This is a fairly straightforward manuscript. Various of the authors have previously developed a new stomatal conductance model, obtained PFT-specific parameters for the model (g1), implemented the stomatal conductance model in the CABLE land surface model, and tested the CABLE model in offline (uncoupled) simulations. Here, the authors conduct climate simulations and document the impact of the stomatal conductance model on the simulated climate.

We thank the reviewer for their constructive comments

1. My biggest concern is that the authors oversell the importance of their stomatal conductance model. I agree with their assertions that their model has a theoretical basis (in contrast to other more widely used empirical models) and that alone is a significant step forward. However, the authors repeatedly state that the model improves the simulated climate, while ignoring the fact that the model enhances temperature biases in many regions of the world. This is seen in the abstract, where the authors tout the improved simulation in boreal regions of the Northern Hemisphere during summer, but do not mention the degradation of the simulation in much of the rest of the world. Similar generalizations are made in the discussion and conclusions. Figures 2 and 3 clearly show Tmax and Tmin warm in much of the world because of the stomatal conductance model, and this enhances biases (e.g., North America, South America). These biases are acknowledged in the results, but then are not mentioned when the authors summarize "improvements".

We modified the abstract to reflect the reviewers concern and now note that we do not fix errors everywhere by any means. We do not agree, however, that we "degrade the simulation in much of the rest of the world". The large errors in North and South America are existing biases in ACCESS documented in detail in Lorenz et al. (2014) and likely due to several factors including atmospheric dynamics, problems with moisture advection and the cloud parameterization (all of these have been noted as possible reasons in earlier analyses). The new scheme makes these biases slightly worse, we agree and make that clear in the abstract as follows:

"We show that the new scheme reduces the latent heat flux from the land surface over the Boreal forests during the northern hemisphere summer by 0.5 to 1.0 mm day^{-1}. This leads to warmer daily maximum and minimum temperatures by up to 1.0 $^{\circ}$C and warmer extreme maximum temperatures by up to 1.5 $^{\circ}$C. These changes generally improve the climate model's climatology of warm extremes and improve existing biases by 10–20%. The bias in minimum temperatures is however degraded, but overall, this is outweighed by the improvement in maximum temperatures as there is a net improvement in the diurnal temperature range in this region. In other regions such as parts of South and North America where ACCESSv1.3b has known large positive biases in both maximum and minimum temperatures ($\sim$ 5 to 10°C), the new

scheme degrades this bias by up to 1°C. We conclude that although several large biases remain in ACCESSv1.3b for temperature extremes, the improvements in the global climate model over large parts of the Boreal forests during the northern hemisphere summer which result from the new stomatal scheme, constrained by a global synthesis of experimental data, provide a valuable advance in the long-term development of the ACCESS modelling system."

We have also made substantial edits to the Discussion to better explain the degradation in the bias over North America:

"The ACCESSv1.3 model is a skillful global climate model, performing at the top end in the evaluation of CMIP-5 models (Flato et al., 2013). While the model's overall climatology is very good, there are regional systematic biases that exist irrespective of whether the LEU or MED scheme is used. However, we note that MED does make the model worse in some regions. Perhaps the most serious bias in temperature is over North America in summer (Figs. 2, 3, 4). The biases in ACCESS using LEU reach 5-7°C in T_{MAX}, 8-9°C in T_{MIN} and 6-8°C in TXx. Using MED makes these biases worse, by ~0.5°C for T_{MAX} and T_{MIN} and by ~1°C for TXx. Given the magnitude of the error in ACCESS, we suggest that using MED does not really make the model significantly worse since it is already very poor in this region. The poor performance of ACCESS over North America, across several model versions, has been previously reported. Bi et al. (2013) show errors of 3-4°C in the mean air temperature over North America. This does not appear to be linked with low rainfall but rather, it appears linked with an underestimation of cloud coverage over North America despite the overall tendency to simulate slightly too much cloud over northern hemisphere land (Bi et al., 2013). Franklin et al. (2013) examined ACCESS1.3's simulation of clouds but did not focus specifically on North America. They did note problems with capturing convective regimes; these showed too weak a dependence on large-scale dynamics in comparison to observations. The change from the LEU to the MED scheme would not be expected to affect large-scale dynamics or how these processes affect clouds. Small decreases in ACCESS's skill in capturing North America's climate linked with the MED scheme are very likely insignificant; if the large-scale cloud fields were simulated well MED might make a small positive impact in this region but any benefits are currently swamped by the poor cloud climatology. A similar problem exists over the Indian monsoon region where there is a very serious rainfall bias (Fig. 7). The MED scheme does not add value here because there are major systematic weaknesses in the generation of the monsoon linked with larger-scale dynamics as distinct from terrestrial processes."

