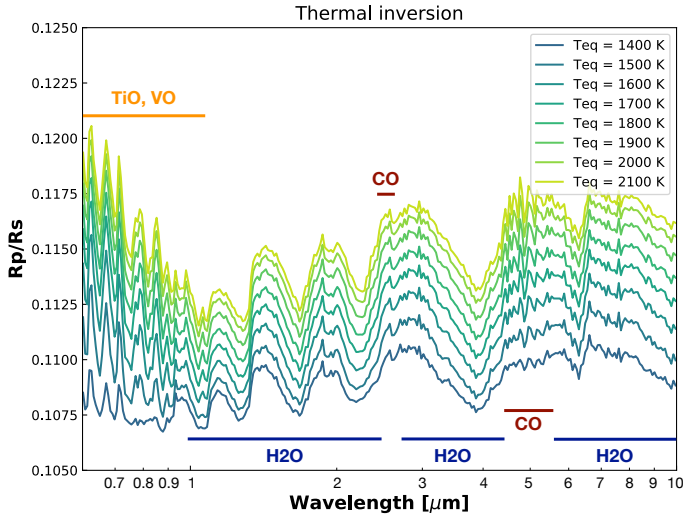


**Fig. 3.** Equatorial cut of the temperature and water abundance for the 8 GCM simulations with optical absorbers (TiO and VO). The equilibrium temperature of the planet (ranging from 1400 to 2100 K) is specified. From the center outward, the five solid black lines are the  $1.434 \times 10^7$ ,  $10^3$ , 1,  $10^{-2}$ , and  $10^{-4}$  Pa pressure levels. The hotter the equilibrium temperature, the larger the day-night thermal and chemical dichotomy. Water is completely dissociated below  $10^3$  Pa on the dayside for the hottest simulations.

dynamics timescale, hence an alignment between the substellar point and the hottest region. Thus, there is an asymmetry between east and west of the atmosphere, which extends well beyond the terminator for the colder simulations. The strong dichotomy in day-night temperature of the hottest atmospheres induces wind dynamics that are sufficiently effective to affect the terminator of these planets. Differences in temperatures and therefore in scale heights are clearly visible in the terminator of these atmospheres. The different east and west terminator signatures will thus mix into the global transmission spectra. Nevertheless, these asymmetries disappear very quickly by moving away from the terminator for the hottest planets.

The overall change in temperature on the dayside increases sharply as the equilibrium temperature increases, implying a significant increase in the scale height of this hemisphere, much greater than that of the night hemisphere. This stratospheric heating by optical absorbers significantly enhances the day-night asymmetry as the planetary equilibrium temperature raises. The temperature on the nightside of each simulation in fact remains very cold overall, and although it increases slightly with the equilibrium temperature, its scale height remains very small compared to that of the dayside (Parmentier et al. 2021).

We also show in Fig. 3 volume-mixing ratio (VMR) maps of  $\text{H}_2\text{O}$ . In our hottest simulations, water thermal dissociation is



**Fig. 4.** Transmission spectra calculated with Pytmosph3R from the eight GCM simulations with optical absorbers (VO and TiO). We added the main spectral bands of CO, TiO and VO, and H<sub>2</sub>O. CO bands become deeper when the temperature increases, while the TiO and VO bands become shallower when the temperature decreases because these species condense.

extremely efficient on the dayside of the planet, gradually disappearing with decreasing  $T_{\text{eq}}$  and becoming finally negligible at  $T_{\text{eq}} = 1400$  K. As the considered equilibrium temperature decreases, we note that the dissociation of H<sub>2</sub>O, which depends on both temperature and pressure, affects regions of lower pressures. The abundance of H<sub>2</sub>O at the terminator remains close to solar abundance for all simulations. Only the hottest simulations reveal a slight decrease in the abundance of H<sub>2</sub>O for the regions with the lowest pressures, even though it remains above  $10^{-4}$  in VMR.

Figure 4 shows the spectra generated with Pytmosph3R from our thermal inversion simulations. The transmission spectra are all shifted to higher  $R_p/R_\star$  ratios for increasingly hotter planets. This vertical shift indicates that higher altitudes are probed, implying more extended atmospheres with larger scale heights. The scale height rises because of the temperature increase and molecular weight decrease that are caused by the thermal dissociation of H<sub>2</sub>. If water were not thermally dissociated,  $R_p/R_\star$  would be higher and would hide most of the CO features. However, as H<sub>2</sub>O dissociates on the dayside, the water features come from the nightside while the CO features come from the dayside of the atmosphere, thus they appear more clearly in the spectra. We also see that adding TiO and VO in the atmosphere adds features in the optical domain, hiding the Rayleigh slope and several water bands in the near-IR. The amplitude of the H<sub>2</sub>O absorption bands increases with equilibrium temperature. As water absorption increases with temperature (Yurchenko et al. 2011), this indicates that higher temperatures are probed. Moreover, the CO absorption bands (at 2.3 and 4.5 μm), which remain quite weak for the coldest simulations, appear more clearly for the hottest simulations. Because CO remains present everywhere in the atmosphere without being dissociated, the absorption bands are indeed dominated by hot dayside features, and this trend increases with the equilibrium temperature as the day-night gradient. This means that when we simulate hotter planets, the greater the equilibrium temperature, the greater the difference between the temperature probed by the absorption bands of H<sub>2</sub>O and those of CO. Figure 4 also shows that the condensation

of TiO and VO occurs in the coldest simulation, implying a decrease in the features in the optical part of the spectra. However, for the hottest simulation at  $T_{\text{eq}} = 2100$  K, the spectra are unaffected by condensation.

### 3.2. GCM simulations without optical absorbers

Equatorial cut temperature maps for GCM simulations without TiO and VO in the atmosphere are presented in Fig. 5. These maps strongly differ from those obtained with TiO and VO. The atmospheres are less extended, especially on the dayside, due to the absence of the visible and near-UV absorbers. Thus these atmospheres do not own stratospheres. In the absence of stratospheric heating, dayside scale heights are significantly smaller. The atmospheres are also much more homogeneous horizontally without a significant day-night dichotomy for  $T_{\text{eq}}$  below  $\sim 1400$  K. A pronounced day-night thermal gradient gradually appears for hotter cases with an increasing eastward shift of the hottest point with respect to the substellar point, which reaches  $33^\circ$  for the hottest simulation (see Table 2). The lower temperatures obtained without TiO and VO heating yield longer radiative timescale than the cases with thermal inversion. This explains these larger displacements of the hot spot.

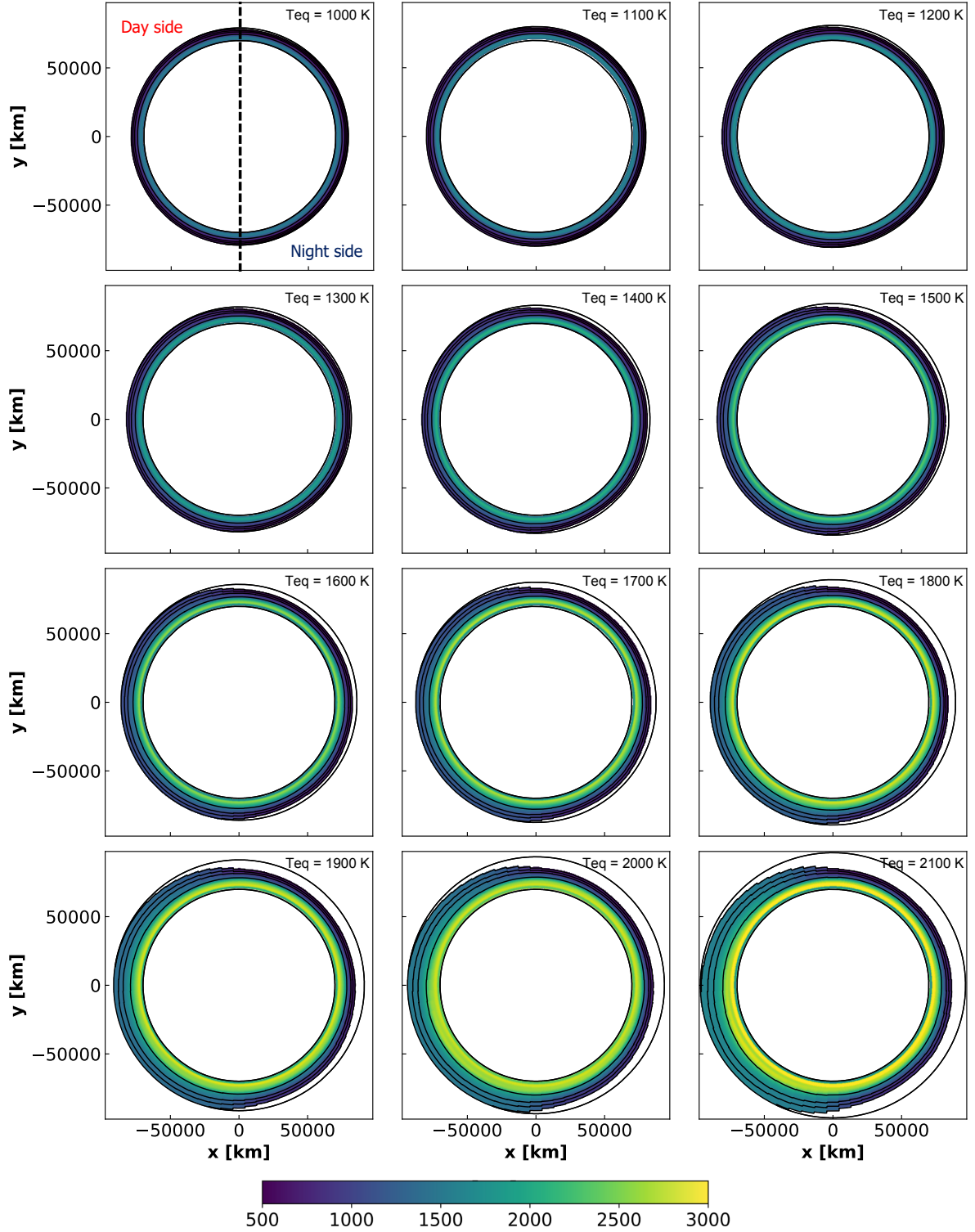
The thermal profiles remain below the temperature required for a thermal dissociation of H<sub>2</sub> or H<sub>2</sub>O, resulting in a total compositional homogeneity. Only a very particular region of the two hottest simulations allows a very low thermal dissociation of H<sub>2</sub>O and H<sub>2</sub> that is associated with a decrease in H<sub>2</sub> and H<sub>2</sub>O abundances by a factor of about 1.2. Moreover, as the concerned regions are contained between  $50^\circ$  and  $-50^\circ$  in latitude,  $10^\circ$  and  $70^\circ$  in longitude, and  $2 \times 10^4$  Pa and  $5 \times 10^1$  Pa in altitude, they are not probed in transmission. For this study, we can therefore consider that these atmospheres are chemically homogeneous.

