Reviewer #5 Evaluations:
Recommendation: Return to author for major revisions
Grammar improvements needed: No
Commentary: No
Willing to review a revision: Yes
Do you have a potential conflict of interest?(Required): No
Annotated: No

Reviewer #5 (Comments to Author):

Overview:

The author attempts to estimate a lower bound for the uncertainty error bars associated with the global temperature projections of current Global Climate Models (GCMs). To do this, he first presents a simple, semi-empirical analytical model for the global temperature response to greenhouse gas concentrations. According to this model, global temperature trends are linearly related to greenhouse gas concentrations. The author shows that the output from this simple linear model is a reasonable approximation of the global temperature projections of current GCMs. He then uses this observation to argue that the appropriate framework for determining the error bars is to use "linear propagation of errors".

Specifically, he argues that the uncertainty errors for each year of the projection should be added to the uncertainty errors of the preceding years. This approach leads to a fairly rapid increase in the magnitude of the error bars. He then estimates a lower bound for the uncertainty associated with a single year from reviewing some of the literature assessing the modelling of clouds in GCMs. When this lower bound is plugged into his error propagation approach, the error bars rapidly increase to encompass a range of {plus minus}15{degree sign}C by the end of a century. This proposed uncertainty range is so large that it would make all the global temperature projections of current GCMs effectively worthless. The author made a similar argument in an earlier paper which he references, i.e., Frank, 2008.

In my opinion, his use of the "linear propagation of errors" is inappropriate and invalid. His proposed error bars are unrealistically large and lead to incorrect conclusions.

Having said that, I do believe that there are several good points and arguments made in the paper which should be of interest and relevance for the climate modelling community as well as those using the results of climate models - e.g., the distinction between accuracy & precision; the fact that climate model uncertainties are typically described in terms of intra- and inter-model comparisons; the fact that there has been relatively little effort made into quantifying climate model uncertainties in terms of observation-based measurements; the differences between modelled and observed clouds.

If the author were to drop his use of "the linear propagation of errors" and restructure the article to focus on these more justifiable points and arguments, I believe there is potentially a paper which could be important and worthy of publication. However, as it stands, with the focus on these unrealistically large and inappropriate error bars, the paper is not publishable.

I appreciate that the author has been promoting this argument about using "the linear propagation of errors" for GCM projections since at least Frank, 2008. So, he might not be prepared to change tack, as I suggest. However, if he does, I think with some restructuring and refocusing of the material, and some additional discussion of the literature on comparisons between models and observations (such as the references included in this review), there could be enough material for a good paper.

With that in mind, I recommend the article should be "returned to the author for major revisions".

Below are some more detailed comments on each of the sections in the paper.

Comments on citation style

Throughout the paper there are several inconsistencies in how references are cited. For instance, on lines 61-62, the references include the initials and surnames of the lead author, sometimes with full stops after the initials and sometimes not: "[R. Knutti et al., 2008; D M Smith et al., 2007; D.A. Stainforth et al., 2005]". But then, on line 63, only the lead author's surnames are included: "[Mu et al., 2004; Murphy et al., 2004]" and on lines 108-9, it almost seems a different approach is taken for each reference! I.e.: "[Collins, 2007; J Curry, 2011; J A Curry and Webster, 2011; Hegerl et al., 2011; D.A. Stainforth et al., 2007]".

In the references section itself (Section 5), there are also inconsistencies, e.g., some references include a doi index, while others don't. Some use journal abbreviations, while others don't. Also, the references to books, reports and chapters in books don't always seem to provide all the relevant information. I would include urls for some of the references which don't have a doi, e.g., Frank, 2008 and Haynes, 2012.

I haven't checked if ESS has a specific citation approach that they require, but at any rate, whatever approach is used, the author should stick to one.

Comments on the Table of terms or acronyms

The author repeatedly assigns the term "general circulation (climate) model" to the acronym "GCM" throughout the paper, including in the glossary. This used to be a common usage, but it is now a rather outdated assignation, as the GCM acronym is now usually given for the term "Global Climate Model".

