We would like to thank the anonymous referee #1 for his/her very helpful review of our publication and provide here the answers to his/her questions.

1) P6192 L18 – P6193 L13
I found the discussion around the ability of SOCOL3-MPIOM to simulate the top-down mechanism, which is thought to involve a modulation of the strength of the stratospheric polar vortex, to be unsatisfactory and incomplete. On P6193 they describe how the post-volcanic European warming in their VOLC experiment is consistent with other studies that have highlighted a top-down mechanism following volcanic eruptions (e.g. Driscoll et al., 2012, JGR). However, in the conclusions (P6201 L24-25) they argue that this feature is only weakly significant and that the top-down mechanism doesn’t appear to be operating properly in their model, which seems to be somewhat contradictory.

We agree that the phrasing was unlucky. We therefore changed P6193, L6-10 to:

In agreement with Robock and Mao (1992), Kirchner et al. (1999), and Driscoll et al. (2012), or to the DM analysis of Fischer et al. (2007), we discern a slight, yet significant winter warming pattern (WWP) over Europe, Russia and parts of Northern America in the years following the volcanic eruptions (Fig. 3b) and a weak cool anomaly during the summer seasons following the volcanic eruptions (Fig. 3a). The signal – most probably due to a too weak representation of the top-down mechanism during volcanic eruptions – is weaker than in the aforementioned studies. The warming in ....

If the model does simulate a top-down mechanism in response to volcanic aerosol forcing, then why would it not also capture a similar top-down mechanism due to solar forcing? The authors make some inconclusive statements (P6193 L1-2) about the role of the climatologically weak polar vortex in the model, but this issue is not addressed in detail. The reader is left wondering whether the results and their validity may be strongly affected by model deficiencies in being able to simulate the relevant processes to capture the surface response to volcanic and solar forcing. I deem this to be a deficiency of the manuscript it its present state, and recommend that a more detailed examination of this issue be included.

It is true that the current description of the problem is very short and maybe inconclusive. Yet, this publication is not meant to be a technical assessment of the model. We recommend the publications of Muthers et al. 2013 (accepted) and 2014 (in prep.), which give a deeper insight of some model deficiencies.

Moreover, the top-down mechanism during volcanic eruptions or grand solar variations stems from different heights in the atmosphere; those disturbances have also different magnitudes. While during volcanic eruptions, strong positive temperature anomalies emerge in the lower stratosphere, solar disturbances imply weaker negative temperature anomalies in the upper stratosphere. SOCOL-MPIOM seems to be less sensitive to changes in temperatures in the upper stratosphere than in the lower stratosphere. We agree that this point should be reconsidered in a future publication to exclude any bias of the results of our investigations.

We changed the sentence on P6201, L27:
It can be related to model deficiencies in the simulation of the vertical coupling, of the polar vortex state or of the wave generation and propagation, which become especially apparent when looking at the weak TD / EPP response.

P6201 L11-13: The apparent contradiction of the results of Feulner (2011) requires a more complete discussion. Do the contrasting conclusions arise from differences between the model simulations presented in the two studies, or is it related to the fact that the two studies use different temperature reconstructions to evaluate their model results? In particular, how different is the temperature reconstruction of Frank et al. (2010) used by Feulner (2011) from the five reconstructions used in your study? Comparing Feulner (2011) Figs 2 and 3 to your Fig 6, it would appear that the modelled temperature anomalies might not be that different, but that the reconstructed temperature data are quite different.

We agree that the discussion might be too short. However, we can assure that the differences between the two studies do not arise from the selected reconstructions, as can be seen in Fig R1 (Muthers et al. 2014, in prep.), showing the temperature evolution of all our 4 transient simulation members (1600-2000). Even if all the reconstruction of Frank et al. 2010 would have been plotted in Fig. 6, no significantly different picture would be given from the comparison reconstructions vs. simulations.

Thus, the differences in the design of the experiments (model complexity) must be the key factor for those significant differences in surface temperature response.

We added following in P6201, L13: The most obvious explanation of this discrepancy between the two studies is the complexity of the model. While Feulner (2011) used an earth system model of intermediate complexity, while this study uses a much more interactive atmosphere ocean chemistry model of high complexity.