We also note to the reviewer that we did in fact highlighted that some biases are made worse in the conclusion:

"Maximum temperatures are generally improved, but major systematic biases in minimum temperatures remain and are in some cases made worse "

2. Rather than touting improvements, a more robust conclusion is that the stomatal conductance model decreases evapotranspiration and warms the surface air temperature in most areas of the world.

Yes, we now make this explicitly clear in the discussion as follows:

"Our results show first and foremost, that replacing LEU with MED results in an overall decrease in ET by 0.1 to 0.3 mm day^{-1} and subsequent warming signal of up to 1 °C, mostly over most parts of the world."

For the ACCESS model, this improves the simulation in boreal regions of the Northern Hemisphere during summer. In other regions, the climate is degraded.

Please see our response to your first comment.

A similar result is quite likely to happen if the stomatal conductance model is implemented in other climate models (i.e., reduced evapotranspiration and warmer surface air temperature). Is this indeed the robust outcome of the simulations?

We agree with the reviewer that the results they describe, i.e. reduced ET and a ensuing warmer surface air temperature, is indeed likely to be a robust result across models if they implemented a similar stomatal scheme and parameterisation. We come to this conclusions based on the consistent responses observed between both offline (De Kauwe et al. 2015) and online simulations (this study). Nevertheless, we would prefer to tread cautiously when speaking to how this result will play out in other climate models. As with some of the existing biases seen here with ACCESS (e.g. over N. America), climate model biases are highly unlikely to be fixed simply due to a change in stomatal conductance. Furthermore, different models assume different levels of canopy-to-boundary layer coupling (cf. Jarvis and McNaughton 1986) and as such may simulate higher or lower ET fluxes. Given all of these factors we feel it may be speculative to comment further on the exact changes different climate models may see.

If so, then that is the key result and gives guidance to other modelers about the generality of the stomatal model. The fact that the ACCESS climate is improved in

one region of the world during one season is not that significant or generalizable across climate models (given the existing temperature biases in ACCESS).

We do not agree. The reviewer has quite reasonably noted that the new stomatal conductance model is more realistic in a physical sense. We have demonstrated that it can be used in a climate model and it performs better than one of the more commonly used alternatives over large parts of the Boreal forests and offers a clear development path moving forward. We have highlighted this as well as instances where biases are degraded in the discussion. Please see our response above to your previous comments.

3. What are the reasons for the large biases in Tmax and Tmin and the diurnal temperature range? Is this related to clouds, radiation, soil water stress? How might these affect the results presented here regarding the performance of the stomatal conductance model? Might the improvements seen in boreal regions with the stomatal conductance model simply compensate for biases in clouds, radiation, etc? And likewise, the degradation seen in other regions may be because the stomatal conductance model accentuates biases in radiation, soil water, etc. The manuscript lacks a thorough discussion of biases, their causes with respect to the ACCESS model, how these affect the stomatal conductance model, and whether the stomatal model is improved or is simply compensating for these biases.

We have added a paragraph to the discussion focussed on North America and the Indian Monsoon region. Please refer to our response to your earlier comments. The ACCESS model is clearly one of the stronger models in the assessment by Flato et al. (2013) for the Intergovernmental Panel on Climate Change. Across a large range of metrics, it performs overall among the top 3 or 4 models. So, the errors in Tmax and Tmin are large in an absolute sense, but it is not known if they are large relative to other models since this analysis is not something that forms part of the Flato et al. (2013) chapter.

