Response to the Community comment by Alexander Kutepov.

We are responding below to the comments by our colleague. His comments are listed in black with our responses in blue.

GENERAL COMMENTS

This is a large manuscript with very large numbers of plots both in main body and appendixes. Although it is well written and structured, many tests of a new routine repeat one another and, therefore, look excessive.

Reply. The point is accepted. The other two referees made similar comments. The number of figures in the main text and the Appendix will be substantially reduced (see reply to referee 1 for the details). The number of Figs. in the main text will be reduced from 28 to 19 (and some of them with a reduced number of panels); and the number of Figs. in the appendix will be reduced from 20 to 8. 21 figures will be moved to a Supplement.

Nevertheless, we would like to state clearly in the manuscript the goodness and limitations of the parameterization; hence all the tests will be retained but the corresponding figures moved to a Supplement.

The work takes us back to the late 70s - early 80s of the last century, when new for that time techniques of the approximate 15-micron CO2 cooling calculations (both for LTE and non-LTE conditions) were developed, see (Kutepov, 1978) and (Akmaev and Shved, 1982). The revised version of the Fomichev et al, 1998 (hereafter F98) routine described in the manuscript does not represent any innovation and does not suggest any new option for the GCM users and developers.

Reply. We recognize in the manuscript that this is based on the same approach of Fomichev et al, 1998. Our major aim was to extend it to higher CO2 vmrs and once we did it, to improve it also in other aspects, as detailed in the manuscript. We think it is a new option, as GCM models can now be run for higher CO2 vmrs and it is more accurate.

This statement may be explained as follows. Any physical parameterization for GCM must be able to react as realistic as possible at steady changing local physical state of the atmosphere in the modeling process. This is particularly true for the infra-red radiation and its effects (local cooling or heating), which are critically important for adequate modeling of energy balance. It is well known that instantaneous p-T distributions in modern GCMs of middle and upper atmosphere exhibit very strong variability, caused by the superposition of tidal and gravity waves of different amplitude and vertical scales. Same is true for the atmospheric p-T distributions measured in both ground and space experiments. Therefore, the parameterization of the 15-micron CO2 cooling, which is a main radiative cooling mechanism of these layers, must properly react to this variability.

However, the matrix parameterizations of the 15-micron cooling are unable to provide adequate reaction to strongly disturbed T distributions. This was well known already 30 years ago for those (I was among them) who worked on developing the first version of the F98 parameterization. Demonstration of large parameterization errors for wavy T profiles was not included in the paper by Fomichev et al, 1993, however, it contained at least the warning addressed to its users ***"It is recommended for the calculation of the radiative cooling for smooth temperature profiles, namely for profiles undisturbed by micro- and meso-scale motions."*** When the updated 1993 routine was released by Fomichev et al, 1998 nothing changed, the revised routine was still unable to treat wavy distributions, see (Kutepov and Feofilov, 2023, hereafter KF23). Again, as in the 1993 paper, the accuracy of F98 was demonstrated exceptionally for smooth T profiles. However, the authors dropped from the paper text the warning for its users cited above. Since that time the F98 routine has been widely used in GCMs of middle and upper atmosphere for any T distributions.

