Response to Review Comment 5 (RC5)

Manuscript: egusphere-2023-2129
Title: A dynamic approach to three-dimensional radiative transfer in numerical weather prediction models: the dynamic TenStream solver v1.0
Authors: Richard Maier, Fabian Jakub, Claudia Emde, Mihail Manev, Aiko Voigt, and Bernhard Mayer

We thank Anonymous Referee #4 for his or her comments on our manuscript, which we will respond to below. To structure our response, the referee's comments are printed on a gray background color, while our answers are displayed on ordinary white background.

Summary

This paper describes an updated version of the TenStream solver, which can be used to solve radiation in high-resolution numerical models such as atmospheric Large-Eddy Models. This new "dynamic" version represents an improvement in terms of computational speed compared to the original TenStream. It relies on the same radiative transfer model but its resolution is accelerated using two fundamental ideas. This first one is that previously computed radiation fields can be used as a first guess in the numerical resolution of the linear system corresponding to the TenStream model, which is refered to as a "dynamic" approach or "time-stepping" scheme because of the similarity with the resolution of advection in the dynamical core of atmospheric models. The second idea is that using an iterative method, namely, the Gauss-Seidel method, to solve the linear system starting from this first guess offers the possibility to stop the calculation after a few iterations, using the resulting field even if it has not converged toward the solution. This is refered to as "incomplete solve". After exposing these ideas and describing their implementation in the dynamical TenStream solver, the authors examine the errors introduced by the fact that infrequent calls to radiation will lead to starting from a "bad" first guess, increasing the error associated with incomplete solves compared to more frequent calls, for the same number of Gauss-Seidel iterations; as well as errors introduced by the fact that the solves are incomplete, by comparing their results with those predicted by the full TenStream solver given the same input fields. Their conclusions are that the dynamic TenStream is significantly faster than the original TenStream, while being mostly as accurate even using as few as 2 Gauss-Seidel iterations at each radiation call.

General comments

I find the paper of great interest. It reports important advances in the field of 3D radiation modeling and its numerical resolution, working towards replacing overly simplified and strongly biased 1D radiation models by their 3D counterparts. I find the manuscript very clear and well organized, and the demonstration of the capabilities of the dynamical TenStream solver convincing. I appreciated the detailed explanations on the models and evaluation methods. I found the part where the results are discussed a little less satisfying but I understand that much more work might be needed to really understand the biases of the different models and that it probably falls out of scope of the present study.

In the following I list some questions and suggestions that I think would make the manuscript even clearer. They are given in a chronological manner rather than per importance. I trust the authors' judgement in the relevance of my suggestions and questions and would recommend publication even if not all my comments are addressed in the revised version.

Specific comments

- Mostly in the Abstract and Introduction but also elsewhere in the paper: the distinction between sub-grid and inter-column "3D effects" is not clear enough and I am afraid it might be confusing for a non-expert reader. For instance in the Introduction L.30-33, it is mentioned that NWP models still use 1D ICA RT schemes, by which I think you mean "solve radiation independently in each model column". Immediatly after this statement comes "such as the McICA" which is indeed a 1D RT solver but here the ICA refers to the neglect of *subgrid* 3D effects (that is, between stochastically generated 1D profiles or "subcolumns"). Later on, you describe SPARTACUS, which is of a very different nature from the TenStream and NCA models, and only there the distinction between inter-column and sub-grid 3D effects is mentioned. I suggest you clarify since the beginning of the Introduction that this distinction exists and that your work relies to the resolution of inter-column horizontal transport. I also feel this distinction is lacking when you write that the 3D effects are becoming more important as the horizontal resolution of NWP models increases. I would rather say that the partition between subgrid and inter-column 3D effects depends on the host model horizontal resolution and that, as we go toward higher resolution, it becomes more important to solve horizontal transfers between columns and less so at the subgrid scale.