The actual biases in Tmax and Tmin were partly explained in the submitted manuscript. The Discussion noted "Lorenz et al. (2014) linked the biases in T_{MIN} to systematic biases in incoming longwave radiation" for example. However, there was no explanation for the biases in Tmax. This was explored by Lorenz et al. (2014) and we have added the text "these were, in part, attributed to an overestimation of evapotranspiration linked to weaknesses in the representation of land processes". The *reason* for the excessive evaporation is unknown though there are several hypotheses. We are testing one of these in this paper. A significant proportion of the total evaporation is controlled by the stomatal conductance, and hence this paper is addressing one element of the potential bias. We note this in the introduction:

"We seek to determine whether the biases identified by Lorenz et al. (2014) can be resolved in part via the parameterization of g_s."

We acknowledge we do not provide "a thorough discussion of biases, their causes with respect to the ACCESS model, how these affect the stomatal conductance model, and whether the stomatal model is improved or is simply compensating for these biases."

This is a lot harder than it sounds, and would represent a significant expansion of this paper, one that we think is well outside its scope. We also note that errors in radiation products from commonly used re-analysis products such as ERA-interim and/or one of the NCEP products would be too large to accurately diagnose errors in Tmax and Tmin. These tend to be unreliable as they relate to land processes since they assimilate land variables which break conservation of energy and water constraints. In the absence of reanalyses, global radiation data sets are averaged over too long a period to enable them to be used to diagnose daily Tmax or Tmin. There are no observations of soil moisture at a high enough temporal resolution, spatial extent, or depth, that would enable the analysis suggested by the reviewer. We can do elements of this off-line using eddy covariance observations, but this has been done by De Kauwe et al. (2015). Of course, one cannot then evaluate Tmax or Tmin since these form part of the meteorological forcing. So, we strongly agree with the reviewer that this is something that would be useful, but we are not aware of a way to approach this that would be applicable to the scales we are examining here.

4. A key result of the study is buried in the discussion and is not highlighted in the abstract or conclusion. Lines 11-13 (page 5247) state that the difference between the two stomatal models relates to the slope parameter g1 and its PFT-specific values, whereas the Leuning model uses a constant g1 (distinguishing only C3 and C4 plants). This is a critical result of the study. It suggests that PFT-specific g1 would similarly change the Leuning model, perhaps so that the two models are not distinguishable in their climate effects. This statement should (i) have some supporting analysis (it is given as a statement of fact without any documentation) and (ii) is a key documentation of the stomatal model that needs to be in the abstract and conclusion (not buried in the discussion).

We agree with the reviewer that a key difference between the two stomatal models relates to different assumed g1 values. We introduced these differences earlier in the text than the discussion, on line 15 of page 5240 and referred to the values in Table 1, we stated:

"*Values of g_1 for the CABLE PFTs (Fig. 1) are shown in Table 1. These have been compiled from a global database of stomatal conductance from 314 species across 56 field sites (Lin et al., 2015) including Arctic tundra, boreal, temperate forests and tropical rainforest biomes. Further details can be found in De Kauwe et al. (2015).*"

Prior to this, on Line 22 of page 5237, we stated:

"Medlyn et al. (2011) proposed an alternative approach, deriving a stomatal conductance scheme from optimal stomatal theory. This approach maintains the same functional form as the empirical models, but attaches a theoretical meaning to the model parameters. Parameters now represent a plant's water use strategy and thus can be hypothesised to vary across climate space. Lin et al. (2015) used a global database of g_s data to show that stomatal behaviour varied between PFTs as predicted by the optimal stomatal theory (and the leaf and wood economic spectrum), in line with the Medlyn et al. (2011) model. De Kauwe et al. (2015) tested an implementation of this g_s scheme within CABLE using flux tower observations and global offline forcing and showed a ~30 % reduction in annual fluxes of transpiration compared to the standard CABLE model across boreal, tundra and C4 grass regions."

However, we don't fully agree with the reviewer that were we to derive PFT dependent g1 values for the Leuning model we would obtain indistinguishable results. The reason for this is the two models have different responses to vapour pressure deficit (*D*). In our previous offline work (De Kauwe et al. 2015), we examined whether differences between the Medlyn and Leuning schemes were due to (a) structural differences (i.e. response to *D*) or (b) different g1 values (i.e. PFT water use). To test this, we conducted offline simulations with MED but with g1 parameters fitted to the LEU model, and showed that for the contemporary period, most of the differences are due to different g1 values (i.e. the point the reviewer is making). Nevertheless, as *D* changes with climate into the future this result may no longer be true. Specifically, conductance predictions from the Leuning model are more sensitive to higher values of *D* than the Medlyn model. On re-reading out text we acknowledge that information about these differences between the models in relation to *D* was omitted and we now add in section 2.2:

"In addition to differences in paramaterisation between the two g_S models (i.e., g_0, g_1), the models also have contrasting sensitivities to D. Further details can be found in De Kauwe et al. (2015); in short, over low to moderate values of D (< 1.5 kPa), the g_S simulated by the new scheme is more sensitive to D. By contrast, at higher values of D (> 2 kPa), g_S simulated by the default model are more sensitive to increasing D than the new model."