There is a bit of history associated with this acronym, e.g., see Edwards, 2011. Basically, the earliest climate models only considered the atmospheric circulation, so the term "General Circulation Model" was understood to refer to a climate model. However, starting in the 1980s, modellers increasingly began distinguishing between "Atmospheric General Circulation Models" (AGCM) and "Oceanic General Circulation Models" (OGCM) and the latest "coupled models" which included both an atmospheric module and an oceanic module.

As computational capabilities advanced, these coupled "AOGCMs" started to become the standard in climate modelling. Moreover, climate modellers began introducing other components into their models, e.g., sea ice, vegetation, etc. So, the climate models no longer were confined to just "general circulations". However, because the acronym GCM was still widely understood among the community to refer to "climate models", modellers decided to keep the acronym, but change the assignation to "Global Climate Model". As a result, the acronym GCM is now generally understood to mean "Global Climate Model" and the term "general circulation model" is usually prefixed with the specific type of "general circulation model", i.e., an AGCM, an OGCM or a coupled AOGCM.

In the paper, the author seems to be focusing on coupled AOGCMs such as the CMIP3 and CMIP5 Global Climate Models, so I would recommend using the "Global Climate Model" terminology (or at least adding in the "coupled AO..." prefixes wherever appropriate if he chooses to stick with the older terminology).

Also, the author uses the term "global average surface air temperature" (GASAT). I think terms such as "Global Mean Surface Temperature" (GMST), "Surface Air Temperature" (SAT) or "Global Mean Surface Air Temperature" (GMSAT) are more commonly used. I would probably recommend using GMST.

The author uses the term "greenhouse gas" to refer to "CO2, CH4, N2O, various chlorofluorocarbons, etc." What about H2O? Or is he intentionally using the term to refer to non-H2O greenhouse gases? If so, this should be mentioned.

On a minor point, I would probably use a lower case "n" for the "ARN" and

"CMIPN" acronyms and accompanying explanation, to highlight that the "N" is a variable index. Also, in the explanation for RCP, he uses "5AR" to refer to the 5th Assessment Report instead of the "AR5" acronym.

Some of the terms in the table seem unnecessary to me - does the reader gain much benefit by having to refer back to the glossary for "Lag-1", "theory-bias" or "wve"? These are only used a few times in the paper and it would probably be easier for the readers to have them repeated or explained whenever they are used.

Comments on Section 1

The author makes some very important points in this introduction - specifically, the distinction between precision and accuracy; the fact that climate model uncertainties are typically described in terms of intra- and inter-model comparisons; and the fact that there has been relatively little effort made into quantifying climate model uncertainties in terms of observation-based measurements.

However, a major focus of the current draft of the paper is his argument that the error bars associated with each year should be summed cumulatively to those of the years beforehand. This is the main reason why he (incorrectly!) ends up with such dramatically wide error bars. As I will discuss in "Comments on Section 2.4", this argument is inappropriate and invalid for estimating the uncertainties associated with global temperature projections.

Hence, the discussion of this argument in the Introduction, e.g., most of lines 81-105, should be removed.

He could simply remove that discussion, or perhaps replace it with a more detailed discussion of some of the recent attempts to compare climate model outputs to observed data.

For instance, in Soon et al., 2015, we found that the CMIP5 hindcasts did not consider any of the "high variability" estimates of Total Solar Irradiance in their "natural forcing" datasets. We also showed that, after correcting for urbanization bias in the "observed" global temperature records, the CMIP5 hindcasts were unable to accurately reproduce the global temperature trends since 1880 regardless of whether they used their "natural forcings" or their "anthropogenic plus natural forcings". Anagnostopoulos et al., 2010 have shown that the current climate models have done a poor job of reproducing both temperature and precipitation trends at local and regional levels.

