In the reply, the referee's comments are in *italics*, our response is in normal text, and quotes from the manuscript are in blue.

Anonymous Referee #1

General comments

This manuscript is a valuable contribution to the geoengineering literature, as it provides a detailed assessment of the simulated tropical circulation response to uniform solar dimming in a suite of coupled climate models. The Hadley circulation does not return to preindustrial conditions in a climate with quadrupled carbon dioxide levels and reduced insolation (the G1 experiment). The authors attribute this result to changes in meridional temperature gradients rather than changes in static stability. The Walker circulation, by contrast, is largely restored to its preindustrial state in G1. The Introduction section effectively describes the many motivations for the study.

The novelty of this work lies in its assessment of the G1 experiment, as much analysis of the impact of elevated atmospheric carbon dioxide concentrations on tropical dynamics has been discussed previously (see references below). At present, the Introduction mentions a few studies on the latter subject (on page 3), but a more thorough review of the existing literature is warranted, both in the Introduction and the Discussion. This manuscript would be more effective if it were shortened so as to emphasize new knowledge; if the circulation changes in abrupt4xCO2 simulations differ from what is reported in the existing literature on the subject, this can be emphasized, but otherwise the geoengineering results should be brought to the forefront.

Reply: Thanks for the suggested additional references. We modify the introduction to include many suggested by both referees:

Climate model simulations with increased greenhouse gas forcing also indicate a poleward expansion of the Hadley circulation, (Hu et al., 2013; Ma and Xie, 2013; Kang and Lu, 2012; Davis et al., 2016). Vallis et al. (2015) analysed the response of 40 CMIP5 climate models finding that there was only modest model agreement on changes. Robust results were slight expansion and weakening of the winter cell Hadley circulation in the Northern hemisphere. It is unclear how closely the model simulations match reality. Choi et al. (2014) and Quan et al. (2014) both suggest that reanalysis trends for the Hadley cell edges may be overstated, especially compared to independent observations., and model trends are in reasonable agreement with the reanalysis trends (Davis and Birner, 2017; Garfinkel et al., 2015), but choice of metric also matters (Solomon et al., 2016) when discussing trends.

Many authors have considered the impact of greenhouse gas forcing on the Hadley circulation, particular in respect of changes in the width of the tropical belt (e.g., (Frierson et al., 2007; Grise and Polvani, 2016; Johanson and Fu, 2009; Lu et al., 2007; Seidel et al., 2008), but far fewer have discussed changes in Hadley intensity (Seo et al., 2014; He and Soden, 2015). The importance of tropical belt widening is of course due to its impact on the hydrological system, especially the locations of the deserts (Lau and Kim, 2015; Seager et al., 2010), which are a critically important for the habitability of several well-populated areas.

We try to shorten the parts related to abrupt4×CO_2, but as the manuscript needs to be self-contained, we do need to discuss the greenhouse gas forced changes as well as the geoengineering results to some extent.

I have questions pertaining to the methodological choices described in Section 2.3. Why is the Hadley cell intensity based on such a broad latitudinal extent (to 40◦ S or N)? This extends beyond the tropics and includes the Ferrel cell. Additionally, the Hadley cell migration is not symmetric in the two seasons (the July-September cell extends further into the summer hemisphere than the January-March cell), so why is the Hadley cell intensity metric hemispherically symmetric?

Reply: We show how the Hadley intensity varies with 2 other choices of cell width in Table R1 below, using 38°-15° or 35°-15° extent in both hemispheres for each model.

There is a consistent difference of 7% for 38°-15° and 14% 35°-15° in all 3 experiments. This shows that while there are differences in the actual intensities computed depending on the latitudinal width of the cells, the results do not show differences between models or experiments. There is just a systematic offset in the intensity calculated, not a change of type of response calculated, or big across-model differences in behaviour, and the offsets are the same for the 3 experiments and both hemispheres. So even though the referee is correct, that we use none-standard definitions so that we capture all the variability in the Hadley cells in all the models and experiments, and also use symmetric cells, it seems that the changes in behavior we observe due to the experiments would not be affected.

