Peer Review File

Manuscript Title: Thermal Emission from the Earth-sized Exoplanet TRAPPIST-1 b using JWST

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

In this manuscript, the authors present the first results of the -- much awaited -- JWST observation campaign of the TRAPPIST-1 planets. Specifically, they present the first detection of thermal emission from TRAPPIST-1b, a planet with dimensions similar to the Earth, and which has a moderate insolation (halfway between that of Mercury and Venus). The performances of the MIRI instrument are truly amazing with a signal to noise ratio quite close to the photon noise limit.

The detection suggests very strongly that TRAPPIST-1b has no atmosphere (and at least, no CO2, which was the most likely atmospheric composition for the planet) and a very low albedo (close to zero).

This has potentially strong implications for the study of terrestrial planets around M stars (i.e., the category of Earth-like planets most easily accessible to our telescopes), because this would be a strong indication that a large part of these planets have no atmosphere at all.

The manuscript is very clear, concise and timely. It is of immediate interest to a very large fraction of scientists working in the field of exoplanetary sciences, planetary sciences, astrobiology, etc.

I strongly support the publication of the manuscript in Nature. I do have a list of mostly minor comments (listed below) that I would like the authors to address:

- l285 : "However, atmospheres dominated by CO2 or O2 (outgassed or desiccated) with surface pressures as little as 10 bar or less should have significant absorptions in the F1500W filter bandpass, reducing the observed brightness temperature to 300 K or less when accompanied by efficient heat redistribution over the planet (3; 19)."

>> It seems to me that the reason for the "O2-atmospheres" to produce such low brightness temperatures in the F1500W filter bandpass is that they contain some CO2. I think we might see it in your Figure 3 (comparing O2 and CO2 curves). I am thus not really convinced that you are really excluding O2 atmospheres, because do all O2-dominated atmospheres necessarily contain CO2?

That being said, I find it interesting that even low CO2 mixing ratios (which is the case – I guess – for the O2-dominated case) can be excluded using your measurement. I think this could be emphasized a bit more in the manuscript, especially given CO2 was one of the favorite candidates for the atmosphere compositions of TRAPPIST-1 planets (Lincowski et al. 2018, Turbet et al. 2020).

The fact that CO2, the most likely molecule a priori for the composition of the atmosphere of TRAPPIST-1b (if it has an atmosphere), is excluded, makes even more convincing the scenario in which TRAPPIST-1b has no atmosphere at all.

ps: I have a similar remark on the caption of Fig. 2; and on l374-376 as well.

- Fig. 2:

The increase in brightness temperature from the 0 albedo blackbody (orange) to the stellar+internal (red) seems really strong. Is it only due to the addition of internal heating? It seems that such increase in brightness temperature would require an increase in the heat flux of about 20-30%, which seems really huge given the level of tidal heating typically expected on the planet (Turbet et al. 2018, Barr et al. 2018). Could you please clarify?

- A more general question: Did you compare the IR stellar emission in the F1500W to predicted synthetic spectra, e.g. from the Phoenix database? Does it match well? Also, did you notice some temporal variability? Could it bring some constraints on stellar contamination (which seems to be a challenge for transmission spectroscopy for TRAPPIST-1)?

- l328 : "We converted this mean isotropic surface temperature to a dayside temperature Tday = 534 K using the geometric and thermal conventions for zero redistribution of stellar heating to the night side (35)"

>> What do you mean by "geometric and thermal conventions"? If you assume a negligible atmosphere, you could simply write that the surface flux at equilibrium is exchanged on an area 2x smaller than the total area (Area_day = A_global_isotrope /2), which means that due to Planck's law, 0.5*T_day^4 + 0.5*T_night^4 = (T_global_isotrope)^4. With T_night = 0, you have T_day increased by a factor 2^(1/4). Is it what you have done here? Or are you doing something more sophisticated? Could you please clarify?