Also, even climate modellers have publicly expressed their concern about various aspects of the current models. For instance, Palmer, 2016 points out in his Figure 1 that there is an unacceptably large range of absolute global mean surface temperatures as simulated by the CMIP5 models. And Fyfe et al., 2013 has criticised the models for "(overestimating) global warming over the past 20 years".

Below are some other comments on this section, which are relatively minor:

• Line 49: Instead of saying "by about 3 Celsius", use the degree symbol, i.e., 3{degree sign}C, and maybe give a range of predicted values too.

• Line 52: "...is central to the question of causality..." is quite ambiguously phrased - causality of what, exactly? Maybe say something like "...is central to this prediction..."

• Lines 124-127: The author claims that he won't attempt to survey or address the physics of climate, or to model the terrestrial climate in any way. Instead, he will focus "strictly on the behaviour and reliability of climate models alone". This seems a reasonable approach to keeping the article short, focused and on topic. Yet, in Section 2.1, he then proceeds to describe a simple model of global mean surface temperature. As I will discuss below, I do not believe the analysis in Section 2.1 is strictly necessary, and so if he decides to remove Section 2.1, then lines 124-127 are good. But, if he keeps Section 2.1, then lines 124-127 should probably be rephrased accordingly.

• Lines 129-182: In my opinion, the outline of the structure of the paper is probably too long. A short outline of just a paragraph or two should be sufficient. He doesn't need to go through all the details of what happens in the rest of the paper (this isn't a review or thesis!). The reader will find out these details in a few pages & it makes the points quite repetitive.

• As a side note, among the climate modelling community there is a lot of consternation over the usage of the terms "prediction" and "projection". Climate modellers argue that they are only estimating "projections" of future climate conditions under a given emissions scenario and that they are not "predicting" which emissions scenario would eventually occur. So, you will often hear climate modellers objecting to their projections being described as predictions. Personally, I think these objections are usually undeserved and I think that in this case the author has used the terms in a manner which is technically correct in that he says, "(the IPCC) has predicted that...unabated... could increase..." However, it is worth realising that this prediction vs projection debate exists, and being careful in how you use words such as "prediction" and "predict" in this context.

Comments on Sections 2.1 and 2.2

The author presents a simple, semi-empirically derived model of the expected global SAT increase from an increase in GHG. This is the same model he presented earlier in Frank, 2008. He then claims (or assumes?) that the physics underlying his model are essentially the same as the physics built into GCMs (and in Frank, 2008 he also suggested that his theoretical model were reasonable substitutes for GCMs in terms of estimating global SAT projections without involving the considerable computational expense of a GCM). This appears to be a key assumption in his argument, yet it is flawed.

There are a number of different theoretical models which have been developed to estimate the expected global SAT response to GHG. Some of them yield rather straightforward expressions that can be evaluated using a pen and paper, such as the author's equation 6; Arrhenius, 1896's classical one-dimensional model; or Benestad, 2016's conceptual "toy model". However, the radiative physics modules used by GCMs are based on a different approach, which is more computationally expensive, i.e., the incorporation of so-called "infrared cooling models" [e.g., Rodgers & Walshaw, 1966; Stone & Manabe, 1968; Fels et al., 1991]. While all of these approaches are often assumed to be equivalent, they frequently invoke different assumptions and/or approximations, and can sometimes lead to (often subtly) different conclusions.

For instance, according to the author's model, the effect of increasing [CO2] on global SAT is a consequence of the absorption of outgoing IR, and therefore mainly determined by Beer's law. On the other hand, Benestad's model suggests that the effect is a consequence of the changes in the mean emission altitude combined with the assumption of a relatively constant lapse rate. Meanwhile, the infrared cooling models used by the GCMs split the atmosphere into a series of 3D-grid boxes, and for each grid box (in each time step) determine the net absorption and emission fluxes through each grid box for each range of wavelengths [e.g., see the "infrared cooling models" references mentioned above].