Fig. R1 NH average 2m temperatures for the four transient simulations (lines) in comparison to the probability range of different NH temperature reconstructions (Frank et al., 2010). The ensemble members L1 and L2 correspond to the large amplitude solar forcing, M1 and M1 were forced by the moderate TSI amplitude. Reconstructions are given as anomalies to the pre-industrial period 1600-1850 for the reconstructions and the model simulations. This allows for a direct comparison of the variability in the pre-industrial period despite the strong 20th century temperature trend. Model time series were decadal smoothing with a cubic-smoothing spline similar to the reconstructions.
Are the timings of ENSO events different in the various sensitivity experiments? If so, I would expect this to be important for the shorter timescale fluctuations in the temperature timeseries in Figs 4 and 6. Furthermore, the timings of ENSO in the model are unlikely to be the same as those which occurred in the real world, so can you factor in the possible effects of this variability into your comparison of the modelled vs. reconstructed temperatures in Fig 6?

We investigated this question more in detail. First, it is clear that due to the coarse resolution of the models, the ENSO events are timed differently in the different sensitivity experiments. By defining an ENSO event as such to have a normalized Niño 3.4 region index above 1, we find following events:

| BU     | 1781, 1782, 1785, 1786, 1801, 1803, 1807, 1816, 1823, 1829, 1835, 1837 |
| VOLC   | 1791, 1804, 1817, 1825, 1832, 1833                                    |
| ALL    | 1781, 1789, 1791, 1796, 1808, 1830, 1836                                |

It is possible that the different timing induces different small-scale fluctuations in the temperature time series. We thus added following sentence on P6197L10:

*ENSO events monitored by the Niño3.4 index are not shown. Yet, it still should be stated that most (> 70 %) of the events differ from one sensitivity experiment to another with a time lapse of +/- 1 year. However, at the periods of interest, which are discussed later in this section, the ENSO events of BU and VOLC happened at +/- the same period. The general…*

We are however unable to plot an uncertainty envelope around the temperature curves of Figure 6 as no reliable proxies for ENSO exist for that period of time.

**P6182 L7 ejecting – replace with ‘injecting’**

This has been corrected.

**P6183 L1 I would have thought the main issue here is not the size of the perturbation (0.3 W/m2), which could in principle be large enough to impact on surface temperatures, but the fact that most of this energy will be absorbed in the stratosphere.**

This is true and it seems that we formulated this sentence not clearly enough. We thus reformulated:

*…: on one hand, a substantial decrease in the UV-C at lambda > 250 nm (0.3 W/m^2) cools down the middle atmosphere and decreases the ozone production due to decelerated oxygen photolysis (Anet et al. 2013), resulting in a very small radiation anomaly on the earth surface. On the other….*

**P6183 L4-14 I think it is important to distinguish between the two types of top-down mechanism that have been proposed – one relates to a modification of the strength of the stratospheric polar vortex (winter-time only), and one relates to the direct impact of the tropical lower stratospheric ‘secondary maximum’ in temperature on the midlatitude jets (e.g. Haigh et al., 2006; Simpson et al., 2009), which could in principle operate during other seasons.**

We understand the two top-down mechanisms not as separate mechanisms, but as linked mechanism. The first mechanism which the reviewer states ("winter-time only") triggers the second mechanism ("secondary maximum"), as explained in Kodera & Kuroda (2002)

Nevertheless, we reformulated P6183, L4-14:
“A negative UV-C anomaly affects the state of the stratosphere and mesosphere (Rozanov et al., 2012a; Anet et al., 2013), from where it may influence the troposphere via a cascade of mechanisms: by cooling down the tropical and midlatitude stratosphere, it decreases the pole-to-equator temperature gradient, weakens the zonal winds and accelerates the Brewer–Dobson circulation. The latter is followed by a cooling in the lower tropical stratosphere (Kodera and Kuroda, 2002), and a subsequent modulation of the Hadley cell (Haigh, 1996) impacting especially the equatorial region and alteration of the tropospheric wave pattern (Brugnara et al., 2013), propagating down to the surface. This is also known as the “top-down” mechanism (Meehl et al., 2009). However, in the present set of simulations the top-down mechanism is shown to be of minor importance when comparing with other mechanisms discussed below.”

P6183 L20 This sounds very deterministic – would suggest changing to 'this is thought to lead to'

We introduced this change

P6183 L22 Both top-down and bottom-up mechanisms are described as having the potential to modify the Hadley cell – how might their relative effects be separated?

The effects are separable especially concerning the time of the year, at which they start to influence the circulation. Generally: TD starts in hemispheric winter time at the polar regions, BU at any possible time especially at the equatorial region.