5. The results show a reduction in evapotranspiration and increase in Tmax and Tmin because of the stomatal conductance model. Presumably this is because stomatal conductance decreases. However, the authors never present maps of stomatal conductance to show this. A map of stomatal conductance (e.g., mid-day maximum) is essential to show readers the magnitude of the change in stomatal conductance and its geography.

Stomatal conductance was one of our original outputs, however it turned out that the outputted flux erroneously related to an incorrect flux (it was a bogus output from the UM/atmospheric component of the code which is not used anywhere, rather than the actual stomatal conductance computed from CABLE). As such, we are not in a position to produce such a map without considerable computational cost (i.e. repeating the entire piece of work). Whilst we fully acknowledge this is not an ideal scenario, we would argue that we have fully documented changes in gs in offline simulations (De Kauwe et al. 2015a) and we are building upon that work here. The changes in ET between offline and coupled simulations are consistent which would support the expectation that changes in gs were similar here.

6. Give values for g0. Is this the same for all PFTs? Is it the same for both stomatal models (Leuning and Medlyn)?

We have added this information now at the end of section 2.2 as follows:

"Additionally, for the MED scheme, $g_0 = 0$ mol H_20 m^{-2} s^{-1}, meaning that g_S goes to zero under low light and importantly, high VPD conditions. By contrast, in the default LEU scheme, $g_0 = 0.01$ and 0.04 mol H_20 m^{-2} s^{-1} for C3 and C4 species respectively"

7. The carbon cycle and NPP are mentioned in the abstract and also in the results (Fig 8). However, the methods (section 2.3) states that carbon pool dynamics was not simulated (line 9, page 5241). What, then, is shown in Figure 8?

The reviewer is correct our original text was unclear. In the version of CABLE we used, carbon fluxes were simulated (e.g. gross primary productivity, autotrophic respiration, etc.), but the dynamics of the pools, i.e. changes in biomass (foliage, wood and roots) were not. However, because LAI was prescribed, it was possible to simulate NPP, as it does not depend on simulating the evolution of C pools. We have amended the text to make this clearer:

"We note that whilst CABLE has the ability to incorporate carbon pool dynamics (i.e., time varying foliage, wood and root biomass) as a function of nitrogen and phosphorus limitations (Wang et al., 2010), these features were not used here."

This paper explains the impacts of including an improved stomatal conductance scheme in the CABLE land surface model on climate biases in the ACCESS global climate model. The new stomatal conductance model is based on global observations with PFT-specific values for the parameter g1, and it has been documented in a few previous papers. The large biases in the climate model are still present with the new stomatal conductance model, but at least this revision shows some improvement. For example, due to lower ET the daily Tmax is increased, and the bias in the annual maximum Tmax is reduced. This manuscript is well written, and the results are clearly presented. It clearly represents an improvement in both the land surface model and in the coupled modelling system.

We thank the reviewer for this summary; we agree that this captures our paper.

I have a few questions for clarification and some recommendations to improve the paper.

General comments/questions for the authors:
1. My first question relates to the impact of the model on Tmin. Mechanistically, what causes the changes in Tmin? Is it small differences in gs, or is it due to residual effects of the higher daytime temperatures (i.e.: changes in boundary layer height or turbulence due to the changes in the surface energy budget)? I would expect stomatal conductance to be 0 overnight since there is no assimilation. What is the magnitude of the g0 term in Equation 3?