Obviously, absorption and emission processes are closely related (in many senses they are two sides of the same coin), and the physics used by each of these models has a lot of common ground. The three approaches described above (i.e., the author's equation 6, Benestad's model & the GCM approach) also agree that increasing [CO2] should lead to warmer global SAT. However, this doesn't mean that they are substitutes for each other and/or that they produce the exact same results. Indeed, Benestad, 2016 explicitly notes that his toy model is not a replacement for GCMs.

For this reason, the author should not assume that the physical implications suggested by his model necessarily also apply to the GCMs.

Having said all of that, from what I can tell, the main relevance of Section 2.1 to the article seems to be his finding in Section 2.2 that the global SAT projections are reasonably well approximated as simple linear extrapolations of the increase in [GHG]. If that is the main point that is being made, then this could presumably also be demonstrated by applying a linear least squares fitting between the GCM projections and the changes in GHG forcing. In my opinion, this would be a simpler and more straightforward approach to reaching the same conclusion.

Indeed, in lines 124-127 of Section 1, the author explains how he will not try to survey or address the physics of climate, or to model the terrestrial climate in any way. Yet, this seems to be exactly what he does with Section 2.1!

Therefore, I recommend the author considers how important his equation 6 model is to this study:

• If he is just using it for the linearity result, then I would remove Section 2.1 and replace his analysis in Section 2.2 with a statistical analysis of the relationship between GCM projections and GHG forcing using, e.g., linear least squares fitting.

• If he wants to use this opportunity to point out that this is consistent with his model, he could do this with a simple sentence, e.g., "This is consistent with an earlier analysis by the author [Frank, 2008]" or something similar.

• If he wants to use this article to highlight his 2008 model and argue that its global SAT projections are a reasonable substitute for those of the GCMs, then in my opinion, he should carry out a far more detailed compare/contrast assessment of the similarities and differences of his model to the GCMs (as well as to other models such as Benestad, 2016). Indeed, I would argue that this would probably be better carried out as a separate study, and should include a more detailed analysis than the original Frank, 2008 paper.

At any rate, regardless of whether his semi-empirical model or a linear least-squares approach (or something else) is used, he should add some analysis of the residuals from the linear fits. Indeed it is quite ironic that the author is criticising the prior discussion of GCMs for not adequately assessing the residuals between the model output and observations, yet he does not include any such assessment of his claim of linearity. Even just by eyeballing Figures 2 and 3, it seems to me that there are noticeable differences between his linear PWM projections and the GCM projections.

I still think the deviations from linearity are probably small enough to justify saying something like, "...the GCM projections are reasonably approximated in terms of a linear relationship..." But, the deviations from linearity (i.e., residuals with regards to the linear model/fit) should still be discussed and statistically analysed. Also, on lines 403-5, he shouldn't claim the projections "...are just linear extrapolations..."

If the author goes with the lesser - yet better justified - claim that "the projections can be reasonably well-approximated through a linear relationship...", this is still quite a powerful and striking finding which should be of interest and relevance to the scientific community.

Comments on Section 2.3

In this section, the author uses the analysis of several other studies (in particular Jiang et al., 2012 and Lauer & Hamilton, 2013) to conclude that there are systematic biases in the current GCMs' simulated total cloud fraction estimates.

In a sense, this is an easy target for criticising GCMs since even climate modellers agree that current models are only able to at best approximate cloud coverage. Part of this is due to the fact that cloud formations are orders of magnitude smaller than the horizontal dimensions of a typical grid-box. As a result, current climate models have to resort to "parameterization" of the average cloud amounts and types in a grid box rather than the explicit modelling of cloud formations. However, the models still have problems in accurately simulating the observed values, as demonstrated by the various studies referred to in the section.

I think the author does a reasonable job of arguing that there appear to be systematic biases in the current GCM cloud modelling. This observation isn't new - after all, he is basing his analysis on the observations of several previous studies. However, I thought Figure S10 in the SI and Figure 5 were interesting, and might be shocking to some readers who are unfamiliar with the cloud-modelling problem.