It is also worth noting here that 15-micron CO2 radiative cooling is strongly non-linear in respect to the temperature variations. Therefore, the zonal mean cooling, which is enthusiastically discussed in this manuscript, is not equal to that calculated for the zonal mean temperature. The authors, however, pay no attention to this fact!
Reply. Precisely for the reasons mentioned above, we tested the parameterizations with realistic pT profiles as those derived from MIPAS. Rather than giving a recommendation we assess its performance and give errors so the users are aware of them. One important point: The individual pT profiles do show a larger vertical structure because they are affected by the instrument noise. They are not smoothed profiles. For that reason, we wanted to show Fig. 20. We have included a couple of Figs. in the Reply to Referee 1 where we can see the large vertical variability of the tested pT profiles.
Let us clarify also the point about "zonal mean cooling" and the zonal mean temperature. We did NOT calculate the cooling for the zonal mean temperatures. We calculate the cooling for each individual p-T profile. Only when showing the differences between the "exact line-by-line" model cooling rates and those from the parameterization they were zonally averaged. For that reason, we also show the RMS. As Referee 1 also pointed out, we agree that we should not consider the mean as the error of the parameterisation. We will still provide the mean values (bias), which will inform us about its accuracy in a global sense (eg the cooling at high altitudes in the previous version is always about 2K/Day larger than in the line-by-line calculations), and by the RMS (not the standard deviation), as an appropriate estimate of its error for individual profiles. Those RMS values will be brought up in the abstract.
Also, note the study was done altitude by altitude, not averaging over altitudes. Of course, the parameterization is not perfect and one would wish to run the full non-LTE model for each p-T profile in each grid of the model for each run. Here we offer a fast parameterisation with reasonable (well-assessed "errors"). Let us the modeller choose what they prefer.

Now back to the current manuscript. The authors invested large efforts to update the F98 routine. They declare that the revised routine (hereafter L-P.23) allows calculations of 15-mucron cooling with higher accuracy for current CO2 vmr. It calculates with reasonable accuracy also the cooling for much higher CO2 concentrations. Nevertheless, again the accuracy of L-P.23 is demonstrated only for very smoothed "standard" T distributions.
Reply. We have shown both, the accuracy for the reference (smooth) atmospheres and also for the MIPAS pT atmospheres, which show many "wavy" profiles and very uncommon nearly isothermal profiles (see previous point and the Figs included in the reply to Referee 1).
As suggested by Referee 1, we will include also tests for lidar pT profiles.

Meanwhile, the authors are aware about large errors F98 routine has for disturbed T profiles. D. Marsh was the co-investigator of a recent NASA grant (Kutepov, 2021), where I was the PI. He and his team received funding for testing the KF23 algorithm for calculating 15-micron non-LTE cooling, comparisons of this routine with F98 parameterization, and for installing KF23 routine in the WACCM model. This study showed that F98 causes very large cooling errors (up to 25 K/day) on wavy profiles. These errors are discussed in KF23.
Reply. We agree that for very wavy individual profiles, the cooling rates might differ significantly from the "accurate" calculations in the upper mesosphere. Fig. 22 gives us the error when averaging over narrow latitude bands (remember they are the mean of the differences of the "accurate" and "parameterized" cooling rates (not the differences of the cooling rates calculated for the mean temperatures), showing that they could be significant (up to 2-3 K/day) as discussed in the text; and Fig. 24 shows the RMS, which near the mesopause they are very significant. Thus we are properly assessing the parameterization accuracy. As suggested by Referee 1, these values will be brought up in the abstract.

One point to be made is that there might be room for multiple non-LTE parameterizations and, for climate simulations, the errors introduced due to small-scale temperature variability may well be tolerated by modelling groups if it comes at a reduced numerical cost. The KF23 parameterization adds significant numerical cost in the running of a GCMs and therefore may not be suitable for centennial-scale ensemble simulations. In limited testing by the software engineer at NCAR it was not insignificant and the option to call it only every other timestep needed to be introduced to reduce the cost. We should point out that KF2023 is about 300 times slower than F98.

Meanwhile, the authors of the manuscript are quite honest when, describing the main motivation of their study, they write: ***"In our case we have the option of developing a completely new parameterization, to adapt other CO2 parameterizations (as those cited above), or to extent and improve the parameterization of Fomichev et al. (1998). Attending mainly to practical reasons of promptness, we opted for the later."*** The keyword here is "***promptness".***

This "***promptness"*** looks somewhat strange after 25 years of no interest of the authors to the problem of fast and accurate calculation of radiative cooling for the Earth's GCMs. Knowing about the drawbacks of F98 routine and other matrix parameterizations we spent these years developing accurate non-LTE radiative transfer techniques, which are free of these drawbacks and are fast enough to be applied in GCMs. The results of this long-term study are summarized in KF23.

