Dr. Lee de Mora
Plymouth Marine Laboratory
Prospect Place
The Hoe
Plymouth
PL1 3DH

Dear Geoscientific Communication editors, referees and reviewers,

We received two review comments and one short comment. We have addressed their comments in this document, and made changes to the main document, also attached. We're like to thank both anonymous referees and the short comment author for their helpful comments. Thanks to their contributions, this work is in a much better state and should be easier to follow.

For clarity, we will reproduce the comments and respond to them in turn. Our responses are marked in bold and begin with **"LdM:".** The new text are included with an indent and may include some latex grammar. Alternatively, a document showing the difference between the old and the new version is also available.

Sincerely,

Lee de Mora – representing the authorship team.

Anonymous Referee #2

This manuscript describes six music pieces that have been produced to make climate data accessible to non-experts. The aims are (1) to generate music pieces using climate model data, (2) to use music to illustrate standard practices in Earth System modelling to non-experts, and (3) to quantify the dissemination of music pieces. The method employed here (i.e., data sonification or turning data into music) is a powerful approach and have been successfully used by others for presenting complex datasets to engage/inspire those outside the expert community. The method is particularly well suited for working with climate data. The authors have done a thorough job of explaining each music video.

LdM: Thank you for the summary and the kind words.

However, there are several sections that need improvements (particularly the method and evaluation). Therefore, I see two major changes:

1. Of the above three goals (also stated in the manuscript), goal 1 was clearly achieved. However, it is not clear how this manuscript addresses goal 2. As for goal 3, there was no systematic, robust documentation of the authors' dissemination strategy, the audience demographics, learning, etc. Therefore, a more quantitative (and systematic) assessment of the video usage is needed. The authors state some strategies for doing this (in discussion and conclusion, e.g., performing to a live audience and surveying the audience to measure impact). I think such strategies are great and should be implemented. It is difficult to indicate if a science communication product is helpful (and if so in what way) without any systematic assessment. Therefore, I highly encourage the authors to consider evaluating their videos, and adding the analysis of their findings to the manuscript.

LdM: We have added more detail to the quantification of reach, results, discussions and limitations sections. There are many changes throughout the paper and we refer the editor and reviewer to the difference document, which lists these changes.

2. The method section (lines 106-192) is difficult to follow for non-musicians. This section explains how the music was produced but fails to explain how it relates to climate data. I think giving some examples may increase its readability. For example, when you state "the lowest value in the dataset is presented by the lowest note. . ." (line 120), it may be helpful to give an example of the lowest value in the dataset (e.g., coldest recorded temperature). The same goes for the highest value in the dataset (line 121).

LdM: We have expanded this whole section and made many clarifications. Unfortunately, music theory is its own discipline (as is music composition) and it's not possible to make a complete, thorough, brief and easy to follow explanation here.