Table R1: The differences (10^{10} kg s^{-1}) relative to the method used in the manuscript, defined as 40°-15° S and N) for the Hadley cells. The number in brackets is the percentage difference percent relative to our method

model	Southern hemisphere Hadley intensity in JAS defined by 38°S-15°N			Southern hemisphere Hadley intensity in JAS defined 35°S-15°N		
	piControl	G1	4xCO2	piControl	G1	4xCO2
BNU-ESM	7.4(7%)	7.3(6.8%)	7.7(6.9%)	7.9(14.4%)	7.8(14%)	8(14%)
CanESM2	7(7%)	6.8(7%)	7.2(6.9%)	7.5(14.3%)	7.2(14.4%)	7.7(14%)
CCSM4	7.7(6.5%)	7.6(6.4%)	7.9(6.3%)	8.2(13.4%)	8(13.3%)	8(13%)
GISS-E2-R	6.8(6.4%)	6.4(6.4%)	7(6.8%)	7.2(13.5%)	6.8(13.4%)	7.5(14%)
HadGEM2-ES	7.4(6.7%)	7.4(6.6%)	7.6(6.8%)	8(14.4%)	8(14%)	8(14%)
IPSL-CM5A-LR	6.4(7%)	6.4(7%)	6.6(7.5%)	7(15%)	6.8(14.7%)	7(15.5%)
MIROC-ESM	6.8(7.4%)	6.4(7.5%)	7.2(7.4%)	7.4(15.5%)	7(15.7%)	7.7(15%)
NorESM1-M	7(6%)	6.8(6%)	7(6%)	7.5(12.8%)	7.3(13%)	7.7(12.8%)
Ensemble	7(6.8%)	7(6.7%)	7.3(6.8)	7.6(14%)	7.4(14%)	7.8(14%)

model	Northern hemisphere Hadley intensity in JFM defined 15°S-38°N			Northern hemisphere Hadley intensity in JFM defined 15°S-35°N		
	piControl	G1	4xCO2	piControl	G1	4xCO2
BNU-ESM	6.8(6%)	6.4(6%)	6.7(6.2%)	7.2(13%)	6.8(12.7%)	7(12.6%)
CanESM2	6.3(6.3%)	6(6.4%)	6.2(6.3%)	6.7(13%)	6.5(13%)	6.6(12.8%)

CCSM4	7.0(6%)	7(6%)	7(6%)	7.4(12.8%)	7.3(12.8%)	7.6(12.7%)
GISS-E2-R	4.7(7%)	4.6(7.4%)	4.7(7.4%)	5(15.4%)	5(15.8%)	5(16%)
HadGEM2-ES	6.5(6.7%)	6.3(6.7%)	6.3(6.8%)	7(14%)	6.8(14%)	6.7(14%)
IPSL-CM5A-LR	5.3(7.6%)	5.2(7.4%)	5.2(8.2%)	5.7(16%)	5.6(15.7%)	5.7(17%)
MIROC-ESM	4.6(8.6%)	4.5(8.7%)	4.4(9%)	5(18%)	5(18.4%)	4.8(18.4%)
NorESM1-M	6.4(5.5%)	6(5.6%)	6.5(5.7%)	6.8(11.4%)	6.5(11.7%)	6.8(11.8%)
Ensemble	6(6.6%)	6(6.7%)	6(6.8%)	6.4(13.8%)	6.2(14%)	6.3(14%)

We add "We experimented with using narrower definitions of the Hadley cell (38°-15° or 35°-15°) in the 3 experiments, finding almost the same systematic offsets in intensities across the models and experiments. This is also true for each hemisphere separately. Departures in model ensemble mean intensity across the three experiments for both hemispheres from an outer latitude of 40° range from 6.6-7% and 13.8-14% with outer latitudes of 38°and 35° respectively. So using the wide latitude bands we chose captures all the variability in the Hadley cells in all the models and experiments without introducing biases due to experiments or hemispheres." in line 230.

Finally, throughout the paper it would be helpful to explicitly distinguish between robust results and results for which the ensemble mean is dominated by inter-model cancellation. For example, it is important to elaborate on this last sentence of Section 3.1 (line 253). Does the substantial inter-model spread undermine subsequent interpretation of the ensemble mean change?