Also, what do you mean exactly by dayside temperature? I believe there might the implicit assumption in your calculation that the dayside has a uniform temperature, but this is not the case if there is a negligible atmosphere due to the distribution of the incident flux depending on cos(zenith_angle). In that case, the dayside surface temperature should be significantly lower than the 534K value due to the Rogers-Hölder inequality. See for instance equation 3 in Leconte et al. 2013 (https://ui.adsabs.harvard.edu/abs/2013A%26A...554A..69L/abstract).

Similarly, at l270, you state that you "compute Td = 508K". Is it a value that assume a uniform dayside temperature? Or is it a mean calculated from the average of dayside surface temperatures (weighted by the cosine of the zenith angle?)?

- l75: an age of 7.6 Gyr : I suggest to add error bars, which are provided in the citer paper.

- l91: I suggest to also cite Bolmont et al. 2017 (https://ui.adsabs.harvard.edu/abs/2017MNRAS.464.3728B/abstract) here.

- l150 : ref 5 is cited twice.

- l302-315:
You could maybe state a bit more explicitly that your MIRI measurements still leave some room for a thin atmosphere with little heat redistribution?

- l169-174 "Spitzer Space Telescope observations did not detect the secondary eclipse of TRAPPIST-1b. However, the measurement precision of the Spitzer data constrained the dayside temperature of TRAPPIST-1 b to below 611 K (3 σ upper limit (15))"

>> I believe ref15 should be replaced by ref31?

Referee #2 (Remarks to the Author):

This paper is a landmark first detection of thermal emission from a terrestrial planet with a temperature of just 500 K (roughly 500 K cooler than the previous record holder, LHS 3844b). The result - a dayside temperature consistent with a bare rock - is of significant interest to the broader astronomical community. The TRAPPIST-1 planets are some of the top targets for rocky planet atmosphere characterisation in the foreseeable future, so the first JWST observations of this system are hotly anticipated, and whether they have atmospheres or not is one of the biggest looming questions in the field. In light of this, I strongly recommend the paper for publication in Nature.

The authors have done a careful job with the data analysis. The modelling is somewhat minimal, but an appropriate level of detail for a first paper (there will surely be a lot of follow-up work). With that, I do have a few minor comments and suggestions that would improve the manuscript.

Modeling
—
Line 278 - rocky surface models actually show a substantial amount of variation in brightness temperature (Hu et al. 2012). Based on these data it may actually be possible to distinguish between some surface compositions. I understand that this may be beyond the scope of this work but it is worth at least a sentence or two of comment.

Paragraph beginning on line 302 - how exactly were the model predictions for the grey atmospheres calculated? Were these calculated analytically based on the scaling relation in Koll 2022? Please add more detail here.

It would be really interesting to also compare the data to a very thin CO_2 atmosphere (P ~ 0.01 bar, like Mars). The models in the paper clearly show that a thick CO_2 atmosphere is ruled out, but suggest that thin grey atmospheres are consistent. I suspect that a thin CO_2 atmosphere would actually also be ruled out, which is a very powerful result worth showing.

Figures
—

Figs 2 and 3 present almost identical information (expect in different units). I find Fig 3 much more informative about which models can be ruled in /out. My recommendation would be to clean up Fig 3 and make this the main "takeaway" plot from the paper, and turn Fig 2 into a table (with columns for the model description, predicted brightness temperature, and discrepancy in sigma between the model and data).

Alternatively if the authors really want to keep Fig 2, I recommend changing the colour scheme so that there aren't multiple bars with the same color. I also suggest avoiding red/green bars to make the figure legible for color-blind people. Add additional dashed lines (maybe in lighter grey) indicating the two- and three-sigma bounds on the eclipse depth - this will help readers identify which models are confidently ruled out.

Data analysis
—

Please comment on whether the data showed any evidence for time-correlated red noise (looks like no, but would be nice to state this explicitly)

For the fits to the individual eclipse observations, please report the measured eclipse times. Those would be useful for future analyses of the TTVs in the system.

Author Rebuttals to Initial Comments:

Referee #1

>> It seems to me that the reason for the “O2-atmospheres” to produce such low brightness temperatures in the F1500W filter bandpass is that they contain some CO2. I think we might see it in your Figure 3 (comparing O2 and CO2 curves). I am thus not really convinced that you are really excluding O2 atmospheres, because do all O2-dominated atmospheres necessarily contain CO2?