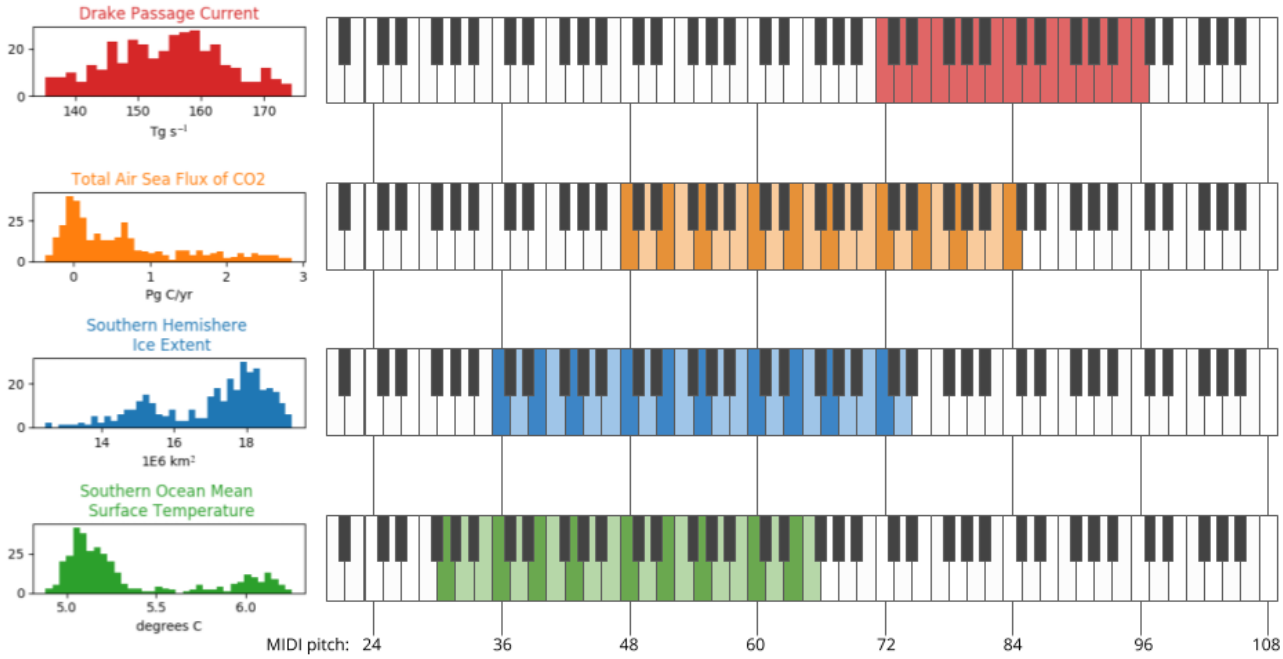
Line 119 also says "each model dataset is linked to a series of notes", so does this mean each note is a data point? Again, translating this into climate data would be helpful. Similar suggestions for Fig 2 (see below) and lines 169-172. Also, many of the comments shown in the next few pages are related to this issue.

LdM: In addition to adding the new keyboard figure (see below), I've changed this entire paragraph to:

> Each model timeseries dataset is converted into a series of consecutive MIDI notes, which form a track. For instance, the Sea Surface Temperature (SST) time series could be converted into a series of MIDI notes in the upper range of the keyboard, forming a track. For each track, the time series data is converted into musical notes such that the lowest value in the dataset is represented by the lowest note pitch available, and the highest value of the dataset is represented by the highest pitch note available. The notes in between are assigned proportionally by their data value between the highest and lowest pitched notes. The lowest and highest notes available for each track are pre-defined in the piece's settings. Each track is given its own customised pitch range, so that the tracks may be lower pitch, higher pitch or have overlapping pitch ranges relative to other tracks in the piece. The ranges of notes available for the piece \textit{Earth System Allegro} is shown in fig.~\ref{fig:histograms}. In this figure, the four histograms on the left hand side show the distributions of data used in the piece, and the right hand side shows a standard piano keyboard which the musical range available to each dataset. For instance, the Drake Passage Current ranges between 135 and 175 Tg s^{-1} in these simulations and we selected a range between MIDI pitches 72 and 96. This means that the lowest Drake passage current values (135 Tg s^{-1}) would be represented in MIDI with a pitch of 72 and the highest Drake passage current values (175 Tg s^{-1}) would be assigned a MIDI pitch of 96, which is two octaves higher.

Figure 2- It would be very helpful if you can connect what you show in the piano keyboard to climate data. See figure 1 of George et al. (2017, American Meteorological Sociey) for example. Again what does each note represent? What does each pitch represent? A bit hard to follow as a non-musician.