Reply: We use standard statistical significance tests to quantify robustness of result throughout, e.g. in Table 2 where we distinguish between ensemble mean changes that are significant at 95% confident level in abrupt4×CO_2, but not G1. The other method commonly used to assess model agreement is if some super-majority, e.g.75% of models agree on the sign of an anomaly. Using that measure implies that in Table 2 the change under abrupt4×CO_2 is not robust. We add the model agreement wherever relevant to the results we present, for example we change:

"However there is some scatter between models (Table 2)." To:

There is significant change in the ensemble mean position and strength under abrupt4×CO_2, but not G1 in Table 2. However, only 5 out of 8 models agree on the sign of the changes, so the inter-model differences are rather large in this case.

Specific comments

The ENSO-related results could be included in the abstract.

Throughout the paper, starting in the abstract, the Hadley circulation changes are discussed in terms of "seasonal maximum Northern and Southern cells." It would be clearer to discuss these together as the "solstitial Hadley cells," or as "the JAS and JFM cross-equatorial cells."

Reply: Unfortunately the abstract length guide in TC is 100-200 words. Presently we are at 250. We think that the changes in ENSO are of less importance than the other aspects we discuss in the abstract.

With respect we disagree that using the term *solstitial* is an improvement or disambiguation. Typically *solstitial* refers to the summer solstice (e.g. Websters disctionary), whereas in this case it would refer to the winter solstice. The vast majority of the cell we refer to "Northern" is indeed in the Northern hemisphere, and we think using the terms "JAS Southern" and "JFM Northern" cells is the least ambiguous choice.

(lines 79-80) It is not entirely clear what is meant by this phrase: "and a similar response of oceans versus land."

Reply: We change "There is also a relative undercooling of the polar regions and overcooling of the tropics and a similar response of oceans versus land with globally uniform SRM." To

The general pattern of temperature change under all abrupt4×CO_2 includes accentuated Arctic warming, and least warming in the tropics. G1 largely reverses these changes, but leaves some residual warming in the polar regions and under-cools the tropics relative to piControl. Geoengineering also reduces temperatures over land more than over oceans relative to abrupt4×CO_2, and hence reduces the temperature difference between land and oceans by about 1°C.

(line 84) Held and Soden (2006) are better known for explaining this P-E scaling, which is derived from the Clausius-Clapeyron relation and only valid over ocean (reference below).

Reply: Thanks, we add the reference to Held and Soden, we did not mean to claim priority for Tilmes, it was a convenient reference that we used elsewhere

(lines 108-110) This sentence is unclear: "The signal to noise ratio [. . .]."

Reply: We change "The signal to noise ratio in the G4 experiment is relatively low with a background of only the modest RCP4.5 greenhouse gas forcing scenario." To

The geoengineering and greenhouse gas forcing in the G4 experiment are both relatively low compared with the G1 experiment, since under G4 the greenhouse gas scenario is the modest RCP4.5, which means that natural climate variability in the 50 year long period of geoengineering period may obscure features.

(line 198-200) Are all 50 years used only for those measures that do not rely on sea surface temperature?

Reply: No. we use 50 years for zonal and meridional stream function and sea surface temperature correlation with STRF of the Walker circulation in Fig. 11. We note that There is no difference in model behavior between the G1 and abrupt4×CO_2 anomalies and △SST explains 83% of the overall variance. Despite a temperature transient of at a decade or so (e.g. Kravitz et al., 2013) in the abrupt4×CO2 simulation and the lack of any transient in STRF (Fig. S1), the relationship with ΔSST is nearly as good as for piControl. This is because the STRF depends on the SST at that time, which we correlate.

We use 30 years for air temperature and land temperature in Figs. 12 & 13.

(line 213) Does the phrase "whole streamfunction" mean for all longitudes? Say that explicitly.
Reply: Yes, and corrected

(line 237) Does mean state refer to annual mean state?
Reply: Yes, and corrected

Labeling multi-panel figures would facilitate the discussion of results in the text.
Reply: OK. We add letters to distinguish the separate panels.