Reply. The recent interest in improving/extending the parameterization, as you have done, is that a) climate simulations conducted for CMIP have only just begun to include models that resolve the stratosphere and mesosphere, and b) the standard climate sensitivity metric includes a 4xCO2 experiment which as incorporated into the DECK for CMIP6. This work paves the way for moving the F98 code to modern Fortran or C++. It is unlikely that, as modellers increase the model height of their model top for climate simulations, they would want to incorporate code that relies on the compatibility of compilers to 30+-year-old coding standards or does not cover the full CO2 range of the DECK.

Finally, what kind of new product these authors developed with the main motivation "to be prompt"?

Reply. To make available a fast and efficient algorithm capable of coping with the very large CO2 VMRs. Some GCM modellers urged us to extend it to large CO2 vmrs.

Again, I do not need my own judgement. It is enough to cite what the manuscript authors write when they describe large errors, much lagers than 0.5-1.0 K/day reported for standard profiles, which they observed for all profiles at latitudes northernmost of 50 N for a single non-standard situation with a pronounced elevated stratopause event: "***Both parameterizations underestimate the cooling in that atmospheric region. The new parameterization has, however, a better performance above about 80 km, but in the strat-warm/elevated stratopause region (80–100 km) it still underestimates the cooling by 3–7 K Day−1 (~10%)".***

Reply. We do not understand this kind of comment. We just did what we scientists are supposed to do, e.g., to recognize the limitations of the parameterization. These elevated stratopause situations are the most difficult situations to handle with this parameterization, and this one of 2009 in particular, which was very strong. However, globally, they are rather unusual and limited to certain regions and times. In any case, the parameterization, although not perfect, calculates the cooling rates with errors of ~10%.

This looks like a confession that in a non-standard situation the new l-P.23 routine works no better than F98.

Reply. It is better in about 2 K/Day above 5e-4 hPa and also around 4e-3 hPa (see Fig. 27 left panel); and true, it is a very unfavourable situation.

And further: ***"It seems clear that part of this underestimation is caused by the fact that such atypical temperature profiles (see Sec. 3.1) were not considered in the parameterization. However, its inclusion would not solve the problem as in the calculations of the coefficients a trade-off of the weighting of the different p -T reference atmospheres have to be chosen (see Secs. 5.1 and 5.2). Thus, it might ameliorate the inaccuracy for these elevated-stratopause events but would worsen the accuracy for other general situations. This manifests the difficulty/limitation of this method to provide accurate non-LTE cooling rates for all temperature structures (gradients) that we might find in the real atmosphere."***

I absolutely agree with this statement: this approach for parametrizing the 15-micron non-LTE cooling in the middle and upper atmospheric layers, which was applied in previous routine versions in 1993, in F98, and repeated in L-P.23 **is a deadened approach**.
Reply. We do not agree with this final statement. The fact that the parameterization is not very accurate (errors of ~10%) for certain atypical and unusual conditions should not be generalized for all conditions. As very often occurs, it is a trade-off between accuracy and efficiency.

In KF23 we discuss in detail its drawbacks. Briefly: it is impossible to adequately estimate the non-LTE cooling in the very variable atmosphere by dividing it in several altitude regions, where different techniques or expressions for cooling calculation (although linked in various ways) are used. Only exact algorithms, which rigorously describe the radiative energy exchange between various altitude layers and the non-LTE radiative field coupling with atmospheric heat reservoir may satisfy the current cooling accuracy requirements.
Reply. We think it depends on what we understand by "adequate estimate". We are giving conservative error estimates so the users can decide if it is (or not) adequate for his/her cooling rate accuracy requirements. We offer an option, not stating that this option is the best.

It is my opinion that this paper in its current shape must not be published. It does not provide any significant improvement compared to previous work(s).
Reply. We have no comment on the first part of the sentence. We will accept the editor's decision. Regarding the second sentence, we think it is already clear in the manuscript the advantages/improvements of the parameterization.