Thank you for pointing this out. Subgrid and inter-column 3D effects were indeed not clearly separated in our paper. To account for this differentiation, we have modified the introduction as follows:

“Depending on scale, we can differentiate between two different regimes of 3D radiative transport: On the model grid scale, 3D radiative transfer allows for horizontal transport of energy in between different model columns, whereas on the subgrid scale, it refers to the 3D transport of radiative energy within a heterogeneous model grid box. The calculation of both of these effects is computationally expensive, largely preventing their representation in operational weather forecasting. This is why up to this date, numerical weather prediction (NWP) models still use one-dimensional (1D) independent column approximations (ICA), such as the Monte Carlo Independent Column Approximation (McICA; Pincus et al. (2003)) currently employed at both DWD and ECMWF (DWD, 2021; Hogan and Bozzo, 2018). These models assume that radiative transport between grid boxes only takes place in the vertical and neglect any horizontal transport of energy – both in between different model columns and within individual model grid boxes.

In this paper, we will focus on 3D radiative transfer that allows for transport of radiative energy in between different model columns. While the approximation of neglecting this transport worked reasonably well in the past given the computational power and model resolution at that time, the increasing horizontal resolution of numerical weather prediction models makes inter-column 3D radiative effects more and more important (O’Hirok and Gautier, 2005). [...]”

We also clarified that our new solver is specifically designed for considering inter-column 3D radiative effects on the subkilometer-scale:

“To address this high computational cost of current 3D solvers, we present a first step towards a new, "dynamic" 3D radiative transfer model that is based on the TenStream solver. Currently designed for the use at subkilometer-scale horizontal resolutions, this new, fully three-dimensional solver accelerates inter-column 3D radiative transfer towards the speed of currently employed 1D solvers by utilizing two main concepts.”

Furthermore, we have clarified which 3D radiative effect we refer to in various parts of the paper.

- One condition for the Dynamic TenStream to work is that the radiation field does not change too much between two radiation calls, so that the field used as first guess is already close enough to the solution that only a few iterations of the Gauss-Seidel algorithm are needed. It made me wonder if the radiation field was advected with the rest of the atmospheric fields so that it still matched an advected cloud field and the largest errors were mostly limited to cloud birth and death between two radiation time steps?

We have not investigated that, but we would assume that the general structure of the radiative field is indeed to a large part advected with the rest of the atmospheric fields, if the time step does not get too large in a sense that cloud birth and death, but also major changes in the structure of the clouds dominate the differences in the radiative field in between two radiation time steps.

But that is actually a very interesting aspect to investigate in the future, as one could choose a more intelligent first guess that already considers advection as a starting point of the incomplete solves . This might speed up convergence even more.

We have added this into the outlook of the revised version of the paper: “Additionally, we could think about an even more sophisticated first guess for the incomplete solves by advecting the radiative field with the rest of the atmospheric fields. As we assume that the radiative field does not totally change in between two different calls of the radiation model, such a first guess should already better account for the updated position of the clouds, so that the incomplete solves could primarily focus on correcting for the changed optical properties of the clouds, which could speed up convergence even more.”

- How are the thermal sources handled by the Gauss-Seidel method? I imagine they are calculated at the beginning of the iterations and somehow part of the first guess but could you explain how it works exactly? Maybe comment on the fact that B is absent from eq. (2)?

The thermal source terms are calculated right before starting with the Gauß-Seidel iterations. They are not part of the first guess, but calculated from scratch within the same routine that retrieves the TenStream coefficients from the corresponding look-up tables. Whenever this routine is called for a grid box, it calculates both the Planck emission and emissivities for every stream, the latter following the pattern lined out in l. 100 of the preprint. We added a sentence at the end of section 2.2.1 to clarify that: “The thermal source terms are not part of the first guess and have to be calculated from scratch before starting with the Gauß-Seidel algorithm.”

In Eq. (2), the thermal source term was indeed missing. We fixed that for the revised version. Thanks for pointing this out!

- In Fig. 3, I don't understand how the fluxes entering the domain at the borders would systematically be "updated right from the beginning"? From what I understand, if the BC are periodic for instance, then the incoming flux at the left-side wall would be updated only after the outgoing fluxes at the right-side wall have been calculated? In a parallelized Dynamic TenStream, the fluxes at the subdomain boundaries would only be updated at the end of the calculation as mentioned at L.201 and hence the incoming fluxes at the borders used at a given time would be the ones from the calculations at the previous radiation call?

