AUTHORS' RESPONSE 1

2

3 Dear Editor,

4

5 thank you and both reviewers for evaluating our manuscript. Please find our summary reply below,

- 6 followed by detailed answers to individual reviewers' comments.
- 7

8 Jiří Mikšovský, on behalf of the authors

9

10 **SUMMARY OF MAJOR POINTS**

11 One of the major concerns raised by both reviewers was the lack of aerosol-related forcing among 12 the predictors employed. As discussed in more detail in the replies to Rev#1's comment (C) and 13 Rev#2's comment (1), our choice of GHG-only predictor was intentional; moreover, its replacement 14 with total global anthropogenic forcing (involving the effect of aerosols) would not result in any 15 major change to the results.

16 The reviewers also suggested several possible extensions to the manuscript, expanding the analysis 17 itself or the presentation of its setup and results. We have incorporated some of the suggested 18 changes into the revised version (such as an overview of the inter-predictor correlations, or 19 visualization of the global temperature series and their regression-based fits). In some cases, 20 however, the suggested changes either turned out to not contribute very interesting results 21 (inclusion of the GISTEMP data with 250 km smoothing), or went outside of the intended scope of 22 the analysis (assessment of local impacts of the detected patterns, discussion/inclusion of climate 23 feedbacks).

24 The reviewers have also expressed concerns about collinearity of some of the predictors and its 25 potential effect on the results. We have somewhat modified the respective sections of the text to 26 better address the related issues and to highlight them for the reader (including the addition of Table 27 1, summarizing the correlations). However, it should be emphasized that many of the related 28 questions are still open: They have been a focus of various past studies (some of them referenced in 29 the manuscript), but final conclusions are yet to be established (this particularly applies to the nature 30 and effects of the Atlantic Multidecadal Oscillation). Our paper has no ambition to be the final 31 answer in this discussion, but rather to provide another piece to the puzzle.

32

33 **MAJOR CHANGES TO THE MANUSCRIPT**

- 34 We added Table 1 to the manuscript and Table S1 to the Supplement, summarizing 35 correlations between the predictor series, as well as respective variance inflation factors
- 36
- We added a new figure (Fig. S4 in the revised version) to the Supplement, presenting time -37 series of global temperature and their approximation by the regression mappings
- 38 We extended the discussion of the implications of our choice of the GHG-related predictor 39 and of the related issues, as well as some of the passages dealing with potential relations 40 between predictors
- 41 We included additional references, either in response to the comments of the reviewers (Zhang and Delworth, 2007; Knudsen et al., 2014), or to refer to a recent related work 42 43 (Rypdal, 2015).
- 44

1 Attached, the revised manuscript and its Supplement with the proposed changes are provided.

2

Below, the reviewers' comments are shown, highlighted in bold with gray background, with ourreplies in regular font.

1 **REVIEWER #1**

2

The submitted paper (by Miksovsky et al) provides a useful information and I suggest that the paper be accepted for publication after a major revision. I have three major concerned (A, B, and C) and a few minor ones. I believe that the three major concerned should be dealt with in the revision. The minor points I would leave for consideration of the authors.

7

8 (A) The paper shows qualitatively the differences between four different gridded global 9 temperature data. However, we learn very little about the causes of these differences. The authors 10 miss the opportunity to analyze and compare two data sets that could clearly identify the cause of 11 potential differences between them. I mean specifically the NASA GISS data sets with 1200 km 12 smoothing radius (data used in the paper) and data set with 250 km smoothing radius. I consider it 13 essential that the 250 km smoothing set be included in the analysis and differences (between 1200 14 km and 250 km smoothing) of estimated contributions of individual predictors to local 15 temperature be shown. Special attention should be paid to polar region where the differences 16 might be significant. I think this could be the only case where the explicit action (smoothing) could 17 be connected to temperature differences.