For all simulations using the Medlyn stomatal conductance model, $g_0 = 0$ mol H_2O m^{-2} s^{-1}, meaning that stomatal conductance goes to zero under low light and, importantly, high VPD conditions. By contrast, in the Leuning stomatal conductance model, $g_0 = 0.01$ and 0.04 mol H_2O m^2 s^1 for C3 and C4 species, respectively. We acknowledge that this information was accidently omitted from our original manuscript and we now add this information at the end of section 2.2:

"Additionally, for the MED scheme, $g_0 = 0$ mol H_2O m^{-2} s^{-1}, meaning that g_S goes to zero under low light and importantly, high D conditions. By contrast, in the default LEU scheme, $g_0 = 0.01$ and 0.04 mol H_2O m^{-2} s^{-1} for C3 and C4 species respectively."

With regards to what causes the increase in Tmin, we suspect that this has to be due to the residual effect of higher Tmax during the day and this is made clearer in the discussion:

"We suspect that the increase in T_{MIN} is a residual effect of the increases in T_{MAX} and TXx during the day as differences in g_S are likely to be minimal at night-time"

2. I also have questions regarding the impact of the changes on NPP shown in Figure 8. The authors state that the differences in NPP between the models is due to changes in precipitation, but shouldn't there also be a first-order effect from the different approaches to gs? In CABLE, is the stomatal conductance equation solved iteratively so that the original A is impacted by the gs? If this is true, the relationship between the primary productivity and gs needs to be explained, and I would like to see more explanation for the changes to NPP. If this has been addressed in one of the previous papers implementing this stomatal conductance model, that could be mentioned here instead of a full explanation in this text.

The reviewer is correct, CABLE does iteratively solve A and g_s together, and this is made clearer at the end of section 2.2:

"Finally, we note that for both schemes in CABLE, A and g_S are solved iteratively."

We agree with the reviewer that this text in relation to NPP was unclear and gave the impression that changes in NPP related principally to changes in precipitation via ET. We have amended this text to:

"The reduction in g_s also results in an associated reduction in NPP across the Boreal forests in JJA (Fig. 8b) and in the tropics in both JJA and DJF (Fig. 8). This reduction in NPP is also a result of a change in precipitation across these regions, which results from a reduction in ET due to g_s. There is also an increase in NPP in JJA in the region to the north and east of the Mediterranean, consistent with an increase in ET in this region (Fig. 6f) using MED."

3. The authors state in the Discussion that the changes 'first and foremost' result in changing the ET, and this causes the changes in temperature and precipitation. So why not lead in the results section with the impacts on global ET?

We agree. We have moved this text to the second paragraph and moved the longer text around the advantages of this Medlyn model to the end of the Discussion. We have retained the first paragraph of the discussion in its former position since this acts as a summary of our paper.

4. It would be useful to provide more explanation of some of the overall biases in the model. For example, at line 14 on Page 5243, it is stated that the biases in Tmin over North America in JJA are due to clouds – can this be elaborated upon?

We have added a significant paragraph to the Discussion. It is very hard to go beyond this level of explanation since it would require a case-by-case analysis and is clearly well beyond the scope of this paper. The added paragraph is as follows:

"The ACCESSv1.3 model is a skillful global climate model, performing at the top end in the evaluation of CMIP-5 models (Flato et al., 2013). While the model's overall climatology is very good, there are regional systematic biases that exist irrespective of whether the LEU or MED scheme is used. However, we note that MED does make the model worse in some regions. Perhaps the most serious bias in temperature is over North America in summer (Figs. 2, 3, 4). The biases in ACCESS using LEU reach 5-7°C in T_{MAX}, 8-9°C in T_{MIN} and 6-8°C in TXx. Using MED makes these biases worse, by ∼0.5°C for T_{MAX} and T_{MIN} and by ∼1°C for TXx. Given the magnitude of the error in ACCESS, we suggest that using MED does not really make the model significantly worse since it is already very poor in this region. The poor performance of ACCESS over North America, across several model versions, has been previously reported. Bi et al. (2013) show errors of 3-4°C in the mean air temperature over North America. This does not appear to be linked with low rainfall but rather, it appears linked with an underestimation of cloud coverage over North America despite the overall tendency to simulate slightly too much cloud over northern hemisphere land (Bi et al., 2013). Franklin et al. (2013) examined ACCESS1.3's simulation of clouds but did not focus specifically on North America. They did note problems with capturing convective regimes; these showed too weak a dependence on large-scale dynamics in comparison to observations. The change from the LEU to the MED scheme would not be expected to affect large-scale dynamics or how these processes affect clouds. Small decreases in ACCESS's skill in capturing North America's climate linked with the MED scheme are very likely insignificant; if the large-scale cloud fields were simulated well MED might make a small positive impact in this region but any benefits are currently swamped by the poor cloud climatology. A similar problem exists over the Indian monsoon region where there is a very serious rainfall bias (Fig. 7). The MED scheme does not add value here because there are major systematic weaknesses in the generation of the monsoon linked with larger-scale dynamics as distinct from terrestrial processes."