The spectra generated by Pytmosph3R for the 12 no thermal inversion simulations are shown in Fig. 6. If the spectra in transit for the coldest simulations (from  $T_{\text{eq}} = 1000$  K to  $T_{\text{eq}} = 1500$  K) are indeed very similar, the CO absorption bands for the warmer spectra are much less marked. The fact that H<sub>2</sub>O does not dissociate in these simulations without thermal inversion implies that H<sub>2</sub>O and CO contribute similarly to the transmission spectra. Thus, the CO bands, although present, do not stand out as clearly as when H<sub>2</sub>O is dissociated and the spectrum probes these molecules at different temperatures (see Fig. 6).

We note that the  $T_{\text{eq}} = 2100$  K spectrum is nearly superimposed on the  $T_{\text{eq}} = 2000$  K spectrum above 1 μm and then very close to the  $T_{\text{eq}} = 1900$  K below 1 μm. This spectral crossover is not found at lower values of  $T_{\text{eq}}$ , for which spectra are well distinct and overall shifted to higher apparent radii as  $T_{\text{eq}}$  increases. This could be explained by the strong east-west asymmetry of the limb produced by the zonal circulation that is also responsible for the hot-spot shift. The atmosphere is much more extended on the east limb compared to the west. This trend is already visible in the simulations at  $T_{\text{eq}} = 1900$  K and  $T_{\text{eq}} = 2000$  K, but it is more dramatic for the simulation at  $T_{\text{eq}} = 2100$  K because the atmosphere shift is more intense here, as shown in Table 2, which indicates the angle between the hottest point and the substellar point of each simulation.

## 4. Retrieval results

We performed 1D retrievals using TauREx on the set of transmission spectra previously described. First, we computed retrievals assuming only H<sub>2</sub>O and CO as atmospheric trace gases in order



**Fig. 5.** Equatorial temperature maps for the 12 no thermal inversion simulations performed without optical absorbers (VO and TiO), corresponding to 12 values of  $T_{eq}$ . The day-night dichotomy is shallower than in Fig. 3 and the atmospheres extend less far because the scale heights are smaller. The substellar point and the hottest region of the atmosphere for the hottest planets are shifted (see Table 2).

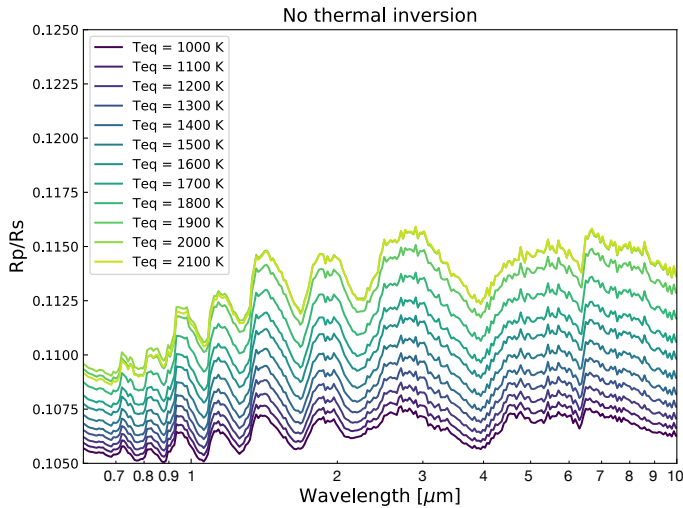
to limit the number of free parameters, save computation time, and better identify biases. For each simulation, we retrieved four parameters: the planetary radius,  $H_2O$  and  $CO$  abundances, and a “gray” cloud layer. We also used the following two different TP profiles: first, an isothermal profile, in which a single temperature is assumed for the whole atmosphere. This assumption

is relevant for cold enough planets in which only a thin part of the atmosphere is probed. Second, a four-point thermal profile that is parameterized by four temperatures and two pressures. The top and bottom pressures are fixed at the extremes of the atmospheric model. This 1D vertical profile assumes a homogeneous atmosphere in latitude and longitude, but introduces more

**Table 2.** Shift of the hot spot.

GCM w/ TI		GCM w/o TI
$T_{\text{eq}}$ (K)	$\Gamma$ ( $^{\circ}$ )	$\Gamma$ ( $^{\circ}$ )
1000	–	$\sim 6$ at $5.3 \times 10^5$ Pa
1100	–	$\sim 6$ at $5.3 \times 10^5$ Pa
1200	–	$\sim 6$ at $5.3 \times 10^5$ Pa
1300	–	$\sim 6$ at $5.3 \times 10^5$ Pa
1400	$\sim 23$ at 11 Pa	$\sim 6$ at $5.3 \times 10^5$ Pa
1500	$\sim 17$ at 4 Pa	$\sim 6$ at $5.3 \times 10^5$ Pa
1600	$\sim 11$ at 8 Pa	$\sim 6$ at $5.3 \times 10^5$ Pa
1700	$\sim 11$ at 32 Pa	$\sim 11$ at $5.3 \times 10^5$ Pa
1800	$\sim 6$ at 44 Pa	$\sim 17$ at $4.6 \times 10^4$ Pa
1900	$\sim 6$ at 63 Pa	$\sim 23$ at $2.3 \times 10^4$ Pa
2000	$\sim 0$ at $2.6 \times 10^2$ Pa	$\sim 28$ at $1.6 \times 10^4$ Pa
2100	$\sim 0$ at $2.6 \times 10^2$ Pa	$\sim 33$ at $5.8 \times 10^3$ Pa

**Notes.**  $\Gamma$  angle (in degrees) between the substellar point and the hottest point of the atmosphere with its pressure indicated for the GCM simulation with and without thermal inversion. The uncertainty on the angle is  $\pm 3^{\circ}$  because of the resolution in longitude.



**Fig. 6.** Transmission spectra calculated with Pytmosp3R from the 12 simulations without optical absorbers (VO and TiO). As these simulations are overall colder without optical absorbers, the CO bands are hidden in the water bands and the apparent radii are smaller than the spectra in Fig. 4.

freedom with a possible variation in altitude. This assumption is relevant when significant vertical variations are expected on the limb. With six parameters added, this profile costs more in computing time.

Then, we proceeded with a full retrieval analysis including TiO and VO abundances (two more parameters), in addition to the abundances of CO and H<sub>2</sub>O, as well as the four-point thermal profile. Although the calculation took more time to converge, these tests allowed us to investigate the biases of a more complex atmosphere and to determine whether TiO and VO spectral features better constrain the retrieval or not.

For every retrieval we report here, the retrieved abundances are constant for the whole atmosphere and the  $\frac{\text{He}}{\text{H}_2}$  ratio is set at the solar metallicity. In addition, TauREx does not take into account the thermal dissociation of species.

#### 4.1. Isothermal retrievals

The isothermal profile is commonly used in retrieval analyses of transmission spectra based on the assumption that only a small and homogeneous region around the terminator is probed (Evans et al. 2018; Tsiaras et al. 2018; Edwards et al. 2020). Moreover, its reduced number of free parameters is consistent with the weak information content of most low-resolution small-bandwidth available spectra. This assumption is often well justified, but deserves to be tested in the context of forthcoming higher-quality spectra that JWST and Ariel will deliver.