18

19 We actually considered both versions of GISTEMP for our analysis, but ultimately chose the variant 20 with 1200 km smoothing to be more suitable. The main reason lied with substantially more limited 21 data coverage in the 250 km version of the dataset: Only relatively small number of grid points 22 contains enough non-missing values over our primary analysis period (i.e. 1901-2010). There is a 23 strong similarity between the two variants of the dataset in areas with good data coverage 24 (especially Europe and much of North America) – see Fig. R1. More noticeable differences do appear 25 in some geographically limited regions, especially around islandic locations – such local contrasts are 26 however likely related to specific behavior of individual measuring sites, and assessing them 27 individually would go beyond the intended scope of our study. Also note that the 250 km version of 28 GISTEMP provided at NASA GISS page is land-based only, unlike the other temperature datasets we 29 employed.

That said, there do not seem to be major systematic differences between the regression analysis outcomes from the 1200 and 250 km version of the data: As seen from Fig. R2, the temperature responses are quite similar in both datasets. The main distinction seems to be higher spatial variability in the response patterns detected from the 250 km variant – an expected result, considering the lower degree of smoothing.

35 <u>CHANGE MADE:</u> Mention of the experiments with the 250 km version of GISTEMP has been added in 36 the revised manuscript (in the Sect. 2.2). However, for the reasons given above, no specific analysis 37 outputs (such as new figures) based on the 250 km version have been included in the revised version. 38

50

(B) The regression coefficients and consequently contribution of individual predictors to temperature depends on the all other predictors present in regression model. How does the contribution of individual predictors change with decreasing complexity of regression model? How they change if you use only the most effective predictors like GHG, volcanic aerosol, and AMO (perhaps also PDO) based on results shown in your Fig. 4?

- 1 Model considering only the explanatory variables most influential on global scale (GHG amount,
- 2 volcanic aerosol, AMO index) produces results very similar to our original setup with all 8 predictors
- 3 (see a demonstration for GISTEMP temperature in Fig. R3).
- 4
- 5 (C) The important topic the collinearity of GHG and anthropogenic aerosols is discussed fully 6 only in the conclusion section. It should be brought in early stages of the paper. Instead of GHG 7 forcing perhaps the sum of GHG+aerosol should be considered (similar to Lean and Rind) and 8 emphasize that in the second half of the 20th century GHG warming cannot be distinguished from 9 aerosol warming (decrease of sulfate aerosols in line with Booth et al).
- We deliberately included the discussion of the representation of anthropogenic forcings in Sect. 5 rather than earlier in the paper, because we felt it fitted better with the rest of the discussion regarding the uncertainties in our results. Note also that the associated challenges are already briefly mentioned in the Data section of our manuscript, so the reader is forewarned about the potential interpretational pitfalls.
- As for the replacement of GHG-only forcing with a composite one including other anthropogenic effects (such as the aerosols), it would have very little effect on the outcomes of our analysis. This may sound somewhat surprising, as the radiative forcing from the aerosols is certainly important from the physical point of view, and should be considered when assessing radiative budget. However, from a perspective of purely statistical analysis, the respective predictor series are almost identical in shape, and so are the results. For instance, over the 1901-2010 period, the Pearson correlation coefficient between our preferred predictor series and other alternatives are generally
- 22 over 0.99, specifically:
- 23 correlation with CO₂-equivalent representing total anthropogenic forcing (incl. aerosols) is 0.992
- 24 correlation with concentration of CO₂ alone is 0.999
- 25 correlation with total anthropogenic forcing (in W.m⁻² rather than CO₂-equivalent)is 0.996
- (all data were obtained from <u>http://www.pik-potsdam.de/~mmalte/rcps/</u>, just as the original
 predictor employed in the manuscript).
- 28 In Fig. R4, the effect of predictor replacement is demonstrated for the 'anthropogenic' component in
- 29 the GISTEMP global temperatures series: It should be clear that the outcomes are indeed very
- similar. (Note that this near-equivalence holds for the temperature responses, i.e. the quantities
 presented in our paper, but not for the regression coefficients themselves, which are scaling dependent.)
- Considering our focus on local scales (for which GHG concentration represents a more universal predictor than forcing itself), and preference for using a single unified set of predictors in all our tests, we have therefore favored GHG-equivalent concentration as our 'anthropogenic' predictor.
- 36 Cooling related to the drop of aerosol amounts may be relevant for some regions (such as Europe),
- but not necessarily universally across the globe, as there was not a substantial aerosol decrease in
- 38 some areas (such as China). Since we do not discuss regional specifics in such depth, we do not
- 39 mention this specifically in the manuscript.
- 40 **CHANGE MADE:** In the Discussion, we have somewhat expanded the section dealing with the problem 41 of distinguishing between GHG- and aerosol-related effects, to make our choices and related
- 42 implications more clear. The respective part of the Data section has also been expanded a bit.
- 43