To be published the manuscript requires significant major revision:
(1) it must demonstrate how revised routine works for T profiles disturbed by the strong waves.
Reply. We will show a few examples of the MIPAS pT profiles, which are rather "wavy" (see, e.g. the figures in the reply to Referee 1, but will also make clear that they cannot be considered as representative of the parameterization performance. We think that a better estimate of its accuracy is given by the RMS (not the standard deviation) obtained from a statistically significant sample. In any case, referee 1 suggested showing the results for some kind of "lidar" pT profiles and they will be included in the revised version.

If the routine fails on the wavy profiles, but the authors still recommend it for further usage in GCMs, then (2) they need to justify that these errors have negligible or no effects on the GCM model results.
Reply. We think that it corresponds to the GCM modellers to decide if the estimated accuracy, given in Figs. 22 and 23, fulfil or not their accuracy requirements.

SPECIFIC COMMENTS
1. Why exact methods of the non-LTE CO2 cooling calculation cannot be applied in GCM?
The manuscript authors write: ***The computation of the cooling under those non-LTE conditions requires the solution of the radiative transfer equation (RTE) which is a non-local problem and requires a large amount of CPU time. Therefore, the solving the RTE in general circulation models (GCMs) or climate models that extend in height above the***

stratopause is impractical and efficient parameterizations of the CO2 infrared cooling have been developed and implemented in such models.

I disagree with this statement. It does not matter whether the LTE approach is applied or the non- LTE problem is considered, in both cases the calculation of cooling requires the solution of RTE, and is, therefore, the non-local problem.

Moreover, the computing costs of exact RTE solution in modern algorithms are not the main problem, which makes ***"the solving the RTE in general circulation models GCMs impractical".*** The authors mean here not just RTE but the entire non-LTE problem.

Reply. You are correct, even in LTE, it is still a non-local problem. About the second sentence, you are correct again, we meant the entire non-LTE problem. We will change the text accordingly.

Inversion of large matrices to get the populations in the developed in 1950s CM (Curtis, 1956) matrix algorithm, which the members of this team utilize since 1980s, is most computational costly part of the non-LTE problem solution. The authors either are not aware of or ignore dramatic progress in the developing the non-LTE techniques, see, for instance, (Hubeny and Mihalas, 2015) and (Frish, 2021). Large matrix inversion costs make CM technique usage impractical in GCMs. We discussed this in the papers by Kutepov et al, 1998, Gusev and Kutepov, 2003, and in more details in KF23.

Reply. The original algorithm was developed based on the Curtis matrix (CM) approach and so we used it here. Nevertheless, we recall that the parameterization does not invert CMs. Coefficients are computed based on pre-calculated CMs but these are not inverted within the parameterization. On the other hand, in the revised manuscript, we will cite the recent fast non-LTE model of KF23 that uses another non-LTE technique.

2. How the 15-micron cooling maximum in the lower thermosphere is formed?

In the manuscript: Above the mesopause, the cooling rate rapidly increases following the enhancement of the kinetic temperature. Above about 130 km, the cooling rates decline because of the depletion of the CO2 vmr (see Fig. 7).

Are the authors sure about this? I recall that in early publications of the 1980s about the 15-micron cooling it was demonstrated that cooling declines at higher altitudes even for constant CO2 vmr. Main problem at these altitudes is the rapid decrease of atmospheric pressure and the CO(v2) collisional quenching rate, which disconnects the 15-micron radiation from the atmospheric heat reservoir. If the authors do not agree with my comment, could they demonstrate that cooling stops decaying above 130 km when CO2 vmr is constant?