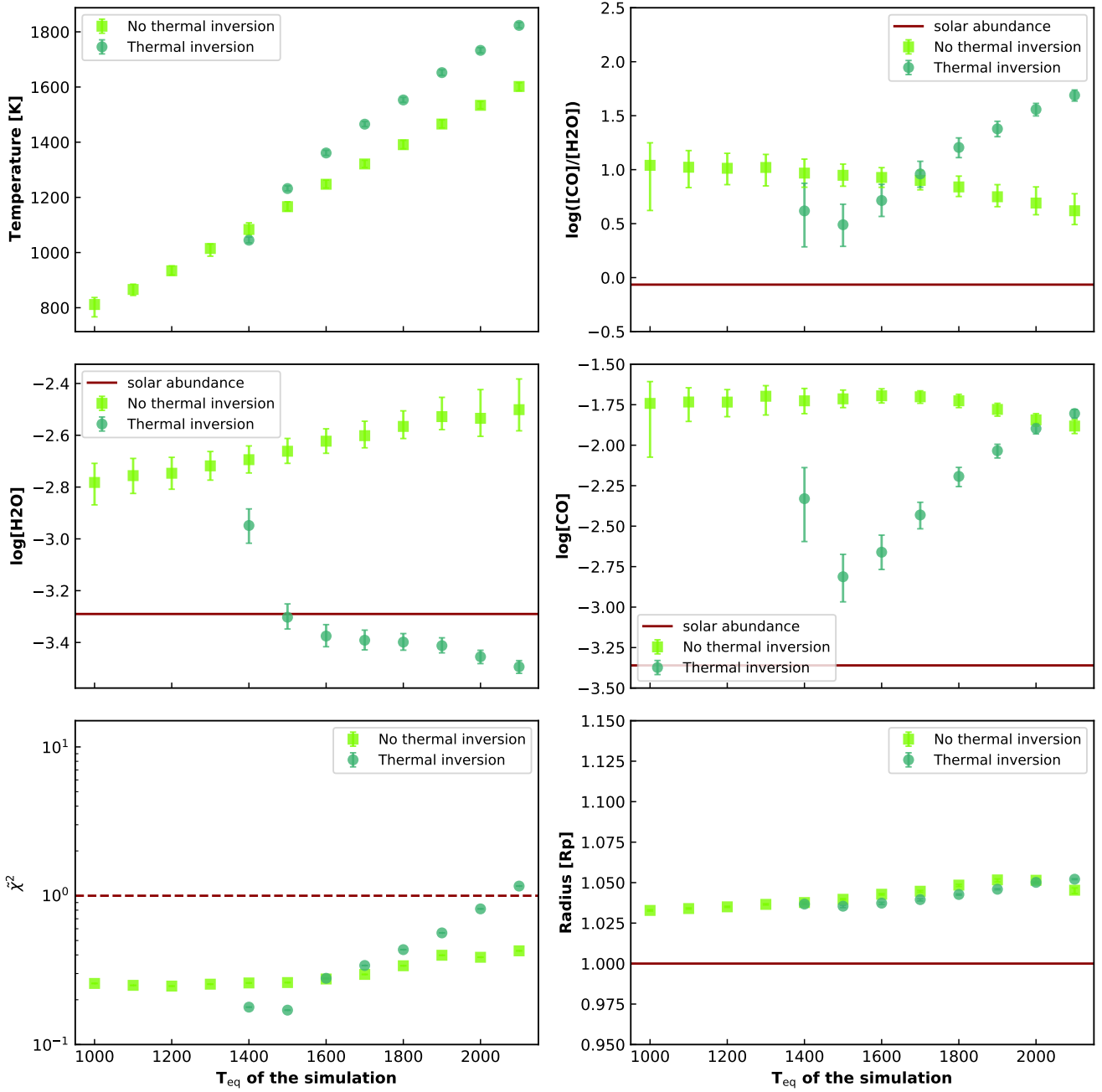
The retrieval results obtained with TauREx using the isothermal assumption are shown in Fig. 7. For each simulation, we show the temperature, the planetary radius, and the H<sub>2</sub>O and CO abundances we retrieved. We also show the CO over H<sub>2</sub>O abundance ratio, which indicates a departure from the solar composition. As explained in Sect. 2.4, we used nonrandomized spectra and we indicate the  $\chi^2$  that is computed with the best-fit parameter values derived from the posterior distribution given by TauREx. We stress (i) that with nonrandomized spectra,  $\chi^2 \ll 1$  is an acceptable value and not the sign of noise fitting, and (ii) that the model comparison was made with logarithmic Bayes factors.

TauREx always converges to a single solution, but the goodness of the fit will differ depending on whether TiO and VO are present or absent. In the case of simulations without TiO and VO, TauREx always produces a good fit ( $\chi^2 \ll 1$ ), with a notable slight increase in  $\chi^2$  for the hottest simulations. The outcome is different for simulations with TiO and VO. For the coldest simulations, therefore with little thermal dissociation in the atmosphere (see Fig. 3),  $\chi^2$  is lower than 1. This is consistent with simulations without TiO and VO in the atmosphere where no thermal dissociation takes place. In contrast,  $\chi^2$  increases very clearly with increasing equilibrium temperature, exceeding 1 for all the other simulations. This behavior is similar to that shown by Pluriel et al. (2020b), indicating that TauREx fails to correctly take the thermal dissociation of H<sub>2</sub> into account.

The results show that none of the retrieved abundances agree with their actual values in the simulations. As a reminder, we simulated atmospheres with solar abundances and where CO does not dissociate. Therefore we would expect to retrieve solar abundances for CO, and abundances smaller than or equal to the solar abundance for H<sub>2</sub>O (because H<sub>2</sub>O can dissociate).

The retrieval results from the simulations without TiO and VO in the atmosphere are wrong. The retrieved [CO]/[H<sub>2</sub>O] is equal to about 10 and the retrieved abundances of CO and H<sub>2</sub>O are always above the solar abundance. This indicates that TauREx is not able to retrieve results consistent with the ground truth, even though we have a visually good fit and  $\chi^2 < 1$ , all of which would give an undue confidence to an observer. Figure 8 shows an example of a good agreement between the data and the retrieval (for the simulation at  $T_{\text{eq}} = 2100$  K) that leads to incorrect parameter values. These results are unexpected because, as shown in Fig. 5, the temperature never reaches a value high enough for species such as H<sub>2</sub>O or H<sub>2</sub> to dissociate. Therefore we do not expect to observe strong compositional heterogeneities between the day- and nightside, and therefore we would expect to retrieve solar abundances. We show in Sect. 4.2 that the assumption of an isothermal profile is the main reason for retrieving incorrect parameter values.

Retrievals are more consistent in the case of thermal inversion simulations including TiO and VO. For the coldest simulations, the retrieved H<sub>2</sub>O abundance is almost solar even if it slightly decreases from the solar abundance for the hottest

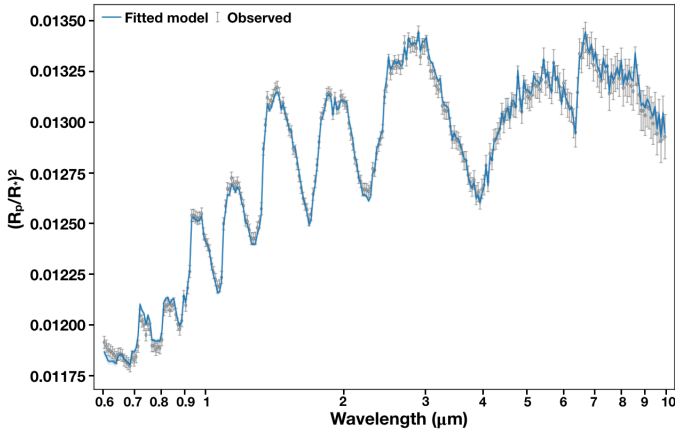


**Fig. 7.** Retrieval results with TauREx assuming an isothermal profile. Cases with and without thermal inversion in the input atmosphere are shown in light and dark green, respectively. We plot the temperature (*top left*),  $[\text{CO}]/[\text{H}_2\text{O}]$  (*top right*), the CO (*middle right*) and the  $\text{H}_2\text{O}$  (*middle left*) abundances, the planetary radius (*bottom right*) within  $1\sigma$  error, and  $\chi^2$  (*bottom left*). These results have been obtained with nonrandomized spectra but  $\chi^2$  calculations, and TauREx assumes a 30 ppm  $1\sigma$  Gaussian noise on the whole spectral domain. The red line represents the input value in the GCM simulations, and the dashed red line shows where  $\chi^2 = 1$ .

simulations. The CO abundance starts to be well constrained in the coldest simulation, then it becomes increasingly biased with the hottest simulations. Thus, we observe an increase in  $[\text{CO}]/[\text{H}_2\text{O}]$  when we simulate hotter Jupiters. This behavior indicates that the 1D retrieval models are less biased when applied to cooler atmospheres. The difference between the scale height on the day- and nightside is reduced when the temperature drops, therefore the difference in altitude of the CO probed on the dayside compared to the  $\text{H}_2\text{O}$  probed on the nightside decreases. On the other hand, the amount of dissociated  $\text{H}_2\text{O}$  also decreases for the coldest planets, implying that the

atmosphere is probed less deeply and at higher temperatures, hence a smaller difference in the spectra appears in Fig. 4.

We observe a regular increase in the retrieved temperature with the equilibrium temperature of the planet for simulations without TiO and VO in the atmosphere, which is expected as the atmosphere are globally hotter (see Fig. 5). For simulations with TiO and VO in the atmosphere, the increase in temperature is regular until a break in the slope from  $T_{\text{eq}} = 1500$  K. This behavior can be explained by the shape of the temperature maps in Fig. 3, where the east-west asymmetry becomes less important for the hottest cases. TauREx seeks here to best fit the very



**Fig. 8.** Best-fit retrieved spectrum using an isothermal profile parameterization (in blue). The input spectrum has been calculated with Pytmosph3R for the  $T_{\text{eq}} = 2100$  K simulation without optical absorbers (VO and TiO) in the atmosphere (in gray).

strong absorption bands of CO, and does so by increasing the abundance of the species, so does the retrieved temperature.

The retrieved radius is almost constant for all the simulations, quite close to the input radius of the simulations. It is therefore difficult to extract much information from this retrieved parameter, which does not seem to play an important role here in improving the fit of TauREx retrievals.

To summarize this section, isothermal retrievals are insufficient to obtain the complexity of hot exoplanetary atmospheres, even for the more homogeneous simulations such as the one without optical absorbers. The discrepancy between the retrieved parameters and the ground truth is usually considerable, much larger than uncertainties estimated from posterior distributions and regardless the goodness of the fit. The following step is thus to determine whether assuming a more complex vertical thermal profile might solve this issue.

#### 4.2. Four-point TP profile retrievals

We present here the results of the retrieval procedure that no longer assumes an isothermal vertical profile, but a six-parameter thermal profile (Sect. 4). The retrieved values for all simulations are presented in Fig. 9.

##### 4.2.1. Without thermal inversion

We first focus on the input simulations (Fig. 1) without optical absorbers (TiO and VO) in the atmosphere that therefore is without thermal inversion. The retrieval results are here consistent with the input models. The retrieved abundances of CO and H<sub>2</sub>O, consequently [CO]/[H<sub>2</sub>O], are now well constrained and fit the solar abundance within  $1\sigma$  in every retrieval. We see here that vertical effects are not negligible when these atmospheres are to be retrieved correctly because isothermal retrievals were biased. Thus, we need to let TauREx fit its own vertical profile. The thermal profiles retrieved within  $1\sigma$  are shown in Fig. 10 (top). We used a log-linear interpolation between the temperature nodes. To quantify the improvement of retrieving a four-point thermal profile compared to the isothermal assumption, we calculated the logarithmic Bayes factor following Eq. (4).