You are perfectly right, the boundary conditions were not properly visualized in Fig. 3. Hence, we have updated Fig. 3 and its caption as follows:

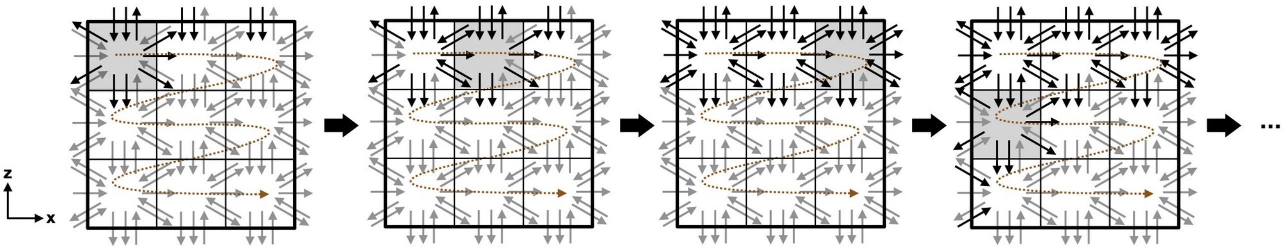


Figure 3: Two-dimensional schematic illustration of the first four steps of a Gauß-Seidel iteration, showing both diffuse and direct TenStream fluxes in case of Sun shining from the west or left-hand side. As one sequentially iterates through the grid boxes, ingoing fluxes are used to update the outgoing fluxes of the corresponding grid box (highlighted in grey). Grey arrows in contrast to black arrows indicate fluxes that have not yet been updated in this Gauß-Seidel iteration. Ingoing fluxes at the domain borders are dependent on the type of boundary conditions used. For this schematic, we applied periodic boundary conditions in the horizontal direction, while fluxes entering at the top of the domain are updated right from the beginning.

- L.154-156, solving for a clear-sky situation does not automatically imply that there is no horizontal variability in the model, e.g. specific humidity or surface albedo could still vary on the horizontal. In which case, shouldn't the spin-up be performed on the entire model grid? Would that still be manageable? Wouldn't it be cheaper to use the classical TenStream solver for initiating the Dynamic TenStream? At L.288, it is said that the classical TenStream is not used for initialization to avoid relying on PETSc library, could you elaborate a little more on that, and maybe mention it when the spin-up is first discussed in Sec. 2.2.2?

You are right, normally there can still be some horizontal variability in the background atmosphere in the absence of clouds. However, this background atmosphere is always one-dimensional in the libRadtran library, which allows us to perform the clear-sky spin-up for a single vertical column. We clarified that for the revised version of the paper: “Since this means there is no horizontal variability in the cloud field and libRadtran does not feature any horizontal variability in the background atmosphere, we can perform this calculation for a single vertical column at a dramatically increased speed compared to a calculation involving the entire model grid.”

In case the background atmosphere is not horizontally homogeneous, this 1D spin-up would certainly be less accurate, but still resemble a better starting point of the Gauß-Seidel algorithm than starting with values of zero for all the radiative fluxes. We also added that to the revised version: “Assigned to the radiative fluxes of all vertical columns in the entire domain, these values then provide a first guess for all the TenStream variables that can be assumed to be much closer to the final result than starting with values of zero – even if the background atmosphere was not horizontally homogeneous and we would have to take the average of that background first.”

For a better spin-up, one could in general of course also use a full TenStream solve. You are absolutely right that the reason for not using it should be given directly in Sec. 2.2.2, which we have done for the revised version, alongside with adding more background to that decision: “However, for the very first call of the radiation scheme, we cannot use a previously calculated result. In order to choose a reasonable starting point of the algorithm for this first call as well, though, we could use a full TenStream solve. However, such a solve would be computationally expensive and rely on numerical methods provided by the PETSc library, that we want to get rid of with our new solver to allow for easier integration into operational models. So instead of performing a full TenStream calculation, we decided to solve the TenStream linear equation system for a clear sky situation as a starting point.”

- L.254, "our solver does not yet take sub-grid scale cloud variability into account": any idea how you would do that? This is probably of great importance for NWP and without it the TenStream solver(s) might be restricted to LES where grid boxes might be considered homogeneous?