I would make some changes, though:

• The author seems to be implicitly assuming that the satellite datasets are perfect and without any systematic biases themselves. This is a big assumption which underlies much of the ensuing discussion. Personally, I think it is a reasonable working assumption to make, but it should be made explicit. There may well be systematic biases in both the models and the satellite datasets. The editor may have more insights into this, since he is the lead author of the Jiang et al., 2012 study. If the author wants to use the assumption that the satellite datasets are unbiased as the basis for his analysis, he should explicitly state that he is doing so. Note that it is not sufficient to comment on the "high resolution" of the A-Train dataset (line 415), since that refers to the precision, not the accuracy of the dataset (which he correctly explains in the Introduction are not the same thing).

• With that in mind, the sentence on line 446 saying "That is, the simulation is inaccurate", and also the "...without an improvement in TCF verisimilitude" on lines 495-6 should be toned down or else removed.

• Also, doesn't Jiang et al., 2012 argue that "...more than half of the models show improvements from CMIP3 to CMIP5 in simulating column-integrated cloud amount..."? This seems to be slightly at odds with the discussion on lines 490-496 on the comparison between AMIP1 and CMIP5. The reasons for the different conclusions should either be discussed, or else the discussion in lines 490-6 modified accordingly.

• Figure 5 is difficult to properly assess without Figure S10 in the SI. Therefore I recommend moving Figure S10 back into the main article, so that the residuals can be directly compared to the model simulations and satellite observations themselves.

• Lines 485-488: As far as I know, the main difference between the GISS e2r simulations and the GISS e2h were to do with the ocean models used (Russell and Hycom ocean models respectively - see https://data.giss.nasa.gov/modelE/cmip5/) and had nothing to do with the atmospheric components. So, it is not surprising that they produced similar results with regards to cloud modelling. Therefore, I don't think this analysis adds to the discussion and lines 485-8 should probably be removed.

• In lines 419-425, when discussing the Jiang et al., 2012 study, the author states, "CMIP5 TCF hindcasts comprising 25-year (1980-2004) annual projection means were compared to the corresponding A-train observations". When reading this, I mistakenly assumed that "corresponding" meant the same 1980-2004 period, and I suspect many readers would also. It was only when checking Section 6 of the SI (which begins on page 13 of the SI, not 12 as the Table of Contents claims) that I realised this wasn't the case. The reason why Jiang et al., 2012 used non-overlapping periods was reasonably well-justified, i.e., the models didn't have any significant trends in clouds or water vapour over those periods. However, this has implications for the author's analysis in Section 2.4 (see below). So, the use of "corresponding" is misleading, and this description should be rephrased.

• Also, on a minor point, on line 413, should it read "...have been carried out" and on line 415, is the use of the word "acute" a typo? Maybe he meant "accurate"?

Comments on Section 2.4

I disagree with much of the analysis in this section. In my opinion, much of this section needs to be variously removed, rewritten and/or replaced. As the article is currently drafted, the most dramatic conclusions of the paper are based on this analysis. However, as I mentioned at the start, I do believe there is much of value in the paper, and if the author is prepared to redraft the paper with a different emphasis, there is probably a paper worthy of publication in here.

With that in mind, here are my main problems with Section 2.4 as it currently stands:

• The author is basically making the same argument that he has already made in Frank, 2008, except using an "annual LWCF error" of {plus minus}4 W/m2 instead of his earlier estimate of {plus minus}2.8 W/m2, and updating his analysis from CMIP3 to CMIP5. I strongly disagree with the conclusion of Frank, 2008 that the uncertainty limits associated with GCM projections are a very dramatic {plus minus}100{degree sign}C/century. However, if the analysis in Frank, 2008 were correct, then its conclusions were already so dramatic that increasing the error bars even further wouldn't make much difference. That is, IF the analysis in Frank, 2008 were valid, then this update would be unnecessary and as a result not worth publishing. But, as I will discuss below, I actually don't think that analysis was correct, and so recommend it be dropped from the paper instead of updated.