1	1. P2341 line 10: It is not clear what you mean by the second half of the sentence "although
2	linear trend"
3	The sentence in question refers to the fact that many studies dealing with attribution use linear trend
4	as the approximation of the long-term component in the temperature series (i.e., time itself is used
5	as predictor).
6	<u>CHANGE MADE:</u> The respective formulation has been modified to " although linear change with time
7	"
8	
9	2. 2343, 3: Are the linear correlations sufficient to quantify the match? Some justification is needed
10	or inclusion of some other variable.
11	The Pearson correlation coefficient has been chosen in our case because it scales linearly to the
12	anomalies in the series investigated, just as (multiple) linear regression, out main analysis tool,
13	considers linear proportionality. While we have not tried "nonparametric" correlation measures such
14	as Spearman correlation coefficient in this particular case, our previous experiments with their
15	application suggest high similarity of the outcomes (at least for the monthly temperature anomalies).
10	Alternatively, fully nonlinear option could be employed, namely some form of average mutual information. Such approach would however require more detailed explanation of the correlation
17 18	estimation methodology and would diverge topically from the rest of the analyses in the manuscript
10	estimation methodology and would diverge topically nom the rest of the analyses in the manuscript.
19	2 2242 40 M/by and station based indians undernal source and deduced form university
20 21	3. 2343, 10: Why are station based indices preferred over ones deduced from principal
21	
22	We do not specifically explain this in the manuscript, but this was done to maintain consistency with
23 24	indices are often the only ontion available
24	Curves are orten the only option available.
25 26	<u>CHANGE MADE:</u> I removed the mention of the preference of the station-based indices, as the
20 27	
27 20	4 2244 11. Anthronogonic corocole played an important role in forming the 20th contury
20 29	temperature profiles Why can we ignore them? How would be results modified by inclusion of
30	anthropogenic aerosols? This is essential: I expect considerable changes in results when aerosol are
31	included.
32	Please see our reply to comment (C) and also reply to Rev.#2's comment (1).
33	
34	5. 2344, 16: The reference for stratospheric aerosol is to Sate et al 1993. How were data updated to
35	2010?
36	We did not make any update ourselves; the respective series (extending beyond 2010) were
37	downloaded from the pages of NASA GISS. Link to this source is provided in the manuscript; the
38	reference to the paper by Sato et al. (1993) seems to be generally used with this dataset, even for its
39	updated versions (e.g. Schmidt et al., 2011).
40	
41	6. 2346, 2: The reference to time delayed correlations is to Wu et al 2011. However, the original
42	paper on the time-delayed correlations was published earlier by R. Zhang et al.

- 1 **<u>CHANGE MADE:</u>** We have added reference to Zhang and Delworth (2007) to the text, while keeping the
- 2 one to Wu et al. (2011) as well (since it shows the relevant delayed correlation plots).
- 3

7. 2346, 8: We need some information concerning the correlation between the predictors. Find the
 way how to provide some information.

CHANGE MADE: We added tables to the manuscript (Table 1) and to the Supplement (Table S1),
 showing the values of correlation for all pair-wise combinations of the predictor series, as well as the
 values of the variance inflation factors.