Reply. We should have said here "mainly" due to the CO2 vmr fall. As you well know, in this region we can assume the cool-to-space approximation, in which (see, e.g. Eq. 9.1 in Lopez-Puertas and Taylor (2001)) the cooling rate is proportional (when expressed in K/Day) to the CO2 vmr, to the [O] concentration (density) and to temperature through exp(-E/kT). As the altitude increases, the CO2 VMR decreases, and so does [O] (density) but T increases. So what we see at the end is the total effect of the three quantities. We will correct this in the revised version.

3. Why the CO2-O quenching rate coefficient ~ 6.0e-12 cm3 s −1 was used as the basic one for revision of the F98 parameterization?

The authors write that they used for updating the F98 parameterization the quenching rate coefficients which are very similar to those used for its development ***"... except the k CO2−O rate (process 1c in Table 1) that has been considered here with its upper limit. That is, about a factor of two larger than in the parameterization of Fomichev et al. (1998). This rate coefficient is not well known with uncertainties of the order of a factor of two (see, e.g., García-Comas et al., 2008). While laboratory measurements are in the range of 1.5 to 2·10−12 cm3 s −1 the values derived from atmospheric observations are close to 6·10−12 cm3 s −1. ...."***

First, measured in laboratory and retrieved from the atmospheric observation values of k differ not by a factor 2, but by 3-4. Additionally, if one accounts for the k retrievals by Feofilov et al, 2012, which involved the ground-based lidar temperature measurements, then this factor will be 3-6.
Reply. That is correct, taking the entire variation (equivalent to 3 sigma), we have roughly a change from 1.5 to 6, e.g. a factor of 4, which would be +/- a factor of 2 for a 3-sigma error. We believe the higher range derived by Feofilov et al, 2012 also accounts for the errors in the measured temperatures and an estimated uncertainty in [O], which was not measured; hence its larger uncertainty.

Table 1 shows that the authors selected k ~ 6.0e-12 cm3 s −1 for updating the parameterization. Although k is supposed to be an input parameter in both F98 and L-P.23 routines, however, the previous one was optimized in the transition region from LTE to non-LTE for k ~ 3.0e-12 cm3 s −1, whereas new one is optimized for 6.0e-12 cm3 s −1. The authors tell in manuscript that this causes additional differences between F98 and L-P.23 in the transition area even when both apply the same rate coefficient, and then explain the reason why they selected higher rate for optimizing L-P.23 as ***"... we have optimized it for the high value (see Table 1), as this value has been used in the most recent non-LTE retrievals of temperature from SABER and MIPAS measurements"***.
It is not enough, however, to say higher k was selected because this rate provides more reasonable T retrievals from space observation. As the authors know to validate these T measurements current GCMs apply twice lower rate 3.0e-12 cm3 s −1. If the authors recommend, which follows from the text, to use L-P.23 with the most high k, then can they demonstrate how this affects the GCM (for instance the WACCM model) runs compared to those with twice lower k? Does this provide better fitting of measured temperatures?
Reply. Of course, it is questionable which value of K to use to optimize the parameterization (e.g. ~3e-12 or ~6e-12, as the lower laboratory values (~1.5e-12) do not fit well with atmospheric measurements). We could have chosen 3e-12 instead of 6e-12. But we did not, as explained above and in the manuscript, because the two major temperature databases of the upper mesosphere/lower thermosphere, SABER and MIPAS, (covering several years, more than 20 years in the case of SABER) both use the larger rate. And the teams responsible for those retrievals made that decision based on the temperature validation performed for both datasets E.g., García-Comas et al (2023) have shown that they obtain a better agreement with temperature measurements of independent "non-LTE" free instruments when this large rate is used. You mentioned the work of Feofilov et al, 2012, who also retrieved from SABER and lidars temperature, and they obtained a k value of 6.5e-12, even larger than that used here.
The argument of using a lower value of ~3e-12 because that is being used in GCM models to reproduce the SABER and MIPAS T measurements we think is less substantiated and weaker. First, it is not clear that the models can reproduce the measured temperature field with that rate, see, e.g. Fig. 1 in Smith (2012). Secondly, the temperature computed by GCMs depends not only on this rate but also on many other factors like the parameterisation of the GWs, the chemistry (related mainly to O3 and O), vertical descent, etc. Further, they should also explain not only the temperature fields but also, for example, the CO2 and CO observations. A good example of the difficulty of e.g. WACCM in simulating those measured temperature fields can be found in this recent work:
https://essopenarchive.org/users/568957/articles/657910-a-novel-gravity-wave-transport-parametrization-for-global-chemistry-climate-models-description-and-validation.
In any case, the errors incurred by the parameterization if using the intermediate K rate are not that much larger. For this reason, we performed such an assessment.