(line 308) What constitutes a "good relationship?"
Reply: We clarify this as : Fig. 7 (B) shows that the modelled motion of the ITCZ explains 73% of the variance in intensity of the JAS Southern cell peak intensity, which is significant at the 95% level. Thus the larger the model reduction in intensity the more the boundary of the ITCZ moves equatorward.

(line 312-314) Note that this was reported by Smyth et al. (2017).
Reply: Noted: The combined seasonal effect of both cell changes is a reduced migration of the upwelling branches of the circulation cells across the equator, as was also noted by Smyth et al. (2017).

(lines 321-344) This section can be made more concise by focusing on similarities or differences from previous studies on the subject (cf. major comment above).
Reply: we modify this section as suggested:

The situation under abrupt4×CO_2 is more complex (Fig. 6 (E) and (F)). The expansion of the tropics has been noted both in greenhouse gas simulations and observationally (Davis et al., 2016; Hu et al., 2011), along with the larger southern expansion. The extratropical changes in the Ferrel circulation are also more pronounced in the Southern hemisphere.

Reduction in strength of the Northern hemisphere winter cell was also a robust result of climate models under RCP8.5, while, the Southern cell exhibited almost no change (Vallis et al., 2015). Our results in Fig. 8 show that the multi-model ensemble mean reduction Hadley intensity under G1 of -18×10^{8} kg s^{-1} and of -7×10^{8} kg s^{-1} for abrupt4×CO_2. The JAS Southern Hadley intensity exhibits a fall of -16×10^{8} kg s^{-1} under G1 but an increase of 23×10^{8} kg s^{-1} under abrupt4×CO_2. At least 6 out of 8 models agree on these sign of changes in both hemispheres and scenarios. Thus the Southern hemisphere results differ for abrupt4×CO_2 from those presented in Vallis et al. (2015). The anomalies for most models are significant, and the ensemble means are 8 standard errors from zero and thus very highly significant.

We move our definition of the anomalies to the fig 8 caption:

Figure 8. Anomalies (10^{10} kg s^{-1}) relative to piControl amongst models in Hadley circulation for the Southern cell in JAS (left panel), defined as the magnitude of the

mean meridional stream-function between 15°N and 40°S, and (right panel) the Northern cell in JFM, defined as the magnitude of the mean meridional stream-function between 15°S and 40°N. The dot size for the models is about 1 standard error of the model mean.

(lines 339-340) Cite other studies which have noted that solar dimming results in an overcompensation of tropical circulation changes induced by global warming.
Reply: we don't see how this question arises from the text at line 339-340 nor anywhere else.

You might consider discussing all of the analysis of the Walker and Hadley circulation responses together, i.e. moving current Section 6.1 to follow current Section 3.2.
Reply: We want to discuss the mechanism of Walker and Hadley circulation by analyzing the relationship between them and temperature. Section 6 is where we do this as that is focused on mechanisms rather than presenting results as in section 3 and 4.

(lines 399-403) This section is confusing. "Monthly temperatures" in which region?
Reply: This sentence is not needed, as we discuss only annual means in the rest of the paper and so we delete it**.**

(lines 433-436) Why do you choose to analyze temperatures over Tibet if the general land-ocean temperature contrast is of interest? This choice should be justified, or the analysis modified. Are you considering surface temperatures? The methodology is not described in enough detail for the results to be reproduced. Additionally, there is an extensive body of literature connecting inter-hemispheric temperatures and the Hadley circulation/Intertropical Convergence Zone, such as Broccoli et al. (2006), which can be referred to here.
Reply: There is difference in surface temperature response of oceans and land under geoengineering. Under pure greenhouse gas forcing very small land-ocean temperature differences are forecast in the tropics, so it is interesting to explore any change in the circulation behavior that may result. There is also well-documented influence of Tibet-Indian Ocean temperature difference affecting Asian monsoon intensity. Hence it seemed worthwhile to explore this in the fig S9. We do think it is worthwhile to show this, though we don't discuss the result in detail in the paper. We modify the text:
Both SRM and greenhouse gas forcing modifies the land-ocean temperature difference relative to piControl and so conceivably affects Hadley circulation, for example by changing the hemispheric temperature and the position of the ITCZ (Broccoli et al., 2006). Under abrupt4×CO_2 land-ocean temperature differences in the tropics (between 30° N and 30°S) are reduced to essentially zero, while under G1 differences in the tropics are 1.2°C which is not significantly different from the piControl difference of 1.4°C. Since the largest continental land masses are in the Northern hemisphere, we would expect any differences in Hadley circulation induced by land-ocean contrasts in the Northern hemisphere to be visible in the Southern Hadley cell. We explored the impact of land-ocean temperature differences by considering differences in the surface