9

8. 2346, 14: You use the GISTEMP data with 1200 km smoothing. However, Hansen and Lebedeff
(1987) warn that there may be significant differences on local and regional scales between the
1200 km and 250 km smoothing. Why not use the 250 km smoothing as another set of gridded
temperature data? This could be an important contribution.

- 14 Level of smoothing/interpolation is indeed one of the factors potentially affecting the properties of
- 15 the gridded data, but GISTEMP250 may not provide sufficient data coverage for our purposes. Please
- 16 see our response to the comment (A).
- 17

9. 2346-2347: I cannot find the latitudinal range of used gridded temperature data. I see a full globe or 60S to 75N for 20CR but not for observed temperatures.

The extent of the gridded data is formally global (i.e., pole-to-pole), but there are missing values for some locations and periods. The mention of the latitudinal range only pertains to the 20CR dataset,

- for which we calculated the global averages ourselves (for all other datasets, global means provided by the respective re
- 23 by the respective research groups were used, as stated in the manuscript).
- 24

25 10. 2348, 8, 21: The statistical regression is relatively simple and un-expensive method. Can you 26 justify the references here to "..high computational cost . . ."?

27 The mention of high computational costs at line 8 refers to the nonlinear alternatives of multiple 28 linear regression (MLR). In particular, artificial neural networks, a popular choice for data-processing 29 tasks, can be quite demanding with regard to the computational resources (particularly when 30 combined with resampling techniques such as bootstrap for estimation of statistical significance). 31 Note that while this is certainly not too prohibitive when processing only a few series, we worked 32 with hundreds of thousands of them across the individual datasets, and the computational demands 33 were further multiplied by the bootstrapping (at least 10 000 resamples required when evaluating 34 statistical significance at the α = 0.01 level) and assessing time-offset responses. Under such setup, 35 even MLR-based calculations can take rather long time to carry out - that is why we explored and 36 used alternatives to bootstrap (as described in the paragraph starting at line 21).

37

11. 2349, 15: You say the results for the first and second half of date are presented in Supplement. It would be nice to indicate here by one sentence how different or not the results are.

This mention at p. 2349 is only a general reference to the materials in the Supplement. The specific results for individual sub-periods (as well as eventual differences and their implications) are

42 discussed in the individual segments of Sects. 4 and 5, when relevant.

1 **12. 2349**, 18: "..outcomes of regression analysis are provided" I do not see any information 2 concerning the regressions. Are all predictors kept for all regions of the globe, even when

3 predictors are significantly correlated?

- 4 The same configuration of predictors was kept for the analysis of all global and local temperature 5 series (this is explicitly stated at the start of Sect. 4.3).
- 6

13. 2376, Fig. 2: What are the latitudinal limits of figures? 90S to 90N? Please, add GISTEMP1200 km smoothing vs GISTEMP 250 km smoothing. This is essential since this would be the only case where you clearly know the differences between the treatments.

- The range of latitudes in the maps is from 60°S to 75°N (the polar-most regions were not shown, as there is generally not enough data for our analysis period, i.e. more than 10% of values are unavailable in most gridpoints/datasets). Regarding inclusion of the 250 km version of GISTEMP, please see our answer to the comment (A).
- 14

14. 2378, Fig. 4: Does the global mean from 90S to 90N? Why solar irradiance has no effect on the temperature? Is the solar variability imprinted on the AMOI? If you use the predictors without the AMOI would the solar variability become significant?