4. Testing the parameterization for measured temperatures
The Fig.21 of manuscript shows an example of the MIPAS nighttime temperature profiles (15 February 2009) used for verifying the parameterization accuracy. The authors note large

variability of the measured temperature profiles. These individual profiles are the good inputs for the revised parameterization to show how it works for strongly disturbed T profiles.
The authors have obviously performed these tests, but they do not show these results. Instead, they write ***"The results are presented in Fig. 22 for the zonal mean of the differences for two days of solstice and two days of equinox conditions and in Fig. 23 as the global mean difference for all latitudes for each of the four individual days."*** Why only these mean values are shown? Obviously, the averaging smashes the errors obtained for individual profiles, for which we observed in our study of F98 parameterization presented in KF23 errors up to 25 K/day. Meanwhile these large errors in our study have generally concentrated in the altitude region around 90 km, exactly where the RMSs in Fig.25 of manuscript are maximized reaching 8-9 K/day. In our paper we explain why F98 works badly in this transition region. L-P.23 has the same problems and is nothing better.
Reply. We think we do appropriately present the results. On one hand, by the zonal mean of the differences as shown in Fig. 22, where the cooling rate differences are averaged over small latitude boxes (5º). And also by using the mean of the differences (to give an idea of "global" biases) and the RMSs. This is the standard procedure followed in the many validation studies we find in the literature when the sample is statistically significant.

5. The accuracy of L-P.23 for smooth T profiles
I wrote above that L-B.23 will not be any better than the F98 for disturbed T profiles. But how about the smoothed standard T profiles? The accuracy of cooling calculations with L-P.23 less than 0.5 K/Day for preindustrial CO_2 for smooth profiles is also questionable.

(1) The non-LTE model, which is used in this study for the reference calculations to optimize L-P.23 and then to check its accuracy, includes only 18 15-micron bands. From my point of view this model itself is not accurate. In the routine we suggest in KF23, which utilizes the exact solution of the non-LTE problem, we use the same bands to calculate nighttime cooling for 400 ppm of CO_2 with an error not higher than 1 K/day for any T profile, including disturbed by strong waves. For smooth profiles this error reduces to ~ 0.2 K/Day. These errors were estimated by comparing KF23 routine with our reference model, which comprises 60 vibration levels of 5 CO_2 isotopic species and hundreds of bands. So, the exact algorithm compared to the exact algorithm showed 0.2 K/day error when a reduced set of bands was used. I doubt that the error of 0.5 K/day the authors report for the L-P.23 routine for smooth profiles after comparing it with a very simplified "reference model" is true. It must be higher.
Reply. Our reference non-LTE model is not simplified at all. We just dropped the contributions of bands beyond the listed 16 bands because we tested that their contribution in the non-LTE region is negligible (below 0.1 K/Day). Some of the not-included very weak bands may have some very small contributions in the LTE region, e.g. around the stratopause and below, but those contributions are not significant for the non-LTE region.