5. For the DTR results, to my eye it looks like the DTR improvements are mostly due to the increase in Tmax - is this true? If anything, it looks like the increase in Tmin in the boreal forest regions would serve to decrease the DTR.

Yes, this is correct and the text has been modified accordingly:

"Figure 5 shows little impact by switching to MED in DJF, but in JJA there are large

areas of the Northern Hemisphere, coincident with the boreal forests, where the 3–5°C error in DTR is reduced by 10–20%. This is caused by the larger increase in T_{MAX} (Fig. 2(f)) versus T_{MIN} (Fig. 3(f)) ”

6. What is meant by "other warm extremes" in the Discussion, Page 5246 Line 8? Does this mean the TXx index, or something else?

Yes, we meant TXx. We have clarified this in the text.

Tables and Figures:

Figure 1 is not referenced in the text.
It was referenced on page 4, line 103.

Minor typographic comments:

Introduction, page 5238: The last two sentences of the introduction could be made clearer. For example, in the phrase "We seek to determine whether these problems, affecting these and other extreme indices…" the repeated use of "these" in this sentence makes it vague.

Resolved – we modified the sentence to make this clearer:

"These were, in part, attributed to an overestimation of evapotranspiration linked to weaknesses in the representation of land processes. We seek to determine whether the biases identified by Lorenz et al. (2014) can be resolved in part via the parameterization of g_S"

Also you could remove the "We also note that" in the last sentence of the section.

We have now deleted this sentence entirely.

Results, page 5244: In reference to the dry bias in the JJA precipitation (lines 15-20), I think this should be the region to the north east of the Mediterranean.

Yes, we have corrected the text.

Conclusions: There are only two goals discussed in the introduction and conclusions, so is the first sentence "three gaols" is a typo (or else something is missing in the paper!)?

We confirm this is a typo and the text has been corrected to:
"We had two goals for this paper."

Dear authors, In my role as Executive editor of GMD, I would like to bring to your attention our Editorial:

http://www.geoscientific-model-development.net/gmd_journal_white_paper.pdf

http://www.geosci-model-dev.net/6/1233/2013/gmd-6-1233-2013.html

This highlights some requirements of papers published in GMD, which is also available on the GMD website in the 'Manuscript Types' section:

http://www.geoscientific-model-development.net/submission/manuscript_types.html

In particular, for your paper it is desirable to name the model (including the version number) you actually describe and publish in this article in the title, e.g.: "Implementation of an optimal stomatal conductance model (CABLE-2.0.1) in the Australian Community Climate Earth Systems Simulator (ACCESS1.3b).

Sincerely yours, Astrid Kerkweg

We thank the editor for this comment. The modelling or parameterization of stomatal conductance is just one part of land surface models generally. The suggested edit would imply that the new stomatal conductance model and parameterization we have introduced is called CABLE, however, it is just one part of CABLE land surface model, which is part of the ACCESS GCM. We agree that the use of the phrase "optimal stomatal conductance **model**" (in particular the word "**model**") in the title warrants a version number as per the GMD policy. However, this stomatal conductance model/scheme comes from a paper by Medlyn et al. (2011) and does not have a version number. We have hence changed the title to use the word "scheme" rather than "model":

"Implementation of an optimal stomatal conductance scheme in the Australian Community Climate Earth Systems Simulator (ACCESS1.3b)"

We also note that we provide information under "Code availability" on access to our modified version of CABLEv2.0.1 from the subversion repository, which includes the new scheme:

"The modified version of CABLE with the new g_s scheme can be viewed and accessed by registered users at https://trac.nci.org.au/trac/cable/browser/branches/Share/CABLE-2. 0.1-

Tagged-plus-Medlyn-Stom-Param. ”

We hope this satisfies the editor.