LdM: Based on the 2017 paper in the Bulletin of the American Meteorological Society by St George et al (*https://doi.org/10.1175/BAMS-D-15-00223.1*) I've added the following image and caption to this paper:



Caption:

> The musical range of each of the datasets used in the Earth System Allegro. The four histograms on the left hand side show the distributions of data used in the piece, and the right hand side shows a standard piano keyboard which the musical range available to each dataset. In this piece, the Drake passage current, shown in red, is free to vary within a two octave range of the C major scale. The other three datasets have their own ranges, but are limited to the notes in the chord progression C major, G major, A minor F major. The dark coloured keys are the notes in C major, but the lighter coloured keys show the other notes which are available. Note that both the C major scale and chord do not include any of the ebony keys on a piano, but these notes would be used if they are within the available range.

Below, my comments are shown line by line:

Line 23 (Introduction)-The authors introduce the topic well, and the references they list are relevant and helpful. Since this study combines sonification with imagery, it would also be helpful to know if this approach has been taken before, and if so, how does this study contribute (or build on) previous work?

LdM: Added the following text.

> It should be noted that all the pieces list here are also accompanied by a video which can explain the methodology behind the creation of the music, shows the performance by the artists, or shows the data development while the music is played.

Line 56 points out the potential for biased-interpretation of data using sonification. However, the authors do not return to this issue later to discuss it. Was this a concern during this study and how was it addressed?

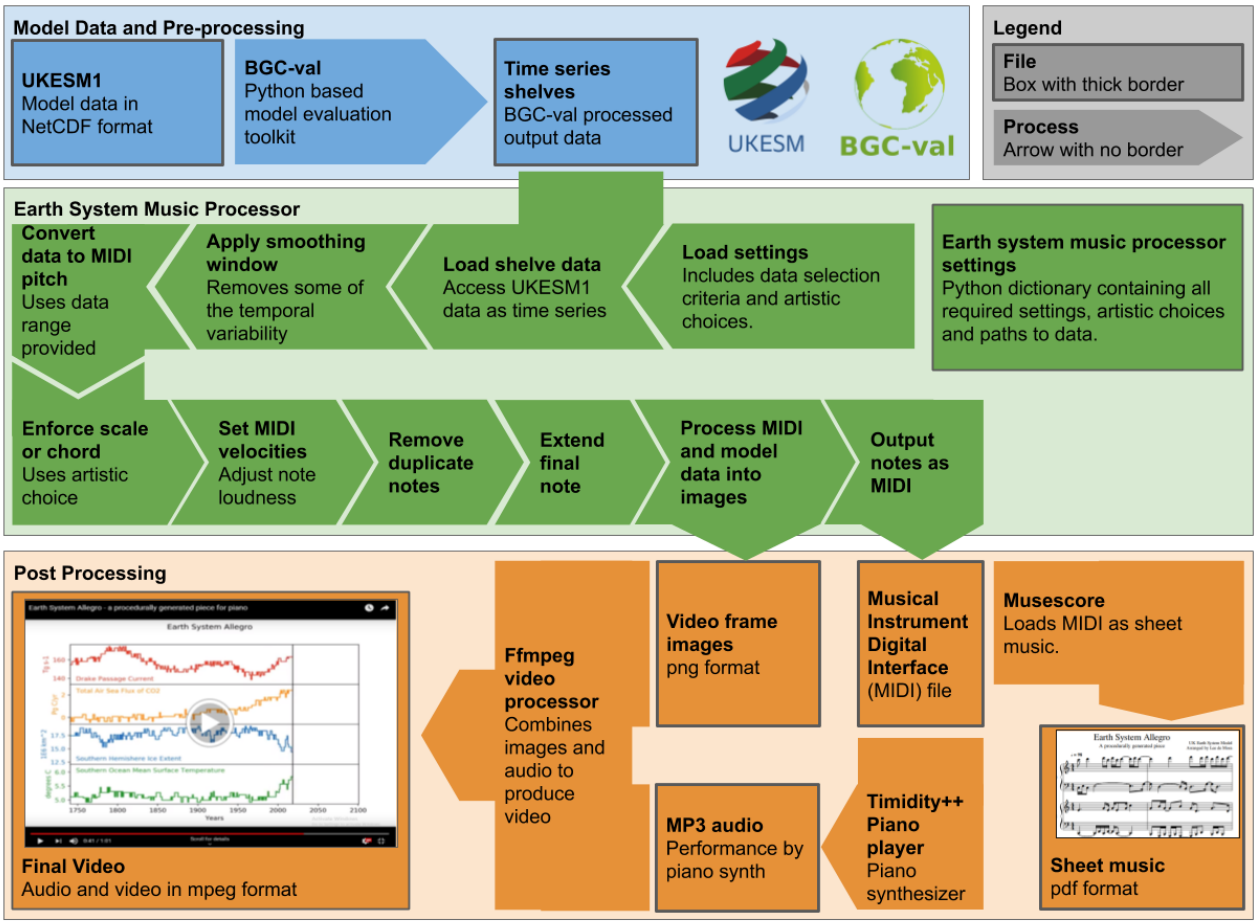
LdM: Upon reflection, we never wanted the musical pieces to be a neutral objective version of the data. We always wanted to try to communicate some of the emotional context of the data.