You are right, accounting for sub-grid scale cloud variability will possibly be the most important thing to consider when going to the NWP scale. To give a first idea of how we could do that, we extended the corresponding sentence in the outlook as follows: “Finally, going to the NWP scale, we will certainly need to consider sub-grid scale cloud variability, for example by extending the TenStream look-up tables to account for cloud fraction.” The implementation of these ideas however is beyond the scope of this paper.

- L.257 "to avoid problems with artificially low LWC at cloud edges [...]" were you able to quantify the error in the radiative field induced by smoothing the cloud field vs. by subsampling it at a coarser resolution? Or could you cite a study demonstrating that one is better than the other?

No, we did actually not quantify this error. The motivation to just use every forth grid box instead of averaging the cloud fields was exactly the one given in the paper: We thought that it might be more wisely to just use data coming directly out of the LES runs instead of producing averages, where artificially low liquid water contents could lead to an underestimation of 3D radiative effects at cloud edges.

- L.426-428, I disagree with "the newly developed solver is able to almost perfectly reproduce the results of the original TenStream solver whenever called". Looking at Fig. 6b, after a few time steps it seems that the Dynamic TenStream for dtrad=30s line is always above the TenStream lines. Similarly, I disagree with "our new solver even performs better than the delta-Eddington solver at a calling frequency of 10 s when it is operated at a calling frequency of 30 s" at L.430-431. Looking at Fig.6b again, it seems that the errors associated with the Dynamic TenStream for dtrad=30s become larger than those associated with the delta-Eddington for dtrad=10s after around 8200 s.

You are right that the dynamic TenStream solver is not exactly reproducing the results of the original TenStream solver whenever called. We actually explicitly noted that in l. 424-426: “Looking closely, we can also see that for both lower calling frequencies, the MAE of the dynamic TenStream solver does not always match the errors obtained at a calling frequency of 10 s when updated.”. However, the phrase “almost perfectly” is certainly not appropriate. We thus changed the statement to “the newly developed solver is almost able to reproduce the results of the original TenStream solver whenever called” for the revised version.

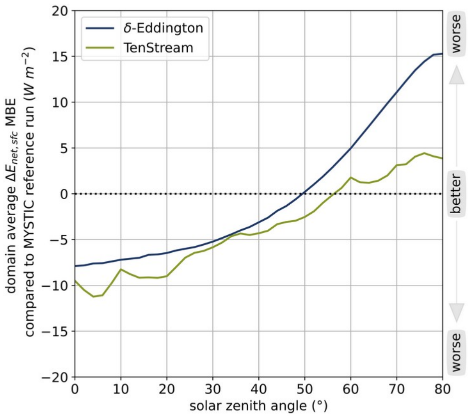
Apart from that, you are right that the maximum error caused by the dynamic TenStream solver at a calling frequency of 30 s exceeds the error of the delta-Eddington solver at a calling frequency of 10 s in the thermal spectral range (as does the original TenStream solver, by the way). To correct that, we have changed the meaning of the sentence to account for time-averages: “Looking at Fig. 6, we can now see that on time-average, our new solver even performs better than the δ-Eddington solver at a calling frequency of 10 s (bold blue line) when it is operated at a calling frequency of 30 s (bold red dash-dotted line) and thus with a similar computational demand as the 1D solver – both in the solar, as well as in the thermal spectral range.”

- Looking at Fig. 7b, it is interesting that the dynamic TenStream solver bias in the thermal partially compensates the original TenStream bias and it might not be for good reasons e.g. the original TenStream is not diffusive enough in the thermal and the incomplete solving in the dynamic approach adds numerical diffusion making the solution closer to the reference but for unphysical reasons?

Thank you for this suggestion that is definitely worth looking into. However, tests conducted with the original TenStream solver involving 24 instead of 10 diffuse streams to account for more diffusion did in general not reduce its bias. However, a sophisticated answer to this question would require a much deeper analysis of the two solvers that is beyond the scope of this paper.