- 18 For the temperature analysis data (GISTEMP, Berkeley Earth, MLOST, HadCRUT), global series were
- 19 used as provided by the respective teams. For the 20^{th} Century Reanalysis, the averaging range is
- 20 60°S to 75°N, as stated in Fig. 4's caption (while the difference compared to the averaging range 90°S
- 21 to 90°N can be seem from Fig. 3).
- The nature and magnitude of the effects of solar irradiance variations have been investigated by a number of studies in the past. While the results differ to some extent, the prevailing conclusion seems to be that the solar imprint is weak in the lower troposphere. We reference some of the relevant studies in our manuscript, though the list is certainly not complete, since the number of existing works targeting this topic is quite large and the respective issue is just one of several
- 27 investigated in our study.
- 28 As seen from Table 1 in the revised paper, the correlation of solar irradiance and AMOI is quite mild;
- also, regression of the AMOI series upon solar irradiance explains only about 2.5% of AMO's total variance. There is therefore no reason to assume that AMO index hides the effect of variations in
- 31 solar activity.
- 32

15. 2379, Fig. 5: It looks like you are keeping all predictors at all locations. Due to probable significant correlations in some regions, the regression coefficients will be uncertain and contribution of individual predictors to total temperature equally uncertain. Atmospheric aerosols are highly correlated to GHG. How would aerosols change the GHG contribution? Similarly the solar activity is correlated to GHG – why is solar activity mostly cooling the NH? Is this to compensate the too strong warming by GHG?

- 39 The uncertainty of the regression coefficients (and thus respective contributions to temperature) due
- 40 to correlations among predictors should now be clear from the Tables 1 and S1, where variance
- 41 inflation factors are shown. The respective uncertainties are of course considered in the estimation
- 42 of the confidence intervals and statistical significance, presented in the manuscript.
- 43 For discussion of aerosol inclusion, please see replies to comment (C) and Rev#2's comment (1).

- Our results show only limited effect of solar activity on temperature, inconclusive in local temperatures (i.e., the respective contributions were found statistically insignificant in most areas – Fig. 5b). I would therefore not dare to conclude that there is a prevalent cooling (or warming) in the northern hemisphere based on our results.
- 5

16. 2379, Fig. 5: An inclusion of the GISTEMP data with 250 km smoothing (as the forth column) will be important here to see the differences especially in polar region. The color scheme is not well chosen. The GHG contribution is just red, even the hatching is not visible.

9 For the 250 km GISTEMP data inclusion, please see the response to comment (A); note, in particular,

- 10 that the 250 km version of GISTEMP has very limited amount of data available in the polar regions,
- insufficient to reliable support the attribution analysis for our primary target period. Furthermore,
 considering the opinion of Rev#2 about figures being already too small, addition of another column
- 13 would be problematic.
- 14 The color scale in the Figs. 5 and 7 was a compromise between encompassing the entire range of 15 temperature responses, providing sufficient number of steps in the scale (with extra resolution 16 around 0) and using a unified scale for all predictors (to facilitate intercomparison of their effects). 17 Due to this, only part of the scale is relevant for some predictors, especially GHG concentration. I 18 tried to pick individual colors to be reasonably distinguishable (and hatching visible) on the computer 19 screens at my disposal; admittedly, on some devices (computer screens or printers), the readability 20 may be diminished. It is however difficult to find a universally best visibility formula without 21 sacrificing details, and occasionally the readability may be suboptimal at some devices.
- 22

17. 2353, 4: The correlation between GHG and solar variability should to be mentioned here. Such correlation affects interpretability and makes conclusions difficult

24 correlation affects interpretability and makes conclusions difficult.

- It actually was mentioned, but perhaps not clearly enough (we spoke of 'aliasing' of the two signals,which might have been misleading).
- 27 CHANGE MADE: I have reformulated the respective sentence to make a specific reference to the
 28 correlation between solar irradiance and GHG amounts in the 1901-1955 period.
- 29

30 18. 2357, 8: The conclusion concerning a weak solar influence depends on other predictors that are

31 collinear with the solar variability (GHG and AMOI?). Does solar influence increases when AMOI is

32 deleted from predictors? Is not a long time solar variability also collinear with increasing GHG? The

33 collinearity makes generally the interpretation difficult.