Table 3 shows that TauREx strongly favors the four-point thermal profile compared to the isothermal one. To understand

why vertical effects are important, we show in Fig. 11 the contribution function of the transmission spectra for the  $T_{\text{eq}} = 2100$  K simulation and the thermal profiles, both retrieved and from the input GCM. We only plot the GCM profile around the limb to focus on the probed region. The contribution function shows that the regions probed covers 6 orders magnitude in pressure, from around  $10^5$ – $10^{-1}$  Pa. Depending on the wavelength, the features in the transmission spectra are therefore coming from regions at different temperatures.

An isothermal profile does not manage to describe this complexity, especially where a broad range of pressure is probed. Figure 11 shows the thermal profile at the equator for every longitude in the hottest simulation. The blue and red curves represent the antistellar and substellar regions of the atmosphere, respectively. The regions probed in transmission are thus mainly represented by the green curves. We also show the isothermal and vertical retrieved thermal profiles in Fig. 11. The TP profiles cannot be well approximated by an isothermal profile because the input profiles move away from an isothermal profile by several hundred kelvin when the vertical retrieved TP profile fits the probed GCM TP profile better. As the contribution plot at the bottom of Fig. 11 shows, however, regions from  $\sim 10^4$  Pa (around  $0.7 \mu\text{m}$ ) to  $\sim 10^{-1}$  Pa are probed. From the shape of the thermal profiles in the terminator region and from the large pressure range probed by the transmission spectrum, we can conclude that an isothermal model is poorly suited and will either fail to yield a correct fit or will yield a correct fit with parameters that significantly depart from the ground truth to compensate for this limitation. We clearly belong to the second case here as TauREx has to increase the species abundances well over their actual values to match the observed features, which is no longer needed when we allow the profile to be nonisothermal.

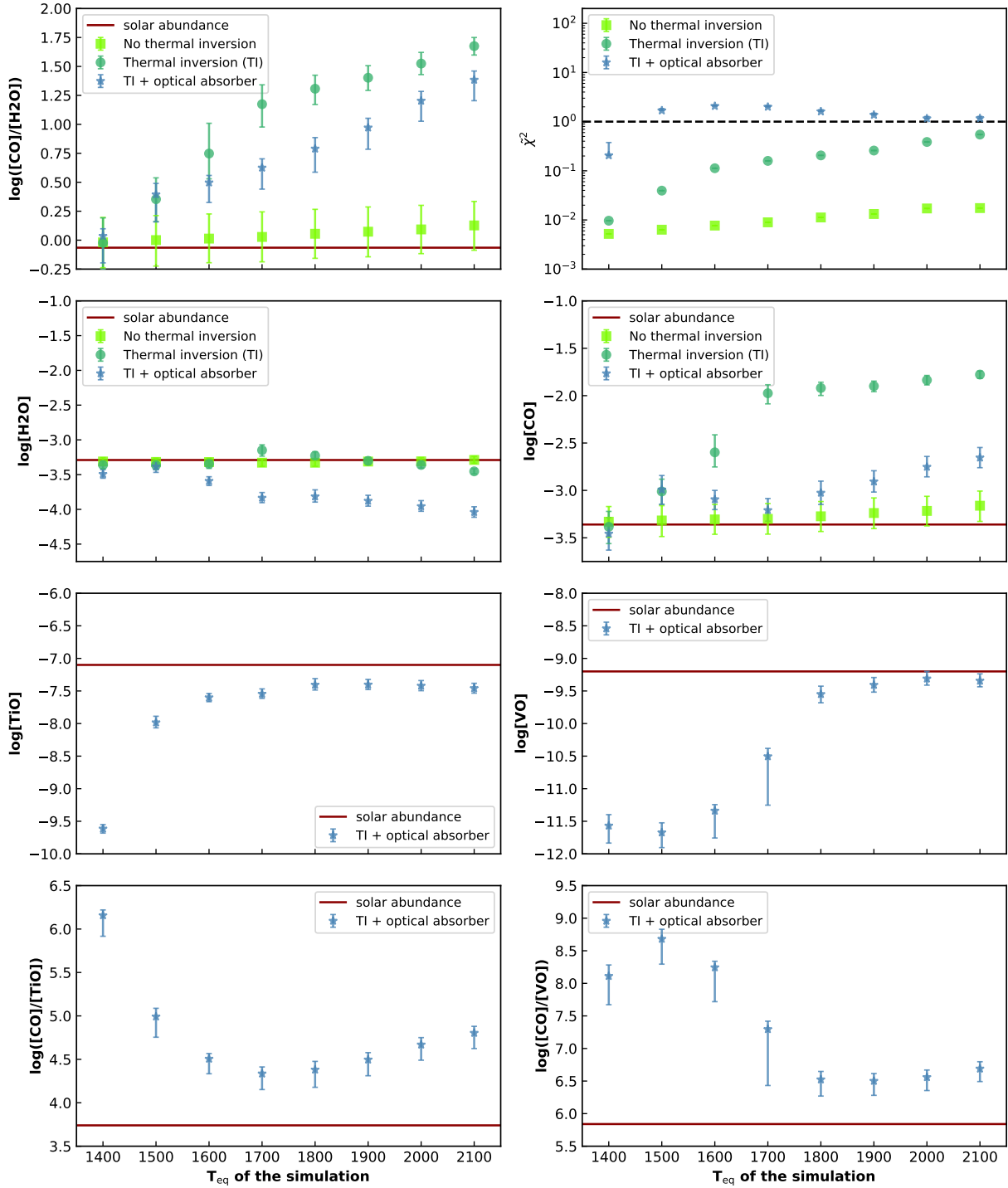
For this reason, the model using a vertical profile retrieves the input abundances, as shown in Fig. 9, with a higher level of confidence than in an isothermal retrieval, as is shown in Table 3. This is an important result because it clearly indicates that the assumptions of 1D retrieval models are justified to analyze and interpret the observations obtained on planets that are not too hot. However, it confirms the caveat about the isothermal assumption (see Sect. 4.1), which leads to incorrect parameter values despite an excellent agreement ( $\chi^2 \leq 1$ ). Of course, this is a simplified model and we can well imagine the presence of a species on one side of the atmosphere and not on the other, which could bias the results, even assuming a nonisothermal vertical TP profile.

##### 4.2.2. With thermal inversion

###### Thermal inversion without optical absorbers in the retrieval

We now analyze with TauREx the spectra from the thermal inversion simulations, but ignore the existence of TiO and VO as absorbers. TiO and VO were included in the GCM simulations, but are not included in the forward radiative transfer model of Pytmosph3R and their abundance is thus not a parameter to be retrieved, in contrast to CO and H<sub>2</sub>O.

Unlike the previous section, the biases that are observed with the retrieved isothermal profiles remain with a four-point thermal profile. Figure 12 helps understand why shaping the vertical profile does not improve the results. It shows the contribution function of the transmission spectra in the  $T_{\text{eq}} = 2100$  K simulation as well as the simulated and retrieved thermal profiles. We only plot the GCM profile around the limb to focus on the probed region. In contrast to the simulations without thermal inversion

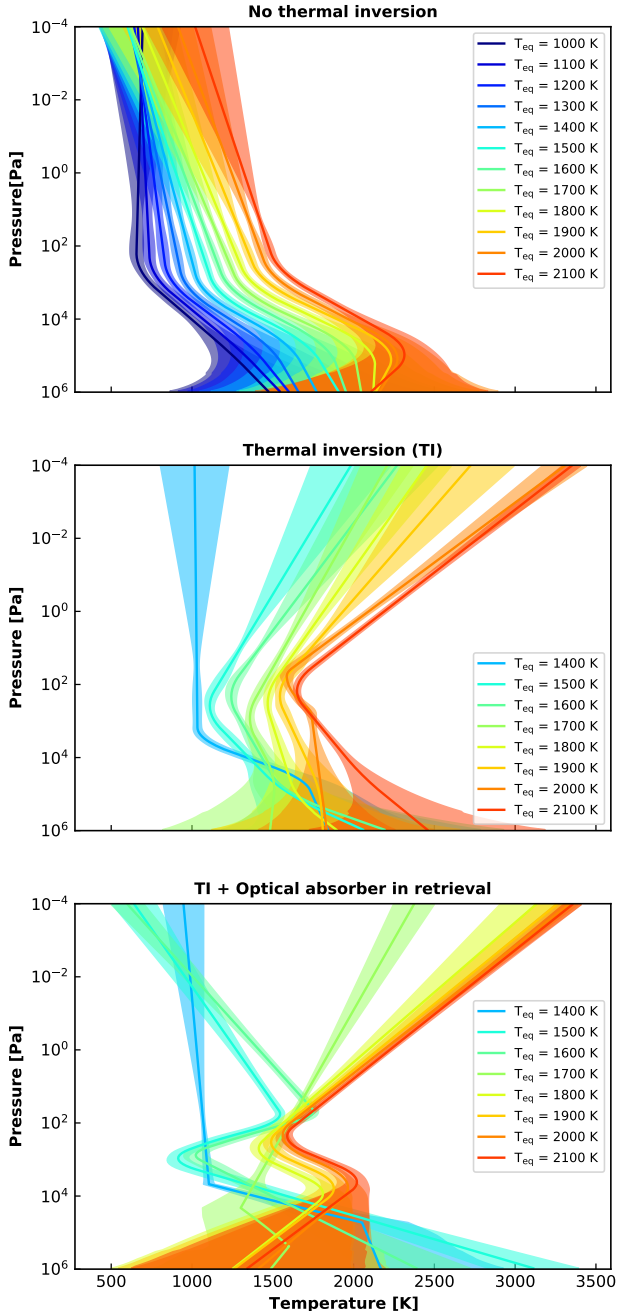


**Fig. 9.** Retrieval results with TauREx using a four-point parameterization for the vertical thermal profile. We present four sets of retrievals. The first set corresponds to the analysis of spectra obtained from the no thermal inversion simulations (light green squares). Then we have three different ways to analyze the spectra from the thermal inversion simulations. The dark green circles show retrievals ignoring TiO and VO, then the light blue triangles and dark blue stars shows retrievals considering TiO and VO without and with condensation, respectively. We plot  $[\text{CO}]/[\text{H}_2\text{O}]$ ,  $[\text{CO}]/[\text{TiO}]$ ,  $[\text{CO}]/[\text{VO}]$ , and the CO, the  $\text{H}_2\text{O}$ , the TiO, the VO abundances within  $1\sigma$  errors and  $\chi^2$ . These results have been obtained with nonrandomized spectra, but  $\chi^2$  calculations and TauREx assumes a 30 ppm  $1\sigma$  Gaussian noise on the whole spectral domain. The red line represents the input value in the GCM simulations, and the dashed red line shows where  $\chi^2 = 1$ .