Re-wrote this paragraph to be:

> In addition to its practical applications, sonification is a unique field where scientific and artistic purposes may coexist \citep{Tsuchiya2015}. This is especially true when in addition to being converted into sound, the data is also converted into music. This branch of sonification is called musification. Through the choice of musical scales and chords, tempo, timbre and volume dynamics, the composer adds emotive meaning to the piece. As such, unlike sonification, musification should be treated as a potentially biased-interpretation of the underlying data. It can not be a true objective representation of the data. Note that the philosophical distinction between sound and music is beyond the scope of this work. Furthermore, even though the composer may have made musical and artistic decisions to link the behaviour of the data with an emotive state, it may not necessarily be interpreted in the same way by the listener.

Figure 1. Though flow charts are generally produced in this way, I suggest to add a few images (one per section) to draw in the readers. The sections are: datasets (top), music (middle) and videos (bottom).

LdM: We have reworked the methods plot now. It should be both clearer, more colourful and include a few images.



The new caption reads:

> The computational process used to convert UKESM1 data into a musical piece and associated video. The boxes with a dark border represent files and datasets, and the arrows and chevrons represent processes. The blue areas are UKESM1 data and the pre-processes stages, the green areas show the data and processing stages needed to convert model data into MIDI data, and orange area show the post processes stages which convert images and MIDI into sheet music and video.

Line 224. When you state “The Earth System Allegro is a relatively fast-paced piece in C Major”, can you describe what C Major sounds like for non-musicians? Also the rest of the sentence starting with “. . .showing some important metrics of the happy to keep. . .” does not make grammatical sense. Please revise.

LdM: Changed this paragraph to:

> The Earth System Allegro is a relatively fast-paced piece in C Major, showing some important metrics of the Southern Ocean in the recent past and projected into the future with the SSP1 1.9. This is the future scenario in which the anthropogenic impact on the climate is at a minimum. The C major scale is composed of only natural notes (no sharp or flat notes), making it one of the first chords that people encounter when learning music. In addition, major chords and scales like C Major typically sound happy. Christian Schubart's `Ideen zu

einer Aesthetik der Tonkunst` (1806) describe C major as "Completely pure. Its character is: innocence, simplicity, naivety, children's talk." As this was the first piece in the series, the link between this seemed an appropriate way to start the Earth System Music project. Through choosing C major and an upbeat tempo, and data from the best possible climate scenario (SSP1 1.9), we aimed to start the project with a piece with a sense of optimism about the future climate and to introduce the principles of musification of UKESM1 time series data.

Line 226. Could you explain how this video demonstrates the principles of sonification using the data series?

LdM: Changed this phrase to:

introduce the principles of musification of UKESM1 time series data.

This was the first piece in the series and does introduce the core-concept of the project, that the music follows the data.