- Indeed, collinearity is something that statistical analysis may struggle with interpretationally. We have tried to somewhat alleviate this problem by including results for individual sub-periods, but this only helps in some cases (e.g. solar and GHG predictors are only weakly correlated in the 1956-2010 time span, unlike in the 1901-2010 period, so distinguishing between their influences is easier). Not always is complete and reliable separation possible, particularly for predictors that are actually tied in the physical sense (such as SO and PDO indices). We have therefore at least tried to highlight the potential problems in the discussion; their identification should also now be easier with the inclusion
- 41 of the correlation matrices in the revised manuscript.

1 **REVIEWER #2**

2

3 This is a review of the manuscript titled "Imprints of climate forcings in global gridded temperature

4 data" by J. MikšovskÃi, E. Holtanová, and P. Pišoft. This work breaks down gridded temperature

5 anomalies into their components attributed to climate variability using multiple linear regression.

- 6 Overall, this paper is well written and a useful addition to the published literature on this
- 7 technique and subject and I recommend publication after revision and some minor copyediting.
- 8 General Comments:

9 1) The decision not to include aerosol forcing in the model leaves this work somewhat lacking. The

10 authors briefly discuss the reasoning for this (i.e. the complexities involved in decoupling the GHG

11 emissions from the aerosol forcing). Can aerosol forcing, as calculated by one of the CMIP5 models

12 be used for this purpose?

13 The CO2 equivalent GHG forcing used in this study needs to be described further including a

14 justification for as to why this was used when the header information for these files state that CO2

- 15 equivalent is for informational purposed and not a CMIP5 recommendation.
- 16 The suggested use of GCM-generated forcing (e.g. from CMIP5 model(s)) would perhaps be possible 17 in theory, but problematic in praxis for several reasons. First, unlike the other predictors, aerosol 18 forcings are only known with substantial uncertainty, stemming from limited knowledge of both the 19 related emissions and their effects. This uncertainty would have to be factored into the calculations 20 somehow. A single model would likely not be sufficient to provide enough data to do this - an entire 21 ensemble of GCMs would have to be employed. In terms of interpretation, a question would arise of 22 what is the contribution of the numerical simulations (and their errors and biases) to the results. The 23 fact that GCMs do create their own temporal trajectory of the climate system would also pose a 24 problem. All in all, our analysis would have to use a substantially different methodology and data to 25 combine the GCM simulations and regression analysis, and it would have to refocus substantially 26 from our intended goal.
- 27 As for the CO_2 -equivalents not being CMIP5 recommendation, we used this series because the
- concentrations of long-lived greenhouse gases can be considered roughly the same across the globe,
- 29 unlike the forcings. However, as shown in our reply to Rev#1's comment (C) and illustrated in Fig. R4,
- almost identical temperature responses would arise from application of the series of actual global
 forcings, including the effects of aerosols (or, for instance, from using just CO₂ concentrations alone).
- forcings, including the effects of aerosols (or, for instance, from using just CO₂ concent
 The specific form of the forcing/concentration series should therefore not matter.
- 33 **CHANGE MADE:** We expanded the discussion of the ramifications of our choice of the predictor 34 approximating anthropogenic effects, mostly in Sect. 5.
- 35

2) Prior studies have been criticized for using AMO as a climate driver and some critics are vehement that AMO is simply a response to volcanic activity (i.e. inclusion of the AMO is essentially "double counting" volcanic forcing). I think that further justification of the inclusion of AMO is warranted. Also, the AMO data set described by Enfield et al., 2001, considers Atlantic Ocean temperatures from 0-70N. Data near the equator will also be affected by El Nino causing some of that signal to leak into the AMO signal.

- The nature and origin of AMO seem to be still quite heavily disputed today we have tried to convey this lack of universal consensus in the Discussion, without speculating what the primary origin of
- 44 AMO actually is. I will specifically mention the possibility of volcanic forcing in the revised manuscript

1 (with a reference to the paper by Knudsen et al., 2014). However, I do not think that AMO can be 2 dismissed as a mere product of volcanic activity at this point (and in our analysis, the predictor 3 representing volcanism explains only about 9% of AMOI's total variability, so their roles as 4 explanatory variables overlap just mildly).