• With regards to his estimate of the "annual LWCF error", his rationale for specifically choosing {plus minus}4 W/m2 is very poorly explained or justified. I had to re-read Lauer & Hamilton, 2013 several times before realising that he was probably referring to this brief sentence in the text - "For CMIP5, the correlation of the multimodel mean LCF is 0.93 (rmse = 4 W m-2)..." (p3833, Lauer & Hamilton, 2013). Is that rmse value the basis for the author's "annual LWCF error" estimate? If so, then the author is being far too cryptic in his rationale. He is also being somewhat misleading, as most readers (including me) would have assumed from his matter-of-fact citation of Lauer & Hamilton, 2013 that this was one of their key findings. If this is indeed the basis for his estimate, he should make it clearer exactly where in the paper he took his estimate from and let the reader know that this was only a passing comment in Lauer & Hamilton, 2013. If it's not the basis for his estimate, then I couldn't find anything else, so he definitely needs to be more explicit in his reasoning!

• Assuming that was the basis for his estimate, I am not sure that this is actually an appropriate value. From what I can tell, the rmse of 4 W m-2 refers to a spatial correlation over the entire globe. From Figure 2 of Lauer & Hamilton, 2013, it can be seen that the values for some regions are overestimated, while other regions are underestimated. However, since the author is estimating the uncertainty associated with *global* SAT, this estimate of the regional variability in the biases is less relevant. I might be wrong in this interpretation of the 4 W m-2 figure, since I am not very familiar with the satellite-based cloud datasets, but if so this would again highlight how the author has failed to *clearly* explain the rationale for his estimate.

• Even assuming that the estimate is correct, this does not in itself mean that this applies to the uncertainties over global mean surface temperature trends, which is what he is attempting to estimate. First, as he notes in the SI, Jiang et al., 2012 didn't find any significant trends in the models for clouds or water vapour over the 1980-2004 period. If the models didn't show significant trends in clouds over this period, then the associated error due to clouds on GMST shouldn't have changed much either. Second, the models still showed trends in GMST over the same period (e.g., Fyfe et al., 2013), i.e., the GMST trends of the models over that period were not heavily influenced by the trends in clouds. Regardless of whether you consider this a problem with the current models or not, this indicates that the 4 W m-2 uncertainty in LWCF only at best indirectly influences the model GMST trends. Third, as Palmer, 2016 notes in his Figure 1, there is considerable disagreement between the various CMIP5 models on the absolute global mean surface temperature, but the models are relatively consistent in the trends for global mean surface temperature anomalies. Perhaps one factor could be related to this uncertainty in LWCF, e.g., the different LWCF estimates alter the model's estimates for the absolute GMST. But, in that case, the associated uncertainty would apply to the absolute GMST, not the GMST trends.

• Additionally, if there is a systematic bias in the way the GCMs model clouds, as opposed to the biases being merely stochastic in nature, then the errors are definitely non-Gaussian in nature and the errors should be treated as systemic, not stochastic.

More generally, the author's argument seems to be that because climate model projections are step-wise in nature (lines 101-105), he believes that the errors associated with the GMST projections for each year should be cumulatively summed together with the errors from previous years using his equation 2 (lines 89-94). This is incorrect for several reasons.

Yes, climate models projections are step-wise in nature. However, while the global surface mean temperatures in a given step do have an influence on the projections of the next steps, this is generally an indirect influence (through feedback loops, etc) and is definitely not the only (or even a major) factor. This can be demonstrated using a simple thought experiment: if the GMST of one year was solely a function of the GMST of the previous year, then GMST would be effectively constant over time. Instead, the global warming projected by the CMIP5 models is mostly a consequence of rising GHG concentrations.