Yes, the “O2-dominated” atmospheres of Linkowski et al. (2018) referred to in the manuscript do contain some CO2 (0.5 bar). We have changed “atmospheres dominated by CO$_2$ or O$_2$” to “atmospheres dominated by CO$_2$ or O$_2$ with some CO$_2$” to reflect this.

That being said, I find it interesting that even low CO2 mixing ratios (which is the case – I guess – for the O2-dominated case) can be excluded using your measurement. I think this could be emphasized a bit more in the manuscript, especially given CO2 was one of the favorite candidates for the atmosphere compositions of TRAPPIST-1 planets (Lincowski et al. 2018, Turbet et al. 2020).

The fact that CO2, the most likely molecule a priori for the composition of the atmosphere of TRAPPIST-1b (if it has an atmosphere), is excluded, makes even more convincing the scenario in which TRAPPIST-1b has no atmosphere at all.

- Fig. 2:
The increase in brightness temperature from the 0 albedo blackbody (orange) to the stellar+internal (red) seems really strong. Is it only due to the addition of internal heating? It seems that such increase in brightness temperature would require an increase in the heat flux of about 20-30%, which seems really huge given the level of tidal heating typically expected on the planet (Turbet et al. 2018, Barr et al. 2018). Could you please clarify?

The stellar+internal brightness temperature was computed from the median of the 50 hottest PACMAN models run by Krissansen-Totton & Fortney (2022) that result in no or minimal atmospheres (as stated in the text). This is only about 10% of all models that result in no or minimal atmospheres, and the resultant brightness temperature should be regarded as an upper limit for the case of stellar plus internal heating with no redistribution.

We have added this sentence to make this point: “This should be regarded as the maximum T$_{\rm day}$ for this particular internal heating model given that it is derived from only the hottest 10\% of model runs that resulted in negligible atmospheres.”

- A more general question: Did you compare the IR stellar emission in the F1500W to predicted synthetic spectra, e.g. from the Phoenix database? Does it match well? Also, did you notice some temporal variability? Could it bring some constraints on stellar contamination (which seems to be a challenge for transmission spectroscopy for TRAPPIST-1)?

Yes, we found that the measured flux in the MIRI F1500W filter was 13% lower that expected from comparison to a Phoenix model that was normalized to the star's 2MASS flux (J = 11.35 mag). The specific model was made using the astropy pysynphot task to synthesize a model from the BT Settl grid (used for T < 2600 K) for the stellar parameters given in Agol et al. (2021).

We added a sentence explaining this in the "Calibrated Stellar Flux" subsection of Methods: "This is 13% lower than than the flux expected from a Phoenix / BT-Settl model produced by the PySynphot package for the stellar parameters and normalized to the star's observed 2MASS J = 11.35 mag flux."

We measured less than 0.4% absolute brightness changes throughout the 5 observations, which is already stated in the Calibrated Stellar Flux" subsection of Methods. Much of this variation is likely a calibration artifact due to the star being placed at a slightly different detector location on every observation. The star has been found to be much more variable at shorter wavelengths.

- l328 : "We converted this mean isotropic surface temperature to a dayside temperature Tday = 534 K using the geometric and thermal conventions for zero redistribution of stellar heating to the night side (35)"
>> What do you mean by "geometric and thermal conventions"? If you assume a negligible atmosphere, you could simply write that the surface flux at equilibrium is exchanged on an area 2x smaller than the total area (Area_day = A_global_isotrope /2), which means that due to Planck's law, 0.5*T_day^4 + 0.5*T_night^4 = (T_global_isotrope)^4. With T_night = 0, you have T_day increased by a factor 2^(1/4). Is it what you have done here? Or are you doing something more sophisticated? Could you please clarify?

Our conversion from the isotropic model was slightly more involved than the above conversion that applies only for a constant temperature at all dayside locations. We started by separating out the internal heating and allowing it to be distributed uniformly over 4 pi str. We explain what we did for the re-radiated stellar flux in our response to the referee's next point.