(Fig. 11), the transition between the day- and nightsides of the planet is sharper at the pressure below the quasi-isothermal layer, the global gradient in order of magnitude (because this depends on the latitude, longitude, and altitude) between the day- and nightside is about  $600 \text{ K}/10^\circ$  when it was more around  $100 \text{ K}/10^\circ$  in the other simulations. For this reason, it is impossible for the

retrieval to find a 1D thermal profile that both fits this simulation and find consistent parameter values.

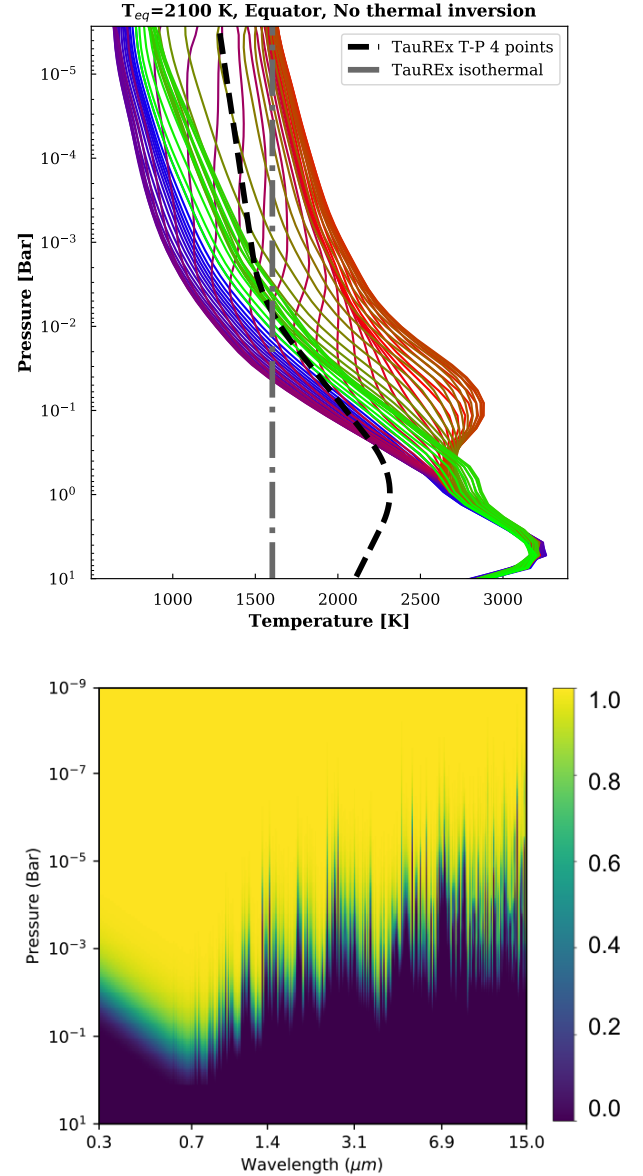
The retrieved water abundances (Fig. 9) are now solar for every simulation, and the CO abundances deviates more than those in the isothermal retrieval. As a result, the retrieved  $[\text{CO}]/[\text{H}_2\text{O}]$  is slightly higher than the one inferred with the



**Fig. 10.** Retrieved temperature profiles for the GCM simulation. Without thermal inversion (*top*); with thermal inversion assuming a simplified atmosphere (*middle*); with thermal inversion assuming a more complex atmosphere with optical absorber (TiO and VO) in the atmosphere (*bottom*).

isothermal model. We also note here a break in the slope at  $T_{\text{eq}} = 1700$  K: above this equilibrium temperature, the biases increase more slowly than below.

The temperature and  $\text{H}_2\text{O}$  abundance maps in Fig. 3 show that for the coldest simulation ( $T_{\text{eq}} = 1400$  K), the spectrum is only sensitive to the shape of the thermal profile because the dissociation of  $\text{H}_2\text{O}$  and  $\text{H}_2$  remains weak, hence a retrieved solar  $[\text{CO}]/[\text{H}_2\text{O}]$ . Then, from  $T_{\text{eq}} = 1500$  K to  $T_{\text{eq}} = 1700$  K, the east-west asymmetry of the atmosphere combined with increasing thermal dissociation on the dayside yields a large increase in the overestimation of the retrieved abundances by a combination of horizontal effects both along and through the limb.



**Fig. 11.** *Top*: equatorial thermal profiles from the no thermal inversion  $T_{\text{eq}} = 2100$  K simulation. Each color represents a longitude from antistellar (blue) to substellar (red). Profiles near the terminator and probed by transmission are green. The dotted gray (isothermal) and dashed black lines are the retrieved profiles. *Bottom*: transmittance map (color bar from 0 to 1) for each wavelength at each pressure.

Finally, above  $T_{\text{eq}} = 1700$  K, although the thermal dissociation of the  $\text{H}_2\text{O}$  and  $\text{H}_2$  intensifies, the east-west asymmetry becomes almost nonexistent, which could explain the break in slope observed in the retrieved  $[\text{CO}]/[\text{H}_2\text{O}]$ . For these hottest simulations, only horizontal effects through the limb seem to dominate the transmission spectra. Even if the retrieval using a nonisothermal profile fits the spectra better, the observed biases remain.

The retrieved thermal profiles are shown in Fig. 10. TauREx favors a thermal inversion in all cases.

We performed the same statistical analysis as in Sect. 4.1 with these simulations with thermal inversion, but ignoring the existence of TiO and VO as absorbers. We calculated the logarithmic Bayes factor (Eq. (4)). We present the value for each simulation in Table 3. They indicate, as in Sect. 4.1 with this simulations with no thermal inversion that TauREx strongly favors

**Table 3.** Logarithmic Bayes factor.

Bayes factor (vertical vs isothermal TP profiles)		
$T_{\text{eq}}$ (K)	w/o thermal inversion	w/ thermal inversion
1000	32.10	–
1100	31.62	–
1200	31.26	–
1300	33.14	–
1400	33.26	20.06
1500	33.98	15.17
1600	35.55	19.82
1700	38.14	22.72
1800	43.10	27.43
1900	51.60	32.82
2000	50.13	50.25
2100	55.30	77.74

**Notes.** Logarithmic Bayes factor comparing retrievals made assuming an isothermal or nonisothermal vertical profile for both the thermal inversion and no thermal inversion simulations. Bayes factors for nonisothermal profiles are 15–78 times larger than those obtained with isothermal profiles, suggesting a very strong preference for the nonisothermal model.

the model with a four-point TP profile compared to the isothermal model. The four-point profile is even more suitable as the equilibrium temperature increases. While  $\tilde{\chi}^2$  was always greater than 1 in the isothermal retrieval, it is now below 1 for each simulation. Therefore, if the fits could have been rejected in this first case, TauREx tells us here that the fits are much better.

#### Thermal inversion with optical absorbers in the retrieval

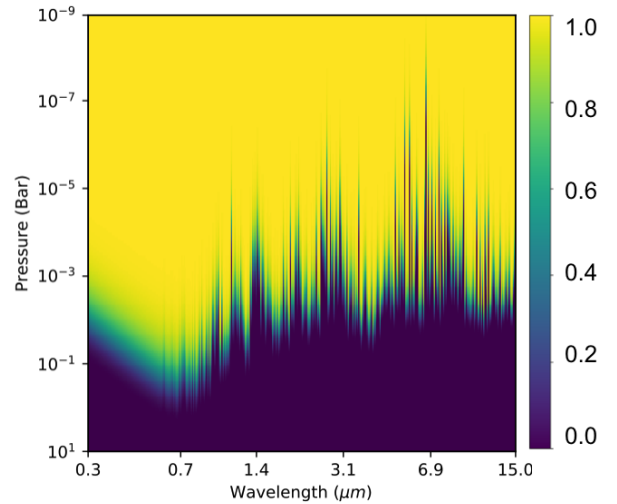
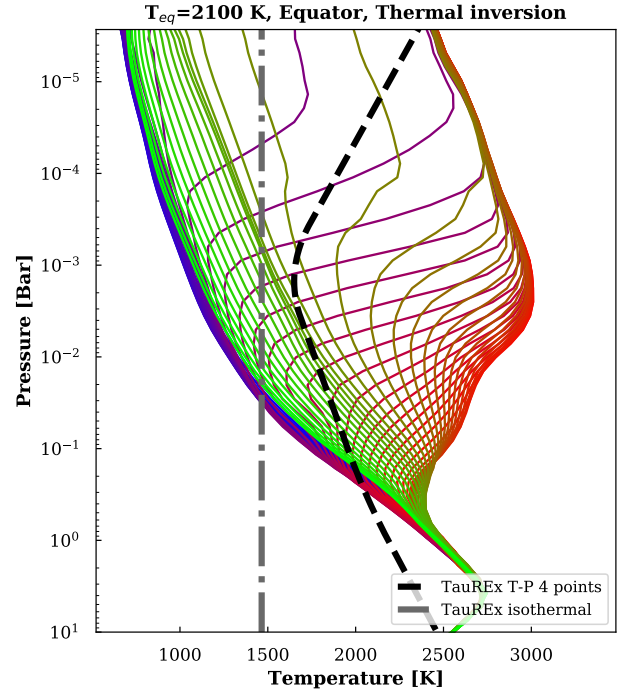
Finally, we studied what happens in more complex atmospheres where TiO and VO are present in addition to CO and H<sub>2</sub>O. Retrieval results, presented in Fig. 9 (blue), remain close to those with an atmosphere containing only CO and H<sub>2</sub>O.