Regarding the AMO-ENSO relation in the equatorial area: This certainly is a possibility as far as the
general dynamics of these phenomena goes. However, in our case the respective predictors are only

7 weakly correlated (r = -0.07 over the 1901-2010 period), so the leakage should be very limited.

8 **CHANGE MADE:** In the Discussion, we expanded the section devoted to AMO, including an additional 9 reference regarding the possible effects of natural external forcings (Knudsen et al., 2014).

10

3) One way to determine possible co-linearities between predictor variables is the use of
 conditional regress analysis. I suggest choosing a sample model framework and run conditional
 analysis to highlight possible co-linearities.

14 **CHANGE MADE:** As described in reply to REV#1's comment (7), we have now included values of pair-

wise correlations between the predictors as well as values of variance inflation factors, in the form of
 Tables 1 and S1. These should provide a clearer picture of the co-linearities.

17

4) This manuscript does an excellent job of breaking down the model results but I feel the
 manuscript could go into more detail as to why there are differences between the simulations,

20 observational temperature records, and the global impact of these results. For instance, can

21 certain predictor variables explain observed floods or drought conditions in certain parts of the

22 world?

Regarding the interpretation of the differences detected, our options are somewhat limited – we can only analyze the final data products, but to properly assess the reasons for their behavior (and not just speculate), experiments with the respective data-creation engines would have to be made. Such

task is hence probably better left to the teams behind the individual datasets.

27 Regarding the impacts of the observed links: Such an extension would be interesting, but it would go28 substantially beyond the scope of our (already quite sizable) manuscript. Many of the local impacts

would be worthy of a separate paper of their own; many such analyses already exist (some – but
 certainly not all – referenced in our manuscript).

31

5) There is no figure comparing the various temperature time series observations to the sum of the explanatory variables. I think this would be a very important contribution to the manuscript to 1) Allow readers to directly compare to similar studies and 2) To see how well the reconstructed temperature time series agrees with the data. This is especially important for the recent "temperature hiatus" time period (if it exists) which has received a lot of attention recently.

CHANGE MADE: I added a new figure (Fig. S4) to the revised version of the Supplement, showing individual global temperature signals and their approximation by the regression model (I did not include it in the main paper itself, as it takes up almost two extra pages; the figure is however referenced from Sect. 4.2 of the main text). We do not specifically discuss the potential "hiatus" (or any other period-specific variations), but our regression fits do not seem to exhibit a greater error for the 21st century than they do for the rest of the series (though to properly assess this, data all the

43 way to 2015 would have to be analyzed).

6) There is no discussion of climate feedbacks and how inclusion of climate feedbacks may influence the results presented here. If they're not to be included in them model framework then

3 some justification is needed.

I agree that feedbacks are an important part of the climate dynamics. However, their proper treatment would be difficult in the purely statistical frame of our analysis, as they are generally something that is induced as a dynamical response; most comparable regression-based studies do not consider them. Honestly, I am not sure how this could be reliably done (besides perhaps analyzing a temporal vector capturing the evolution of the explanatory variables; maybe taking inspiration from Rypdal, 2015). On the whole, assessment of feedbacks is perhaps best left to the dynamical simulations of the climate system, or to studies that target this issue specifically.

11

12 Minor Comments:

P2342, line13. Should specify these are surface temperature data sets. Surface data sets show
 different long term trends than the satellite data which measure tropospheric temperatures.

- 15 **CHANGE MADE:** The nature of the analyzed temperatures as near-surface data is now explicitly 16 mentioned in the Introduction and Data sections.
- 17

Figures: In general, the figures are often very small and it is difficult to see finer details, especially figs 5 and 6. If there's any way to increase the size of the figures, perhaps by breaking them into separate panels, it would be useful.