temperatures over Tibet and the whole tropical ocean temperature (Fig. S9). Results were similar as for Fig. 13, with significant correlations for G1 in the Southern Hadley cell.

(lines 445-448) Specify the location and season on which this is based. Are these near-surface potential temperature gradients?

Reply: As we use the Seo et at. (2014) method, the potential temperature gradients are defined here as the average between 1000 and 400 hPa. And we use JFM for Northern hemisphere and JAS for Southern hemisphere. We specify this in the Table 3 notes:

& Function 1 is $\frac{5}{2}\frac{\delta H}{H}+\frac{5}{2}\frac{\delta\Delta_H}{\Delta_H}-\frac{\delta\Delta_V}{\Delta_V}$ and is based the model of Held and Hou (1980).

Function 2 is $\frac{9}{4}\frac{\delta H}{H}+2\frac{\delta\Delta_H}{\Delta_H}-\frac{3}{4}\frac{\delta\Delta_V}{\Delta_V}$ which is derived from the model by Held (2000).

Δ_H is meridional temperature gradient defined as $\frac{\theta_{eq}-\theta_{higher\ lat}}{\theta_0}$ which is the tropospheric mean meridional potential temperature gradient with θ_0 denoting the hemispheric troposphere mean potential temperature and θ_{eq} calculated between 10°N and 10°S. We follow Seo et al. (2014) in taking $\theta_{higher\ lat}$ as the average potential temperature between 10°-50°N for the Northern hemisphere winter and 10°-30°S for the Southern hemisphere. Potential temperature gradients are defined here as the average between 1000 and 400 hPa. $\Delta_V=\frac{\theta_{300}-\theta_{925}}{\theta_0}$ is the dry static stability of the tropical troposphere. H is the subtropical tropopause height estimated as the level where the lapse rate decreases to 2°C km^{-1}.

! The Hadley intensity ψ_m is described in section 2.3 and we use JFM in the Northern hemisphere and JAS in the Southern hemisphere.

(lines 468-470) The sentence "However only the eastern and western [. . .]" is confusing.

Reply: Rewrite it as "However the eastern and western boundaries of the Walker circulation shift westward during El Niño in G1 relative to piControl."

The analysis of the Walker circulation should be better framed. For example, He and Soden (2015) explain that weakening of the Walker circulation due to carbon dioxide forcing is mostly driven by the change in mean sea surface temperatures (SST).

Reply: Thank you, we now include this as a frame to the discussion: He and Soden (2015) conclude from experiments designed to elucidate the role of various forcings on tropical circulation that weakening of the Walker circulation under greenhouse gas forcing is primarily due to mean SST warming. They also note that increased land-sea temperature contrast results in strengthening of the circulation, and also that while the pattern of greenhouse gas warming is close to an El Niño, there are sufficient differences to produce quite different responses in the Walker circulation. We may therefore expect that changes under G1 compared with pure greenhouse gas forcing would manifest themselves given the changes in both the direct and indirect CO_2 forcings. What we observe though is that changes in the Walker circulation are modest, and examination of the dependence on intensity as a function of zonal Pacific Ocean temperature differences (Fig. 11) show no differences between the greenhouse gas and G1 forcings. Similarly we find no change in the intensity with land-ocean temperature gradients.