- In Fig. 9a, it is also interesting that the mean bias is larger in the TenStream solvers than in the delta-Eddington. I think this might be very dependent on the solar zenith angle: 3D effects on the mean surface fluxes go from positive to negative as the sun goes from zenith to horizon and are usually close to zero for angles between 40 and 50 degrees from zenith in cumulus cloud fields (depending on cloud and surface properties). This is because the overestimation by 1D models of direct flux reaching the surface compensates the underestimation of diffuse almost perfectly at these angles. This solar angle dependence would not explain Fig. 9b though, but here the TenStream and delta-Eddington errors are of the same magnitude albeit of opposite sign.

Thank you for pointing this out. We actually had the same idea that the 3D effects in the domain-average net surface flux probably cancel at the zenith angle of 50° we are investigating. Initially, we have not evaluated the time series for different zenith angles, which is why we did not give an explanation for that in the paper. For the revised version, we investigated that in more detail. The following figure shows the mean bias error in the net surface irradiance as a function of the solar zenith angle for both the 1D delta-Eddington solver and the original TenStream solver, evaluated for the very first time step in our time series:



It basically confirms that the surface irradiance bias at a solar zenith angle of 50° that we used in our evaluation is very beneficial for the 1D solver, although the TenStream solver performs worse than the 1D delta-Eddington solver at all solar zenith angles below 50°. However, the difference in the MBE between the solvers is very small, as the absolute MBEs for solar zenith angles below 50° shown in Fig. 3 both result in relative mean bias errors of about -1% (not shown here). For higher solar zenith angles, we can however clearly see that the TenStream solver outperforms the 1D delta-Eddington solver.

- Even if the TenStream solvers clearly perform radically better than delta-Eddington, it is difficult to imagine how the remaining errors with respect to MYSTIC might affect the simulation once it is used online. Do you have any insights on that, from the literature maybe? For instance it is not obvious to me if it would be preferable to have the right mean flux but with the wrong spatial structure, or the opposite?

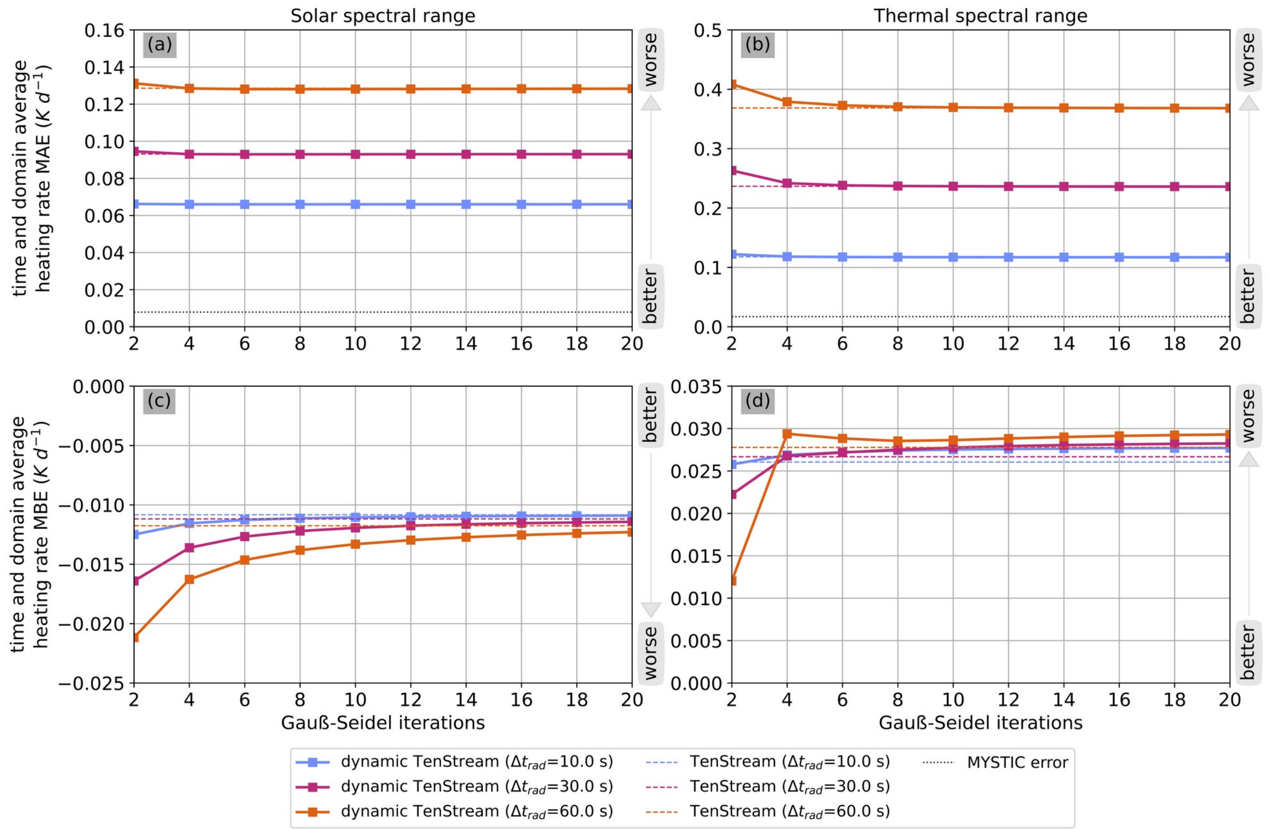
Currently, we do not really have any insights on how the errors introduced by both TenStream and the incomplete solves would affect simulations driven by our new solver. And although this topic is highly interesting for the future, it is somehow beyond the scope of this paper that was mainly focused on exploring first steps on whether incomplete solves could be an option to consider inter-column 3D radiative effects at much lower computational cost.