- 21 The readability of the Figs. 5-7 will hopefully improve in the non-discussion version of the journal in
- the discussion paper, they were squeezed onto a landscape-positioned paper, making them smallerthan intended.
- 24

25 Figure 3: In a printed version of the manuscript, it is difficult to distinguish between the colored

26 lines used for land and land+ocean curves and within each category. The shades of blue for the

27 land+ocean curves are hard to tell apart and same can be said for the land curves. It is easier to see

28 these differences on an electronic version of the manuscript but a few people still like to read hard

- 29 copies.
- 30 **CHANGE MADE:** I have extended the range of colors in both groups of lines (land & global); the lines
- 31 pertaining to individual datasets should now be easier to identify.
- 32

1 References:

- Knudsen, M. F., Jacobsen, B. H., Seidenkrantz, M. S., and Olsen, J.: Evidence for external forcing of the
 Atlantic Multidecadal Oscillation since termination of the Little Ice Age, Nature Communications,
 5, 8, doi:10.1038/ncomms4323, 2014.
- 5 Rypdal, K.: Attribution in the presence of a long-memory climate response, Earth Syst. Dynam., 6,
 6 719-730, doi:10.5194/esd-6-719-2015, 2015.
- Sato, M., Hansen, J. E., McCormick, M. P., and Pollack, J. B.: Stratospheric aerosol optical depths,
 1850-1990, J. Geophys. Res.-Atmos., 98, 22987-22994, doi:10.1029/93jd02553, 1993.
- Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., Delaygue, G.,
 Joos, F., Krivova, N. A., Muscheler, R., Otto-Bliesner, B. L., Pongratz, J., Shindell, D. T., Solanki, S. K.,
 Steinhilber, F., and Vieira, L. E. A.: Climate forcing reconstructions for use in PMIP simulations of
 the last millennium (v1.0), Geoscientific Model Development, 4, 33-45, doi:10.5194/gmd-4-332011, 2011.
- Wu, S., Liu, Z. Y., Zhang, R., and Delworth, T. L.: On the observed relationship between the Pacific
 Decadal Oscillation and the Atlantic Multi-decadal Oscillation, J. Oceanogr., 67, 27-35,
 doi:10.1007/s10872-011-0003-x, 2011.
- Zhang, R., and Delworth, T. L.: Impact of the Atlantic Multidecadal Oscillation on North Pacific climate
 variability, Geophys. Res. Lett., 34, 6, doi:10.1029/2007gl031601, 2007.



3 Fig. R1. Local Pearson correlation coefficient between GISTEMP temperature series with 1200 km

4 and 250 km smoothing over the period 1901-2010. The grey areas correspond to locations with

5 insufficient amount of data, i.e. less than 90% of monthly values available over the target period

6 (note that for most gridpoints, this lack of data is related to the version with 250 km smoothing).



Fig. R2. Regression-estimated temperature responses to various explanatory variables, calculated for the 250 and 1200 km versions of the GISTEMP dataset, employing the same configuration of predictors as in Fig. 5. The missing data mask pertaining to the 250 km version of the dataset was

5 applied in both cases.



2

3 Fig. R3. Regression-estimated GISTEMP global temperature responses derived from the original 8-4 predictor setup (blue) and from a configuration employing only GHG amounts, volcanic aerosol 5 optical depth and AMOI as predictors (white). The analysis setup is otherwise identical to that in Fig.

6 4.



1 2

Figure R4. Regression-estimated global GISTEMP temperature response obtained for five different
 forms of predictor approximating anthropogenic forcing (from the left):

- 5 CO₂-equivalent concentration of GHGs controlled by the Kyoto protocol (i.e., the original
 6 predictor used in the manuscript)
- 7 CO₂-equivalent concentration representing all anthropogenic effects
- 8 Total anthropogenic forcing, expressed directly in W.m⁻² (i.e., not a CO₂-equivalent)
- 9 CO₂ concentration (ppm) (i.e. carbon dioxide only, not GHG equivalent)
- 10 Time (i.e. time index was used as predictor instead of actual forcing)