Line 232. I think there may be a typo here. Could it be “year 2030/2040” as oppose to year 2100?

LdM: Changed this phrase to:

Even under SSP1 1.9, UKESM1 predicts that this value would rise from around zero during the pre-industrial period to maximum of approximately 2 Pg of carbon per year around the year 2030, followed by a return to zero at the end of the century.

Line 235. Consider deleting this sentence as it is repetitive.

LdM: removed and replaced with:

The fourth field is the Southern Ocean mean surface temperature, shown in green, which slightly rises from approximately 5 degrees in the pre-industrial period up to a maximum of 6 degrees.

Line 240. Consider deleting the sentence starting with “Effectively, . . .” It is redundant.

LdM: Removed.

Line 248. Change “there’s” to “there is”, and “doesn’t” to “does not” in line 292. And reflect this change throughout the manuscript.

LdM: Done

Line 250. Again “a very common 4 chord song: C Major, G Major, A Minor, and F Major” does not mean anything to a non-musician. Please clarify this by giving an example for each or give a word to describe what they sound like.

LdM: Added the sentence which should add some context.

> This chord progression is strikingly popular and may be hear in songs such as: Let it Be, by the Beatles, No Woman no Cry by Bob Marley and the Whalers, With or without you by U2, I’m yours by Jazon Mraz, amoung many others.

This is a bit of a trick, as some of these songs are the same chord progression in a different key (For instance, Africa by Toto is written in A Major). I fear that adding the complexity of the roman numeral notation is a step too far for this work! Out of interest, the following songs are written using this progression:
https://en.wikipedia.org/wiki/List_of_songs_containing_the_I%E2%80%93V%E2%80%93vi%E2%80%93IV_progression

Lines 250-253. Draft a similar paragraph for section 3.1.1. This helps connect the music structure with what the dataset represents.

LdM: Done, see above.

Line 256. Add a reference to Figure 3, pane 3, at the end of this sentence.

LdM: Done

Lines 257-259. Add the name of scenarios (e.g., SSP5 8.5) to Figure 3, pane 3.

LdM: I’m not able to do this, these figures show the final frame of each video, not the data itself. As the video is already published, it’s not possible to do this. The colour code is described in the text in lines 255-260.

Line 270. Add reference to Figure 3, pane 4, after “E minor”.

LdM: added the sentence:

> The final frame of this video is shown in pane 4 of fig. 3.

Line 273. How are these 15 historical simulations are shown in the figure? Only 6 lines are shown. Have they been grouped?

LdM: As mentioned above, figure 3 only shows the final frame of the videos. I’ve changed figure 3’s caption to be clearer:

> Figure 3: The final frame of each of the six videos. These frames of the videos are shown in the order that they were published. The videos 1), 3), 5) and 6) use a consistent x-axis for the

duration of the video, but videos 2) and 4) have x-axes which changes over the course of the video. This means that panes 2 and 4 show only a small part of time range.

Line 274. “This piece uses a repeating 12 bar blues structure in E minor”, what does this mean to a non-musician, and how is this connected to the dataset it is reflecting?

LdM: Added the following paragraph to make the point clear about the choice of a 12 bar blues.

> This piece uses a repeating 12 bar blues structure in E minor and a relatively fast tempo. This chord progression is was exceptionally common progression, especially in the blues, Jazz and early rock n roll sounds. It is composed of four bars of the E minor, two bars of A minor, 2 bars of E minor, then one bar of B minor, A minor, E minor and B minor. The twelve bar blues can be be heard in songs such as: Johnny B. Goode by Chuck Berry ,Hound Dog by Elvis Presley, I got you (I feel Good) by James Brown, Sweet Home Chicago by Robert Johnson or Rock n Roll by Led Zeppelin. In the context of Earth System Music, the 12 bar pattern with its opening set of four bars, then two sets of two bar and ending for four sets of one bar between key changes drives the song forward before starting again slowly. This behaviour is thematically similar to the behaviour of the ocean acidification in UKESM1 historical simulation, where the bulk of the acidification occurs at the end of each historical period.