As it is a very important topic, though, we have included it into the outlook of the paper: "Coupled to dynamics, it will also be very interesting to investigate how the incomplete solves in the dynamic TenStream solver influence the development of clouds."

- I find it a little frustrating that all simulations have been performed with two Gauss-Seidel iterations. No information on convergence speed is provided in the paper whereas from what I understand of the method there is a tradeoff to be found between frequency of radiation call and number of iterations of the Gauss-Seidel method?

The current version of the paper is indeed just presenting results for a very low number of two Gauß-Seidel iterations per call. We limited the results to this setup, as it serves as a kind of worst-case setup for the new solver and already lead to promising results. You are however right that the implications of using more iterations are also very interesting and important. For the revised version, we will thus include a short additional section exploring the effects of using more than just two Gauß-Seidel iterations per call. This section will be centered around a modified version of the figure shown below (and another one concentrating on net surface fluxes).

It shows the time and domain average mean absolute error (MAE) and mean bias errors (MBE) in the heating rates for both the solar and thermal spectral range for three different radiation time steps as a function of the number of Gauß-Seidel iterations used. The plot clearly shows that in terms of the mean MAE, independent of the three radiation time steps considered, only two Gauß-Seidel iterations are already sufficient to basically reach convergence, with more Gauß-Seidel iterations adding little to no improvement in both spectral ranges. It is only in the thermal spectral range where more than two Gauß-Seidel iterations per call lead to a noticeable decrease in the time and domain average MAE. On the other hand, the mean MBE significantly improves – or at least converges towards the original TenStream MBE – the more Gauß-Seidel iterations one takes into account. Hence, at least for the shallow cumulus cloud time series investigated, more Gauß-Seidel iterations per call of the radiation model seem to mainly reduce the bias built up by the incomplete solves.



- L.559 I disagree with "almost perfectly". This formulation is not great anyway, as something that is not "entirely" perfect is by definition imperfect.

You are right that this formulation is not making much sense. We got rid of the word “perfectly” for the revised version: “Comparing these results to those of our newly developed dynamic TenStream solver, we can see that also in the thermal spectral range, it is almost able to reproduce the results of the original TenStream solver, even when operated at lower calling frequencies.”.

- L.570 I disagree with "full three-dimensional radiative transport" as it is far from being full considering the limited number of streams and other remaining approximations.

You are certainly also right with that. For the revised version, we have thus modified the corresponding sentence as follows: “In contrast to these results, however, the dynamic TenStream result features horizontal transport of radiative energy, resulting in much more realistically distributed heating rates and net surface irradiance patterns.”

Technical corrections

- First paragraph of Introduction, I would also mention the importance of surface fluxes and not just heating rates.