[CO]/[H<sub>2</sub>O] is still biased, all the more so with hotter atmospheres. Differences appear in the retrieved abundances that are globally not solar in every simulation. TiO and VO abundances are underestimated in the coldest simulations, and they are close to the solar abundances in the hottest. This behavior can be explained by the condensation of these species, which decreases their observed VMRs.

In addition, absorption features of VO in low resolution are hidden by the TiO bands, hence they weakly constrain the retrieval that reaches the limits of the priors (Fig. A.1). CO abundances are still overestimated, but less so than in the previous retrievals (one order of magnitude less in average). However, retrieved H<sub>2</sub>O abundances are underestimated, which explains the biased [CO]/[H<sub>2</sub>O].

It is interesting to note that allowing the code to add TiO and VO, which are present and affect the spectra, does not lead to a better agreement, except partially for CO. The retrieved H<sub>2</sub>O abundances are even worse. We assume that this is due to the model having a smaller margin to find a degenerated 1D solution. We also note that even when using a four-point TP profile,  $\tilde{\chi}^2$  is above 1 for  $T_{\text{eq}}$  in the 1500–2100 K range. Thus a  $\tilde{\chi}^2$  test in these cases might be used as a warning.

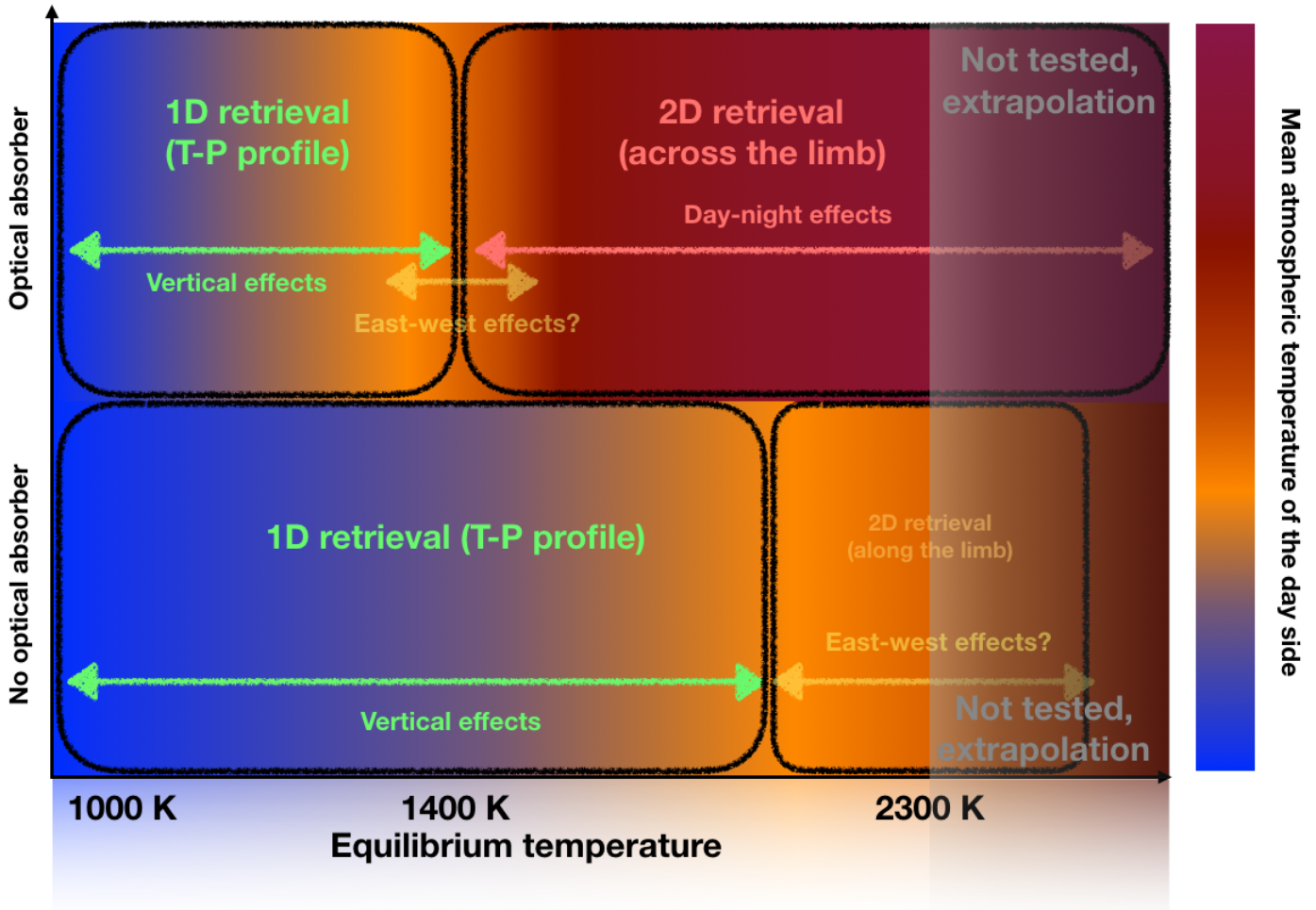
As only CO does not dissociate or condense, we also plot [CO]/[VO] and [CO]/[TiO] to determine how far they are from the expected solar abundances. They are both biased, especially



**Fig. 12.** *Top:* equatorial thermal profiles from the thermal inversion  $T_{\text{eq}} = 2100$  K simulation. Each color represents a longitude from anti-stellar (blue) to sub-stellar (red). The dotted gray (isothermal) and dashed black lines are the retrieved profiles. *Bottom:* transmittance map (color bar from 0 to 1) for each wavelength at each pressure.

in the coldest simulations, but this is due to both the condensation and the fact that VO features are hidden by TiO features in low resolution.

We conclude that 1D retrieval models are able to retrieve abundances within  $1\sigma$  of their actual values in warm atmospheres, where the atmospheres are homogeneous in latitude and longitude. The errors bars estimated by TauREx underestimate the departure, however. In the atmospheres we studied, the transmission spectra are dominated by vertical effects at the limb, which can be well reproduced by 1D models. However, these models cannot correctly reproduce the complexity of the 3D structure of hot exo-atmospheres, starting from  $T_{\text{eq}} \approx 1500$  K. The biases highlighted by Pluriel et al. (2020b) therefore cover a larger number of objects, from HJs to UHJs (see Sect. 5 for more



**Fig. 13.** Summary of the different geometries required in retrieval codes to avoid biases as a function of the equilibrium temperature of the planet and the presence or absence of optical absorbers (hence thermal inversion). 1D retrieval models appear to provide relatively satisfactory parameter estimates for planets with equilibrium temperatures lower than 1400 K when optical absorbers (TiO, VO, K, Na, metals, ionized hydrogen, etc) are present in the atmosphere. However, they lead to biased parameter estimates above this limit where 2D or 3D retrieval codes are mandatory. When no optical absorbers are present in the atmosphere, the validity of 1D retrieval code extends to an equilibrium temperature of 2000 K. Above this temperature, the estimated parameters become biased, probably due to east-west effects. Although we suggest clues for the effect of the east-west asymmetry, further investigations are needed to quantify their effects.

details). Moreover, if metals such as Fe or Mg and ionized hydrogen were present in the atmosphere (as discussed in Sect. 2.1), these species would increase the atmospheric dichotomy, hence the magnitude of 3D effects, thus pushing the 1D-model safe zone to even colder equilibrium temperatures.

We also conclude that the model with a nonisothermal vertical profile is not relevant for the study of WASP-121 b with  $T_{eq} \approx 2350\text{ K}$  (Pluriel et al. 2020b), where the horizontal heterogeneities through the limb have a greater contribution on the transmission spectrum than those due to vertical differences. Even if the Bayes factor of the simulations had been better with a four-point thermal profile, we have shown here that this would not modify the biases observed. A simpler retrieval with fewer parameters is meaningful because the posterior distributions are very similar.

### 5. 3D effects on 1D retrievals

The three-dimensional structure of HJ and UHJ atmospheres strongly affects transmission spectroscopy and will bias the 1D

retrieval models that are used to analyze and interpret future observational data from the JWST and Ariel. These spectra carry information coming from various regions of the atmosphere, and it is often difficult to disentangle them. The 3D structure implies variations in the physical and chemical properties of the atmosphere that affect the transmission spectrum along three main axes: (i) variations as a function of altitude, that is, the vertical effects; (ii) the north, south, east, west variations that we refer to as horizontal effects along the limb; (iii) and the variations between the day- and nightside, also referred to as horizontal effects through the limb. By ranking the effects of these three contributions, we can determine how biased the retrieval models are. We have quantitatively characterized the observed biases, thus allowing a more exhaustive understanding of the effects involved, and we have highlighted the origins as well as the limits of these biases. We present in Fig. 13 a summary of our study, where we suggest to the community which types of retrieval should be used depending on the equilibrium temperature of the planet and the presence or absence of optical absorbers in the atmosphere.

### 5.1. Vertical effects

We highlight here the effect of the physical and chemical variations with altitude. These effects are currently considered in atmospheric studies because most of them assume that the probed area in transmission remains a thin annulus around the terminator (Tinetti et al. 2007; Redfield et al. 2008; Tsiaras et al. 2018; Skaf et al. 2020; Pluriel et al. 2020a). 1D retrieval models are able to reproduce vertical effects with good accuracy because they are able to retrieve vertical TP profiles. When the transmission spectrum is mainly affected by temperature and/or chemical variations with altitude, models such as TauREx therefore accurately fit the observations and derive physical and chemical characteristics consistent with the input simulation or observation.