Line 285. What initial conditions are the authors referring to?

LdM: Changed the text to

> When we produce models of the Earth System, we use a range of points of the pre-industrial control as the initial conditions for the historical simulations. All the historical simulations have slightly different starting points, and evolve from these different initial conditions, which gives us more confidence that the results of our projections are due to changes since the pre-industrial period instead of simply a consequence of the initial conditions.

Line 286. When you state “. . .the results of our projections are due to changes. . .” what changes are the authors referring to?

LdM: see above

Line 294. Please give an example of what it is meant by “inherent change” and “underlying drift”.

LdM: I’ve re-written this paragraph for greater clarity and this should address most of the issues raised in this section.

> This piece combines the spin up of the United Kingdom Earth System Model with the chord progression of John Coltrane's Giant Steps. The spin up is the process of running the model from a set of initial condition to an equilibrium steady state. When a model reaches a steady state, this means that there is no significant trend or drift in the mean behaviour of several key metrics. For instance, as part of the C4MIP protocol, Jones et al (2016) suggest a drift criterion of less than 10 Pg of Carbon per century in the absolute value of the flux of CO2

from the atmosphere to the ocean. In practical terms, the ocean model is considered to be spun up when the long-term average of the air sea flux of Carbon is consistently between -0.1 and 0.1 Pg of carbon per year.

The spin up is a crucial part of model development. Without spinning up, the historical ocean model would still be equilibrating with the atmosphere. It would be much more difficult to separate the trends in the historical and future scenarios from the underlying trend of a model still trying to equilibrate. Note that while a steady state model does not have any significant long term trend or drifts; it can still have short term variability. This short term variability can be seen in the pre-industrial simulation in the Pre-industrial Vivace piece. It can take a model thousands of years of simulation for the ocean to reach a steady state. In our case, the spin up ran for approximately 5000 simulated years before the spun up drift criterion were met , Yool 2020.

Line 295. The spin up ran for 5000 simulated years. Why 5000 years? How was this time selected? A reference is provided, but it would be useful to add a sentence explaining why.

LdM: I've added more details on the spin up criteria, please see above. Also note that the UKESM spin up paper has recently been submitted to JAMES after review and is expected to be accepted for final publication before this paper. A suitable reference will be added when it is available.

Line 305. It may be useful to label these lines in Figure 3, pane 5 or describe them in the caption.

LdM: These lines do appear in figure 3, pane 5, I've added the following text to the caption.

Pane 5 includes two vertical lines showing the jumps in the spin up piece. Pane 6 shows a single vertical line for the crossover between the historical and future scenarios.

Line 311. Why was the musical progression slowed to one chord per four beats? What does it mean in terms of the climate dataset?

LdM: To be honest, this was a happy accident due to a bug in the original code. The original version at full speed just sounded too chaotic. Changed the text to:

This change occurred as an accident, but we found that the full speed version sounded very chaotic, so the slowed version was published instead.

Line 335. Decapitalize "Global total ice" and insert a space between "extent" and "blue".

LdM: Done

Line 365. Change "view" to "video" in "the percentage of the view that the average audience viewed".

LdM: Done

Line 366-369. Consider deleting this part starting from “Aside from the metrics. . .” These numbers are too small to be meaningful, and are not discussed.

LdM: The first reviewer had a stronger interest in metrics beyond the youtube statistics. We have moved this part into that section and extended the discussion to include these metrics.

Lines 370-378. Consider deleting the whole paragraph or move to discussion.

LdM: Moved to discussion.

Figure 7. Consider removing it from the manuscript, but keep the text (lines 387 onward). The figure does not add much to the manuscript, especially when half of the data is unknown.

LdM: Figure removed but kept the text.

Lines 390-392. Consider deleting this or move to discussion.