(2) The authors write that contribution of the heating due to the absorption of solar radiation in the near infrared CO_2 bands at day time is negligible compared to the 15-micron cooling. However, 2-3 K/day (for current CO_2) do not look negligible for the routine, which accuracy is declared to be about ~0.5 K/day. We tested in detail the (Ogibalov and Fomichev, 2003) parameterization of this heating, found this warming increasing for some wavy T profiles, and made sure this parameterization cannot guarantee the error of the KF23 routine at daytime to be lower than 1 K/day. As a result, we extended our KF23 daytime model up to 26 CO_2 vibrational levels and 56 bands to satisfy this requirement.
Reply. It seems there is a misunderstanding here. We are NOT neglecting the cooling by the CO_2 15 μm bands during the daytime. It is just that their contributions during the daytime are accounted for by the NIR heating routine. Hence, if included in this parameterization it would be included twice.
Just one precision. Our day/night differences of the CO_2 15 μm cooling rates are smaller than 1K/day for all pT profiles, at any altitude and for CO_2 vmrs up to 5 times the pre-industrial

value, except for MLS and SAS pT profiles near 105 km for CO2 vmr 4x and 5x the pre-industrial values.

(3) The authors say nothing about how they account for the cooling effect of the micro-scale sub-grid T disturbances. Kutepov et al., 2007 and Kutepov et al., 2013 showed that these temperature fluctuations cause near the mesopause an additional cooling up to 3 K day-1. I draw the authors' attention to the results shown by Kutepov et al, 2013 in Fig.1.5. It demonstrates one of the runs of the Leibniz Institute Middle Atmosphere (LIMA) model with the 15-micron cooling modified to account for the sub-grid T disturbances. It is shown that very minor variation of cooling (not higher than 2-3 K/day) lead to significant changes of the monthly and zonal mean temperatures for July 2005.
I am sure that errors of L-P.23 routine are much higher than 2-3 K/day and can be hardly reduced due to the deficiencies of the methodology applied. These errors will obviously have a strong impact on the GCM results.
Reply. Effectively, we do not include (e.g. do not provide a routine) the cooling rates induced by thermal structure at a grid smaller than the input grid (or, more properly, than the internal grid of the parameterisation). To properly account for them it would be necessary to know (or make assumptions) about how the temperature varies between grid points. We assume that the cooling induced by non-resolved GWs, propagating with a vertical wavelength of the order of/or smaller than the parameterisation grid, would be taken into account in the GCMs by using an appropriate GW parameterisation (eg. see Intro of Kutepov et al, 2013). Currently, some new parameterizations are being developed to account for these effects (see, https://essopenarchive.org/users/568957/articles/657910-a-novel-gravity-wave-transport-parametrization-for-global-chemistry-climate-models-description-and-validation).

6. The code availability
It seems the manuscript was submitted as the GMD "Development and technical paper". If it is correct, then **"The code should be made available, and a model availability paragraph must be included"**. The code is, however, not available.
Once the code is available, I will demonstrate that its errors are much higher than reported in the manuscript.
Reply: We were not aware of that and thought of providing it during the review process. We will provide it in its final version as a Fortran 90 code jointly with the revised version of the manuscript. In the meantime, the parameterization is now available provisionally as a Python routine, please see the reply to the Editor comment 1 (https://zenodo.org/doi/10.5281/zenodo.10547026).
That version of Python is very slow. We will provide the computational cost when translated into Fortran.
A model availability paragraph will be included in the section about the code availability (see Reply to the Editor-in-Chief).

References

García-Comas, M., Funke, B., López-Puertas, M., Glatthor, N., Grabowski, U., Kellmann, S., Kiefer, M., Linden, A., Martínez-Mondéjar, B., Stiller, G. P., & von Clarmann, T. (2023). Version 8 IMK–IAA MIPAS temperatures from 12–15-μm spectra: Middle and Upper Atmosphere modes. *Atmospheric Measurement Techniques*, *16*(21), 5357–5386. https://doi.org/10.5194/amt-16-5357-2023

Smith, A. K. (2012). Global Dynamics of the MLT. *Surveys in Geophysics*, *33*(6), 1177–1230. https://doi.org/10.1007/s10712-012-9196-9

https://essopenarchive.org/users/568957/articles/657910-a-novel-gravity-wave-transport-parametrization-for-global-chemistry-climate-models-description-and-validation).