We demonstrated this with the retrieval on simulations without thermal inversion (Sect. 4.2.1). As a reminder, in these cases, the planet atmospheres remain relatively homogeneous and do not present a very strong day-night dichotomy, as shown in Fig. 5. Thus, on the one hand, the regions probed by the transmission spectra remain close to the terminator, and on the other hand, this terminator is fairly homogeneous with slight east-west variations. The vertical effects at the terminator then dominate the spectrum shapes, thus allowing relevant and accurate 1D retrievals (Fig. 9). However, we also demonstrated that the isothermal hypothesis on the retrieval model should no longer be used for planets with an equilibrium temperature higher than 1000 K as it results in the inference of incorrect abundances despite a good fit ( $\chi^2 \leq 1$ ).

1D retrieval models therefore remain suitable for studying HJs atmospheres that are not too hot ( $T_{\text{eq}} \leq 1400$  K), where horizontal variations along and through the limb can be neglected. However, we must remain aware that the region probed in this case mainly comes from a thin layer around the terminator. To determine the characteristics of the entire atmosphere, we need to compare these results with GCM simulations that model the entire atmosphere.

Finally, we find that when the atmosphere presents a very hot stratosphere ( $T_{\text{eq}} \geq 1500$  K), the shape of spectral features is no longer controlled by vertical gradients, but by horizontal ones. The results of the retrieval depart from the actual values far beyond estimated error bars and despite an excellent spectral match ( $\tilde{\chi}^2 \leq 1$ ). This indicates that 1D vertical models provide an unrealistic solution to the observed spectra with good fits, but with biased  $\text{H}_2\text{O}$ , VO, TiO, and CO abundances.

### 5.2. Horizontal effects along the limb

A second geometric effect can affect the shape of transmission spectra: differences along the limbs, in particular between the east and west limbs. This occurs when strong jets are present in the atmosphere or when the atmosphere is in super-rotation. This then creates an east-west asymmetry with sometimes strong temperature differences, which significantly affects the spectrum. This effect is highlighted in Figs. 3 and 5, which compare the west and east limbs of two simulations, a cold ( $T_{\text{eq}} = 1400$  K) and a hot one ( $T_{\text{eq}} = 2100$  K). They show an east limb that is generally colder than the west limb. The temperature difference becomes significant for the simulation with a hot stratosphere (with TiO/VO in the atmosphere) down to  $T_{\text{eq}} = 1700$  K. This is a clue to explain the break in slope in the retrieved  $[\text{CO}]/[\text{H}_2\text{O}]$  ratios, which is observed in Fig. 9.

Several teams have studied east-west heterogeneities, whether in simulated or real observations. In particular,

Line & Parmentier (2016) showed that the presence of a nonuniform cloud layer at the level of the terminator can affect the interpretation of transit observations. Powell et al. (2019) studied the impact of a heterogeneous cloud cover on the limbs of HJs, showing that the difference of the cloud properties between the east and west limbs affects the transmission spectra. GCM simulations also highlight these east-west asymmetries at the terminator, which seem to be consistent with observations by Showman et al. (2008a). MacDonald et al. (2020) also studied the biases generated by these east-west differences and demonstrated their importance for the results of retrieval models. Taking this heterogeneity at the terminator into account also allowed them to explain the unexpectedly cold temperatures retrieved for exoplanets WASP-17 b and WASP-12 b.

It would be possible to solve this problem by increasing the temporal resolution between ingress and egress of the transit in order to differentiate at least two spectra, one coming from the eastern limb, the other from the western limb. Thus, we would end up with only the vertical differences that 1D data inversion models can handle. Another way to separately obtain the information on each limb would be to analyze the phase curve. If the phase curve contained the observations of the transit ingress and egress with sufficient sampling, we could obtain separate transmission spectra for the east and west limbs.

### 5.3. Horizontal effects across the limb

We focus here on the effect of the day-night thermal and chemical gradients on the transmission spectrum. These horizontal effects are very often overlooked, assuming that the transmission method only probes a thin annulus at the terminator, which is verified only for small enough atmospheric scale heights (see Fig. 2 by Caldas et al. 2019). As shown by Pluriel et al. (2020b), the inflated dayside atmosphere of UHJs and their effect on the geometry of the observation is shown in Fig. 3: the regions probed during transit are extended on the dayside and are neither thin nor centered on the terminator. In these planets, the vertical effects become negligible compared with horizontal gradients because the day-night temperature contrast reaches thousands of kelvin, while vertical temperature variations probed by the spectrum do not exceed hundreds of kelvin. In addition, gradients along the limb remain weak for the warmest atmospheres, hence a dominance of the effects through the limb in this case.

With nightside signatures hidden by the inflated dayside and mixed signatures at different temperatures and compositions in the observed spectrum, this configuration causes major biases in the properties retrieved by 1D retrieval models. Thus, regardless of the nature of the assumed vertical profiles (isothermal or not), 1D retrieval models fail to reproduce atmospheres faithful to the input GCM models. Furthermore, the thermal dissociation of species adds complexity to these geometric considerations, implying that the transmission spectrum probes a broad range of regions depending on the wavelength.

The effect of the day-night dichotomy on the transmission spectrum remains negligible for cold enough atmospheres because for most of them, they are not hot enough to present a detectable dichotomy. On the other hand, when we study warm and hot atmospheres, we need to consider it because this effect can become dominant in the observations and therefore lead to erroneous interpretations. We note that there is an atmosphere regime (around planet with  $T_{\text{eq}} = 1700$  K) where the three effects described above are of the same order of magnitude, making the analysis of their transmission spectra extremely complex.

## 6. Conclusions

We have demonstrated that the shortcomings of 1D retrievals in the interpretation of transmission spectra are not limited to UHJs and can affect cooler planets as well. The 1D assumption in retrieval models will be an issue when the molecular abundances with future observations provided by JWST and Ariel are to be estimated accurately.

In particular, we have shown that the isothermal assumption leads to incorrect retrieved abundances in every case we studied, even though the retrieved spectrum fits the observational data well. This means that isothermal atmospheres could give suitable spectrum fits, but with very incorrect abundances. Thus, we encourage the community to refrain from using this assumption when a planet with an equilibrium temperature higher than  $T_{\text{eq}} = 1000$  K is studied. While parameterized thermal profiles yield retrieved abundances much closer to the actual abundances in the simulation, we nevertheless note that they produce inaccurate results for very hot atmospheres. This limit is reached for  $T_{\text{eq}} \geq 2100$  K and even down to 1500 K when optical absorbers create a thermal inversion. Our findings confirm that these biases are mainly due to the strong day-to-night dichotomy, as shown by Pluriel et al. (2020b).

Based on our findings, Fig. 13 summarizes our recommendations for the minimum model assumptions necessary to avoid incorrect interpretations and biased retrieved parameters. It can be used as follows:

1. Estimate or calculate the equilibrium temperature of the planet.
2. Verify from a first analysis of the data or estimate if the atmosphere is expected to contain optical absorbers (TiO, VO, K, Na, metals, ionized hydrogen, etc.) that might create a thermal inversion.
3. Adapt the retrieval analysis and its interpretation according to Fig. 13.

If  $T_{\text{eq}} \geq 2100$  K, the parameter values (molecular abundances in particular) and their associated errors derived from a 1D retrieval are very likely to be incorrect. A different retrieval method that accounts for the horizontal structure is then needed. If  $T_{\text{eq}} \leq 1400$  K, the 1D hypothesis yields consistent parameter values and a 1D retrieval analysis can be used. If  $1500 \text{ K} \leq T_{\text{eq}} \leq 2100$  K, either (i) there are no optical absorbers in the atmosphere, and in this case, the 1D retrieval can lead to consistent parameter values, or (ii) there are optical absorbers in the atmosphere, so that a hot stratosphere is likely present and the parameter values (molecular abundances in particular) inferred from a 1D retrieval procedure, as well as their estimated uncertainties, are very likely to be incorrect. In case (ii), we suggest to use a different retrieval framework that takes 2D effects across the limb into account.

We consider that 3D retrieval models based on GCM simulations would face two main issues: first, we would need massive computation power for Bayesian inference, and second, we would have to deal with numerous degeneracies inherent to 3D structures. It would be extremely complex to break these degeneracies, even with the resolution and the accuracy of the JWST or Ariel.

We therefore suggest that 2D retrieval models with a horizontal parameterization across the limb are developed to be able to address the unavoidable imprint of horizontal gradients on the spectra. 2D retrieval models probably have the right level of complexity for this task, therefore we developed a 2D retrieval

code (Falco et al. 2022). We will study its impact on UHJs in a forthcoming paper.

*Acknowledgements.* We thank Michiel Min and Quentin Changeat for the useful discussions about the Bayesian statistical analysis. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 679030/WHIPLASH). We thank the Programme National de Planétologie (CNRS/INSU/PNP) and the CNES for their financial support. W.P. acknowledges financial support from the Swiss National Science Foundation for project 200021\_200726.