LdM: Split this paragraph into two and put second half in the discussion.

Line 406. The study goals stated here differ from those stated in page 10. Please keep the goals consistent.

LdM: Fixed.

Line 411. The authors conclude that once the concept was demonstrated, there was reduced enthusiasm from the audience to return to it. How do they know that? Another possibility could be that the audience didn’t feel the need to return to it, or it could also be that the videos sparked their interest further so that they ended up checking out similar videos outside the playlist. These are all possibilities, and there is no evidence for or against them. I suggest sticking to the facts, and only interpret the data when it is actually possible (which is not the case here).

LdM: Removed this sentence.

Line 412. The last sentence may be true but it is irrelevant to this paragraph. Was the goal to grow a YouTube channel? Why do the authors mention this here?

LdM: Removed this sentence

Lines 420-426. The sea surface temperature aria is also the most visually simple animation when compared with the rest. The viewer is not required to keep track of multiple datasets and listen to the music at the same time. Could this be also why this piece has the highest audience retention?

LdM: added:

> The sea surface temperature aria is also the most visually simple animation of the six pieces. Only one pane is visible in the video and much of the piece only includes one or two voices at a time. It may be possible that this simplicity holds the audience's attention.

Line 430. There is no documented evidence that the music pieces and animations improved the wider public's understanding of climate change modelling. The authors mention this in the next paragraph. So I suggest to delete the "perhaps, improve the wider public's understanding of climate change modelling". One could hope for that, but this study was not designed to assess that, and certainly did not do that.

LdM: Changed text to:

> While we hoped to improve the wider public's understanding of the methods used in climate change modelling, the tools available to us within YouTube studio do not allow any way of assessing this. Please see the Limitations and Future Work section, below.

I suggest to move some of the content currently placed in the discussion section to two new sections: Limitations and Future Work. This means most of what is shown in page 19-20 can be reorganized to fit into one of these two sections. This might help the readers.

LdM: Created this section and reorganised the discussion into two parts.

Line 439. Here the authors suggest hosting live events to fully explain the methodology used by the modelling community. But is this something non-experts are interested in, or is this the aim of this study? I thought the idea was to use a unique communication method (sonification and imagery) to explain complex datasets to nonexperts. If this method requires a live event for further explanation, then it does not fulfill what it was supposed to do: to engage non-experts.

LdM: At the moment, the videos by themselves do not even attempt to include an explanation of the methods used in Earth System Modelling. All explanation was in the video description below the video. However, the point of this paragraph is confused by the addition of a sentence about a live performance, so it was removed.

We changed this paragraph to:

> The videos themselves only include the music and a visualisation of the data, they do not include any description about how the music was generated or the Earth system modelling methods used to create the underlying data. The explanations of the science and musification methodologies are held in a text description below the video. Furthermore, viewers must expand this box by clicking the ``show more`` button. Using YouTube studio, it is not currently possible to determine whether the viewers have expanded, read or understood the description. When we have shown these videos to live audiences at scientific meetings and conferences, it has always been associated with a brief explanation of the methods. In the future, this explanatory preface to the work could be included in the video itself or as a separate video.

Line 462. The authors state that it was hard to distinguish the different datasets in the music. One solution would be to insert a very short silence in music between different datasets. Just an idea.

LdM: The confusion here is that one sentence is trying to cover two ideas. We changed this sentence to:

> These pieces were all performed by the same instrument, a solo piano, which limits the musical diversity of the set of pieces. In addition, each dataset within in a given piece was performed by the same instrument, making it difficult to distinguish the different datasets being performed simultaneously.

Line 504. Insert a space between the word viral and the references.

LdM: done

Line 510. Insert a space between the word afternoons and the references.

LdM: done

Line 514. “they reached an audience of 251 unique viewers and a total view count of 553”

LdM: done

Table 1. Add unit of time for “duration”. Minutes?

LdM: done