I agree that GMST of the previous years have SOME influence on the next in the models through feedbacks (changes in sea ice, atmospheric circulation, etc). This can be seen from the difference between "Equilibrium Climate Sensitivity" estimates of the long-term expected GMST increase from a step doubling in CO2 and "Transient Climate Response" estimates derived using simulations involving a gradual increase in CO2. However, it is only an indirect influence, and it is not a simple one either.

So, the author's use of equation 2 for estimating the uncertainty of GMST projections is invalid.

Comments on Section 3

In the Summary and Discussion section, the author focuses almost entirely on his analysis in Section 2.4. However, in my opinion, this is probably the weakest part of the paper.

There are many reasons for questioning the reliability of the current GCMs' GMST projections including the uncertainties over their ability to model clouds, as well as the references I mentioned in the Introduction. However, the author's application of equation 2 is inappropriate and leads to an overly dramatic and unrealistic dismissal of their reliability. In my opinion, the author should drop his focus on his equation 2 (which he has been doing since Frank, 2008). Instead, he should focus on the more justifiable concerns over the current GCMs, such as the ones I mentioned in the introduction.

I also think he should try to be more constructive in his criticism of the current GCMs. Even climate modellers agree the current models have many flaws (e.g., Fyfe et al., 2013; Palmer, 2016). However, this doesn't mean that they are useless or can't be improved. The author should try to provide practical and realistic recommendations for how the modelling community can improve the GCMs rather than implying that they are useless.

If the author could generate more realistic lower bound estimates of the uncertainty in the accuracy of the projections, this could be of help to the modelling community, as well as the rest of the climate science community.

However, the estimates he has calculated in this draft are completely unrealistic, mostly because of his inappropriate use of his equation 2.

Reviewed by:
Dr. Ronan Connolly
Independent scientist, Dublin, Ireland

(I have chosen not to remain anonymous to the author)

Additional articles referred to in this review

The references below are the ones mentioned in my review above, which the author has not already referenced in his paper. Whether or not these should be added to the paper depends on how the author rewrites the paper.

• Anagnostopoulos, Koutsoyiannis, Christofides et al., 2010. A comparison of local and aggregated climate model outputs with observed data. Hydrol. Sci. J., 55:1094-1110. doi:10.1080/02626667.2010.513518
• Arrhenius, 1896. On the influence of carbonic acid in the air upon the temperature of the ground. Phil. Mag. J. Sci., 41:237-276. doi: 10.1080/14786449608620846
• Benestad, 2016. A mental picture of the greenhouse effect. Theor. Appl. Climatol., in press. doi: 10.1007/s00704-016-1732-y
• Edwards, 2011. History of climate modeling. WIREs Clim Change, 2:128-139. doi: 10.1002/wcc.95
• Fels et al., 1991. Infrared cooling rate calculations in operational general circulation models: Comparisons with benchmark computations. J. Geophys. Res., 96(D5):9105-9120. doi: 10.1029/91JD00516
• Fyfe, Gillett & Zwiers, 2013. Overestimated global warming over the past 20 years. Nature Clim. Change, 3:767-769. doi:10.1038/nclimate1972
• Palmer, 2016. A personal perspective on modelling the climate system. Proc. R. Soc. A, 472:20150772. doi:10.1098/rspa.2015.0772
• Rodgers & Walshaw, 1966. The computation of infra-red cooling rate in planetary atmospheres. Q.J.R. Meteorol. Soc., 92:67-92. doi: 10.1002/qj.49709239107
• Soon, Connolly & Connolly, 2015. Re-evaluating the role of solar variability on Northern Hemisphere temperature trends since the 19th century. Earth Sci. Rev., 150:409-452. doi: 10.1016/j.earscirev.2015.08.010
• Stone & Manabe, 1968. Comparison among various numerical models designed for computing infrared cooling. Mon. Wea. Rev., 96:735-741. doi: 10.1175/1520-0493(1968)096<0735:CAVNMD>2.0.CO;2