Also, what do you mean exactly by dayside temperature? I believe there might the implicit assumption in your calculation that the dayside has a uniform temperature, but this is not the case if there is a negligible atmosphere due to the distribution of the incident flux depending on cos(zenith_angle). In that case, the dayside surface temperature should be significantly lower than the 534K value due to the Rogers-Hölder inequality. See for instance equation 3 in Leconte et al. 2013 (https://ui.adsabs.harvard.edu/abs/2013A%26A...554A..69L/abstract).

Next we calculated Tday, the *apparent* observed dayside temperature of this model. We assumed that the dayside absorbed by the star did have the above zenith angle dependence with no heat redistribution, so the surface planet temperature was no longer isothermal and was distributed as described in the cited Cowan & Agol (2011) reference (hottest at the substellar point). We computed Tday as the observed dayside brightness temperature that corresponds to radiating all internal heat uniformly over 4 pi str plus radiating all absorbed stellar flux over the same solid angle as for the case with only re-readiation of only stellar heating with zero redistribution (i.e., computed by Cowan & Agol (2011). The resultant Tday = 534 K should be compared to the effective Tday = 508 K for the no internal heat scenario (per Cowan & Agol 2011and Koll 2022), not the mean temperature of the dayside of the planet.

Stellar heating with no redistribution causes the planet to appear to have a higher effective temperature than the mean dayside because the hottest parts of the planet near the substellar point face the observer directly during secondary eclipse and the cooler regions project a smaller area on the two-dimensional planetary disk seen by the observer. For example, the flux-averaged mean dayside temperature for 2 pi str reradiation of stellar flux only is about 476 K vs. the predicted observed Tday = 508K. This apparent dayside temperature (508K for no internal heating and 534 K including internal heat) is what should be compared to the Tb brightness temperature measured from the observations.

We added some text to emphasize that the T_day = 534 K is an apparent observed value and was calculated using the same conventions used in defining Td above (Cowan & Agol 2011).

`Similarly, at l270, you state that you “compute Td = 508K”. Is it a value that assume a uniform dayside temperature? Or is it a mean calculated from the average of dayside surface temperatures (weighted by the cosine of the zenith angle?)?`

This is the *apparent observed* dayside temperature for a planet with no redistribution and heating being a function of stellar zenith angle. This apparent dayside temperature of the planet that should be compared to the observed Tb brightness temperature (Cowan & Agol 2011).
We added that this is the apparent observed dayside temperature and also added this explanatory sentence:
“The planet's substellar point has temperature T_0, and all other locations have lower temperatures and contribute less relative area to the observed planet flux in a secondary eclipse observation.”

`- l75: an age of 7.6 Gyr : I suggest to add error bars, which are provided in the citer paper.`

We have changed the age to 7.6 +/- 2.2 Gyr as suggested by the referee.

`- l91: I suggest to also cite Bolmont et al. 2017 (https://ui.adsabs.harvard.edu/abs/2017MNRAS.464.3728B/abstract) here.`

We added this reference as suggested.

- l150 : ref 5 is cited twice.

We have fixed this error and thank the referee for pointing this out.

- l302-315:

You could maybe state a bit more explicitly that your MIRI measurements still leave some room for a thin atmosphere with little heat redistribution?

The text stated "our measured dayside brightness temperature of TRAPPIST-1 b is consistent with an atmosphere having surface pressure P = 0.1 bar and total gray optical depth $\tau = 0.1$ (within 1 σ)."

We changed it to this:
"Table 1 shows that our observations are consistent with a thin atmosphere with little heat distribution; the flux for the surface pressure P = 0.1 bar and gray optical depth $\tau = 0.1$ model [25] is within 1 σ of our measurement."

- l169-174 "Spitzer Space Telescope observations did not detect the secondary eclipse of TRAPPIST-1b. However, the measurement precision of the Spitzer data constrained the dayside temperature of TRAPPIST-1 b to below 611 K (3 σ upper limit (15))"
>> I believe ref15 should be replaced by ref31?