## References

- Barton, E. J., Yurchenko, S. N., & Tennyson, J. 2013, *MNRAS*, 434, 1469
- Barton, E. J., Chiu, C., Golpayegani, S., et al. 2014, *MNRAS*, 442, 1821
- Caldas, A., Leconte, J., & Selsis, F. e. a. 2019, *A&A*, 623, A161
- Changeat, Q., Edwards, B., Waldmann, I. P., & Tinetti, G. 2019, *ApJ*, 886, 39
- Charnay, B., Meadows, V., & Leconte, J. 2015, *ApJ*, 813, 15
- Cowan, N. B., Greene, T., Angerhausen, D., et al. 2015, *PASP*, 127, 311
- Drummond, B., Tremblin, P., Baraffe, I., et al. 2016, *A&A*, 594, A69
- Edwards, B., Changeat, Q., Baeyens, R., et al. 2020, *AJ*, 160, 8
- Evans, T. M., Sing, D. K., Goyal, J. M., et al. 2018, *AJ*, 156, 283
- Falco, A., Zingales, T., Pluriel, W., & Leconte, J. 2022, *A&A*, 658, A41
- Feng, Y. K., Robinson, T. D., Fortney, J. J., et al. 2018, *AJ*, 155, 200
- Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, *ApJ*, 678, 1419
- Goody, R. M., & Yung, Y. L. 1989, *Atmospheric Radiation : Theoretical Basis* (Oxford: Oxford University Press)
- Greene, T. P., Line, M. R., Montero, C., et al. 2016, *ApJ*, 817, 17
- Guerlet, S., Spiga, A., Sylvestre, M., et al. 2014, *Icarus*, 238, 110
- Heng, K., Menou, K., & Phillipps, P. J. 2011, *MNRAS*, 413, 2380
- Kataria, T., Showman, A. P., Lewis, N. K., et al. 2013, *ApJ*, 767, 76
- Kataria, T., Sing, D. K., Lewis, N. K., et al. 2016, *ApJ*, 821, 9
- Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007, *Nature*, 447, 183
- Lacy, B. I., & Burrows, A. 2020, *ApJ*, 905, 131
- Leconte, J., Forget, F., Charnay, B., et al. 2013, *A&A*, 554, A69
- Line, M. R., & Parmentier, V. 2016, *ApJ*, 820, 78
- Lothringer, J. D., Barman, T., & Koskinen, T. 2018, *ApJ*, 866, 27
- MacDonald, R. J., Goyal, J. M., & Lewis, N. K. 2020, *ApJ*, 893, L43
- Marley, M. S., & McKay, C. P. 1999, *Icarus*, 138, 268
- Menou, K., & Rauscher, E. 2009, *ApJ*, 700, 887
- Parmentier, V., Guillot, T., Fortney, J. J., & Marley, M. S. 2015, *A&A*, 574, A35
- Parmentier, V., Fortney, J. J., Showman, A. P., Morley, C., & Marley, M. S. 2016, *ApJ*, 828, 22
- Parmentier, V., Line, M. R., Bean, J. L., et al. 2018, *A&A*, 617, A110
- Parmentier, V., Showman, A. P., & Fortney, J. J. 2021, *MNRAS*, 501, 78
- Pluriel, W., Whiteford, N., Edwards, B., et al. 2020a, *AJ*, 160, 112
- Pluriel, W., Zingales, T., Leconte, J., & Parmentier, V. 2020b, *A&A*, 636, A66
- Powell, D., Loudon, T., Kreidberg, L., et al. 2019, *ApJ*, 887, 170
- Redfield, S., Endl, M., Cochran, W. D., & Koesterke, L. 2008, *ApJ*, 673, L87
- Showman, A. P., & Guillot, T. 2002, *A&A*, 385, 166
- Showman, A. P., Cooper, C. S., Fortney, J. J., & Marley, M. S. 2008a, *ApJ*, 682, 559
- Showman, A. P., Menou, K., & Cho, J. Y. K. 2008b, *ASP Conf. Ser.*, 398, 419
- Showman, A. P., Fortney, J. J., Lian, Y., et al. 2009, *ApJ*, 699, 564
- Showman, A. P., Wordsworth, R. D., Merlis, T. M., & Kaspi, Y. 2013, *Atmospheric Circulation of Terrestrial Exoplanets*, eds. S. J. Mackwell, A. A. Simon-Miller, J. W. Harder, & M. A. Bullock (Tucson: University of Arizona Press), 277
- Skaf, N., Bieger, M. F., Edwards, B., et al. 2020, *AJ*, 160, 109
- Spiegel, D. S., Silverio, K., & Burrows, A. 2009, *ApJ*, 699, 1487-500
- Stevenson, K. B., Lewis, N. K., & Jacob L. Bean, e. a. 2016, *PASP*, 128, 094401
- Tan, X., & Komacek, T. D. 2019, *ApJ*, 886, 26
- Tennyson, J., & Yurchenko, S. N. 2012, *MNRAS*, 425, 21
- Tinetti, G., Vidal-Madjar, A., Liang, M.-C., et al. 2007, *Nature*, 448, 169
- Tinetti, G., Eccleston, P., Haswell, C., et al. 2021, *ArXiv e-prints* [arXiv:2104.04824]
- Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, *AJ*, 155, 156
- Venot, O., Agúndez, M., Selsis, F., Tessenyi, M., & Iro, N. 2014, *A&A*, 562, A51
- Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, *ApJ*, 802, 107
- Wordsworth, R. D., Forget, F., Selsis, F., et al. 2011, *ApJ*, 733, L48
- Yurchenko, S. N., Barber, R. J., & Tennyson, J. 2011, *MNRAS*, 413, 1828
- Yurchenko, S. N., Tennyson, J., & Bailey, J. e. a. 2014, *Proc. Natl. Acad. Sci.*, 111, 9379

### Appendix A: Nested-sampling posteriors

We present here one nested-sampling posterior for the  $T_{\text{eq}} = 2100$  K case with a thermal inversion in the GCM that has H<sub>2</sub>O, CO, TiO and VO as absorbers in the atmosphere. This is a typical corner plot from our study. As we did 28 retrievals, it would not have been relevant to show every nested-sampling posterior in the paper. We therefore focus only on the mean value retrieved within  $1\sigma$  error in Figs. 7 and 9.

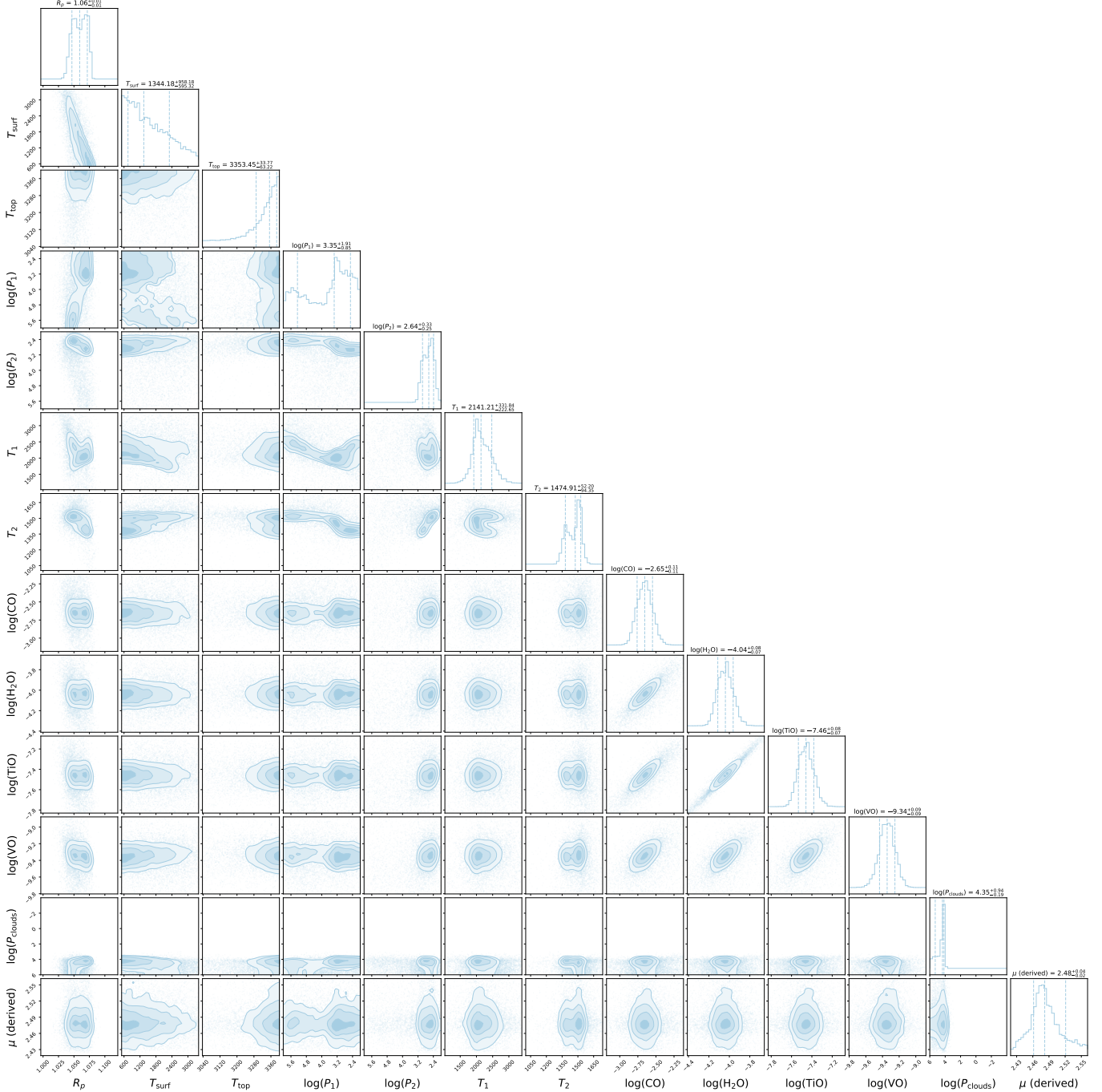


Fig. A.1: Nested-sampling posteriors retrieval from TauREx for the thermal inversion simulation considering TiO and VO as absorbers in the atmosphere. We retrieved 12 free parameters, which are H<sub>2</sub>O, CO, TiO, and VO abundances in log<sub>10</sub>(VMR), the cloud pressure in Pa, four temperature points in Kelvin, two pressure points (in Pa), and the planetary radius in Jupiter’s radius. The molecular weight is derived from these parameters.  $T_1$  and  $T_2$  are the temperature points corresponding to the pressure points  $P_1$  and  $P_2$ , respectively.