(line 500) The second half of this sentence is unclear.
Reply: we rewrite this as 2 sentences:
Furthermore the intensity of the Hadley circulation is expected to decrease as it expands and also in response to an accelerated hydrological cycle. An enhanced hydrological cycle is expected under greenhouse gas forcing, but not solar geoengineering which leads to net drying (Kravitz et al., 2013).

(lines 514-516) Rephrase this sentence for clarity: "But we observe [. . .] G1 forcing."
Reply: We change it to: Nor it is consistent with increases simulated in the Southern Hadley cell intensity and simultaneous decreases in the Northern one relative to piControl, although both are stronger than under the G1 forcing.

(line 545) Does "ocean heating" refer to warming sea surface temperatures or ocean heat uptake?
Reply: Yes, we rephrase this for clarity: intuitively mean reduced heating, sea surface temperatures and moisture flux

(line 548) Define "Rx5day extreme"
Reply: We were inaccurate in the phrase and change it as : shows that the annual wettest consecutive five days are drier

Figure 10: Is this based on annual data from El Nino years, or data from a particular season? In a few places in the paper it is not immediately clear what averaging periods and spatial domains are used for calculations.
Reply: The method is defined in section 2 lines 159-165. "Composite analysis is applied for the study on the influence of ENSO. We follow Bayr et al. (2014) and use detrended and normalized Nino3.4 index (monthly averaged sea surface temperature anomaly in the region bounded by 5° N - 5° S, from 170° W - 120° W) as a criteria to select ENSO event. An index > 1 represents an El Niño event and < -1 a La Niña one (Bayr et al., 2014). We concatenate variables in all El Niño and La Niña events for each individual model to get El Niño and La Niña data sets and then calculate ensemble results."

Technical comments
(Line 10) capitalize "Earth"

Reply: done.

(Line 25) "good correlations" should be rephrased more precisely
Reply: Change it as "There are significant relationships between Northern cell intensity and land temperatures"

(Line 55) typically capitalized "Northern Hemisphere"
Reply: done.

(line 29) "response to" should say "responses of"
Reply: done.

(Line 63) "compliment" should be "complement"
Reply: done.

(line 157) not a sentence
Reply: Changed to We used monthly-mean model output data.

(line 349-351) This is not a sentence.
Reply: Changed to Hadley circulation shrinks and strengthens during El Niño events, while expanding and weakening during La Niña.

(line 542) "While under [. . .]" is not a sentence.
Reply: Changed to: There are clear changes of Hadley cells under the latitudinal varying forcing of G1.

Additional modifications: We also revise the following sections.
(Line 59) Rewrite the sentence.
Model results suggest a significant eastward movement with weakening intensity under greenhouse gas forcing (Bayr et al., 2014), and He and Soden (2015) propose that the sea surface temperature warming plays a crucial role in both the eastward shift and the weakening of Walker circulation. They also note that this weakening may be reversed by rapid land warming.

(Line 207) Correct the writing mistake (5°S - 5°N, 160°W - 80°E) to (5°S - 5°N, 160°W - 80°W)

(Line 450) We change "respectively in south and north cell" to "respectively in Southern and Northern cells"

(Line 475) Delete "Davis et al., (2016) note an expansion in the Hadley cells in proportion to the temperature rises in the models under both G1 and abrupt4×CO_2."

(Line 543) Rewrite it as "The reduction in incoming shortwave radiation in G1 would intuitively mean reduced heating, sea surface temperatures and moisture flux in the ITCZ"

(Line 545) Delete "Reduced ocean heating would then tend to mean a smaller amplitude of seasonal movement of the ITCZ"

(Line 269-273) We redraw Fig.1 to show the difference of ERA-piControl rather than piControl-ERA and rewrite the relevant text "Fig. 1 (D) shows that the ERA-Interim circulation has an eastward displacement and the intensity measured by STRF is overestimated by 26% relative to ensemble piControl. There is a similar structure to the stream function differences between NCEP2 reanalysis and piControl, and the STRF is only overestimated by 3% relative to ensemble piControl."

(Line 316-317) We redraw the Fig.4 show the difference of ERA-piControl rather than piControl-ERA and rewrite the relevant text "The intensity anomalies relative to piControl from both the reanalysis data sets are less than 21% (Fig. 4)."