Thanks for pointing this out. We have changed the corresponding sentence to: “They are quantified by heating rates and net surface irradiances and are calculated using radiative transfer models, which describe the transport of radiative energy through Earth's atmosphere, ideally allowing for full three-dimensional (3D) transport of energy.”

- L.39-40, add "in the solar spectral range"?

We changed that as suggested.

- 2.1 title: I think you describe more than the TenStream "solver"; you describe the underlying radiative transfer "model". Would it be fair to say that this same model can either be solved as in the original TenStream solver, or as in the Dynamic TenStream?

Yes, that is certainly a good point. We have changed the title of Sect. 2.1 to “The original TenStream model”.

-L.88 I was bothered by the use of "transmittance" here as the a-coefficient also account for incoming scattering and I thought that transmittance was defined as the complementary to extinction along a given line sight; but I might be wrong.

You are probably right that transmittance just refers to the complementary of extinction along a given line of sight. Thus, we have changed the corresponding sentence as follows: “While the "a"-coefficients describe the transport of diffuse radiation, the "b"-coefficients quantify the fraction of direct radiation that gets scattered, thus providing a source term for the ten diffuse streams.”

- In Fig. 3, it took me some time to understand that horizontal arrows between horizontally adjacent grid-boxes, as well as one of the two vertical downwelling arrows between vertically adjacent grid-boxes, represent direct solar radiation propagation. It might be worth it to mention it in the caption or to distinguish them somehow or maybe remove them from the schematics?

We added additional information to the caption of Fig. 3 clarifying that it visualizes both direct and diffuse streams: “Two-dimensional schematic illustration of the first four steps of a Gauß-Seidel iteration, showing both diffuse and direct TenStream fluxes in case of Sun shining from the west or left-hand side.”

Having these direct streams in Fig. 3 is crucial to understand the iteration direction through the domain, which is why we leave them in the figure. We also decided against distinguishing them colorwise, as we really want to focus on the information whether fluxes are updated or not and not distract the reader from that by adding another color.

- In Fig.5, consider using a more contrasted color palette for the various circles?

The color palette in Fig. 5 was chosen so that it matches the shade of blue used in various other plots such as Fig. 2. Using a darker blue as base color does not add significantly more contrast to the plot, which is why we decided to stay with that color scheme.

- Figs. 6-9 are impossible to read for color-blind people.

We invested a lot of time and tried a wide range of different color palettes to make Figs. 6-9 as accessible to color-blind people as possible. In the end, these colors achieved the best results in the Coblis color blindness simulator referred to at the GMD website, while still providing a pleasant experience for people without color deficiencies. In fact, the plots should be able to read even for people with a monochromatic color blindness, as we use different line styles for the different solvers (solid for the delta-Eddington solver, dashed for the original TenStream solver and dash-dotted lines for the dynamic TenStream solver) and different levels of brightness for the different radiation time steps, making every line in the plot unique. We are aware that the plots are certainly still not ideally suited for color-blind people, but in the end they offered the best trade-off between readability for people without major color deficiencies and color-blind people that we could find.

- L.448 "the dynamic TenStream solver overestimates thermal heating rates" is not very clear here, do you mean overestimates their magnitude knowing that they are negative (i.e. they are more negative than the classical TenStream)?

Exactly. For the revised version, we clarified that we refer the magnitude of the thermal heating rates here: "But in contrast to the solar spectral range, these heating rates get more negative the less the dynamic TenStream solver is called, so that the dynamic TenStream solver overestimates the magnitude of these thermal heating rates when compared to the original TenStream solver it is based on."

- L.548 I was bothered by the use of "emission" here as I think it might be confusing; consider sticking to "flux" or "irradiance"?

That is a good point. We changed that for the revised version: "This also leads to a very distinct pattern of strongly negative and not so negative net surface irradiance areas at the ground in the 1D results, whereas the net surface irradiance is almost uniform in the MYSTIC benchmark result.".

- Page 27, why not use the more precise term of quadrature point instead of bands?

Thank you for the suggestion. We used the term "spectral bands" instead of "quadrature points" as it seemed easier to understand for a general audience, but given that "quadrature points" is the usual term used in the literature, we have changed that for the revised version.