Again we thank the referee for finding this error, and we have corrected it.

Referee #2 (Remarks to the Author):
Modeling
—
Line 278 - rocky surface models actually show a substantial amount of variation in brightness temperature (Hu et al. 2012). Based on these data it may actually be possible to distinguish between some surface compositions. I understand that this may be beyond the scope of this work but it is worth at least a sentence or two of comment.
(Hu 2012:
https://ui.adsabs.harvard.edu/abs/2012ApJ...752....7H/abstract)

We added this sentence:
"These and other materials could potentially be identified with low-resolution spectroscopic observations at other wavelengths."

Paragraph beginning on line 302 - how exactly were the model predictions for the grey atmospheres calculated? Were these calculated analytically based on the scaling relation in Koll 2022? Please add more detail here.

These model results were taken from Figure 2 or Koll 2022 and adjusted slightly (~2%) to account for slightly different Td due to using slightly different system parameters. We added model references to the new Table 1 and changed the text to make this more obvious:

"Table 1 shows that our observations are consistent with a thin atmosphere with little heat distribution; the flux for the surface pressure P = 0.1 bar and gray optical depth $\tau = 0.1$ model [25] is within 1 σ of our measurement."

```
It would be really interesting to also compare the data to a very thin
CO2 atmosphere (P ~ 0.01 bar, like Mars). The models in the paper
clearly show that a thick CO2 atmosphere is ruled out, but suggest that
thin grey atmospheres are consistent. I suspect that a thin CO2
atmosphere would actually also be ruled out, which is a very powerful
result worth showing.
```

This is indeed an interesting question. Mars's atmosphere has a high optical depth at 15 microns due to CO2 absorption, and radiation at this wavelength has a brightness temperature considerably lower than wavelengths outside of this feature (https://www.esa.int/ESA_Multimedia/Images/2007/12/Mars_-_thermal_radiation_spectra). Therefore, a model TRAPPIST-1 b Mars-like atmosphere would appear less bright and colder in the JWST MIRI F1500W filter than at adjacent wavelengths. We know of no existing model atmosphere like this for TRAPPIST-1 b, and one would have to be created and run before we could determine whether our observations could rule it out or not. We may get some information about whether there is some small amount of CO2 in the atmosphere from the upcoming observations in the F1280 filter. We have not edited the text to address this point since we have no relevant information at this time.

```
Figures
—
Figs 2 and 3 present almost identical information (expect in different
units). I find Fig 3 much more informative about which models can be
ruled in /out. My recommendation would be to clean up Fig 3 and make
this the main "takeaway" plot from the paper, and turn Fig 2 into a
table (with columns for the model description, predicted brightness
temperature, and discrepancy in sigma between the model and data).
```

We thank the referee for this suggestion and have converted Figure 2 to a table. We also modified the fonts and legend in Figure 3 to explain the curves better and make it more legible.

```
Alternatively if the authors really want to keep Fig 2, I recommend
changing the colour scheme so that there aren't multiple bars with the
same color. I also suggest avoiding red/green bars to make the figure
legible for color-blind people. Add additional dashed lines (maybe in
lighter grey) indicating the two- and three-sigma bounds on the eclipse
depth - this will help readers identify which models are confidently
ruled out.
```

We have converted Figure 2 to a table as suggested by the referee.

```
Data analysis
—
Please comment on whether the data showed any evidence for time-
```

correlated red noise (looks like no, but would be nice to state this explicitly)

We now note in Methods that only the eclipse from November 20th showed evidence for a small amount of unmodelled red noise in its Allan variance plot.

For the fits to the individual eclipse observations, please report the measured eclipse times. Those would be useful for future analyses of the TTVs in the system.

We thank the referee for this suggestion and have listed the mid-eclipse times for the five observations in the new Extended Data Table 1. We explain this in the penultimate paragraph of the “Individual Fits #1” subsection in Methods and note there that all times are within a factor of 2 of the uncertainties in the adopted timing model assuming zero eccentricity. We also note in the Joint Fit #1 subsection that the measured eclipse duration is consistent with the planet’s published transit duration.

Sincerely,