We added following sentence in P6183L23:
Both mechanisms thus finally influence the atmospheric circulation, differentiable by the time at which and where they start to influence the atmosphere. Generally, one can say that the top-down effect essentially starts to influence polar regions in hemispheric winter time, whereas the bottom-up effect literally can influence especially tropical regions during the entire year.

P6184 L16 Is there a reference for the 60Mt sulphur? Please add.

This is the estimated load of the Tambora eruption by Gao et al. (2008).

P6184 L25 Could this warming of the polar vortex also be due to dynamical modulation rather than purely radiative effects?

We want to specify that it is not a warming of the polar vortex, but rather a warming of the entire polar night region. Hence, this cannot be a purely dynamic feature.

P6185 Which solar forcing did Shindell use? This 0.6-0.8 K seems large compared to other estimates for the effects of a grand solar minimum (e.g. Jones et al., 2012, JGR).

The used solar forcing of Shindell et al. (2000) is unfortunately not known. We agree the temperature response is one of the largest.

We added following to the manuscript, P6185, 15.
Shindell et al. (2000) compared the long-term influence of volcanic eruptions to grand solar minimum conditions with focus on the DM and on the Maunder Minimum (MM) – which occurred about 150 yr before the DM. Unfortunately, the exact solar forcing used for their modelling study remains unknown, but they concluded that…

P6188 L11 Define ‘phi’.

We changed to “The GCR ionization rates depend on the solar modulation potential Phi, …
Yes, but is this likely to be an upper estimate of what the ‘real world’ influence is?

We are unfortunately unable to understand what the reviewer means here.

CMIP4 should be CMIP3, I think.

This is correct, thank you. We changed that accordingly.

How long is the control run? Please explain in more detail how the three ensemble members are initialized from the control run and how the differences between the three transient ensembles and the control run are constructed?

As reviewer #2 raised the same question, we slightly modified the paragraph. See answers for reviewer #2 for more details.

Why is the Student’s t-test computed for the ensemble mean differences for the 20 year period rather than the differences across all 3*20=60 data points as was done in Anet et al., 2013, ACP? How is autocorrelation taken into account? It appears that the df = 2N-2, as is the case for independent data samples.

It seems like the text is misunderstanding. We reformulated:
“The statistical significance of the global distribution of the 2m temperature anomalies were computed using a 2-sample Student’s t test across all 3*20 = 60 data points as was done in Anet et al. (2013) on a 5% significance level, taking autocorrelation into account. The latter was done by calculating the number of independent data points over the 3*20 time steps. The statistical…”

Is this annual-mean temperature? Please state this.

We changed to “The regional pattern of the annual mean 2m temperature difference between the ALL and the CTRL1780 simulation is illustrated in Fig. 2a.”

Presumably you have the diagnostics to check this?

We re-investigated this point thoroughly and agreed that the diagnostics were too small in significance to argument with the albedo. Many thanks for that critical consideration of this aspect.
We deleted the sentence completely.

Remove ‘being’
L assure – replace with ‘ensure’
L neither – replace with ‘either’

Done

It does not look like Fig 2(b) + Fig 2(c) = Fig 2(a) in many regions, and therefore the solar and volcanic responses are not additive. If this is the case, why can some features, such as the warming over the Bering Sea (P6191 L11-12), be directly attributed to the BU or VOLC forcing by comparing Fig 2(c) to 2(a), whereas the combined (BU + VOLC) changes in other regions, such as the cooling over Northern Europe and Australasia, do not directly correspond to ALL?

We thought it would be clear for the reader that there are significant nonlinearities to be considered between BU+VOLC and ALL.
We added following to P6192L4:
One might speculate that the warming pattern shown in ALL results from a combination of volcanic and solar influences, although certain nonlinearities prohibit the direct comparison of BU+VOLC and ALL, as was already shown in Anet et al. (2013).

P6192 L7 add ‘(not shown)’
P6192 L20 exchange – replace with ‘coupling’
P6193 L1 reword to ‘, but it may originate in the weaker winter vortex’
P6193 L17 add units to standard deviations.
P6196 L19 amply – replace with ‘considerably’
P6197 L10 that means – replace with ‘which shows’

Many thanks for all these suggestions. We implemented all of them.

P6197 L19 Do the differences in the timings of an 11 year like signal imply that the dating is not sufficiently accurate on decadal timescales to detect the effects of solar forcing. If so, what are the implications for investigating the detailed evolution of temperature over the Dalton Minimum?

We would rather say that the reconstruction of the 11-year-cycle from tree-rings might not be directly comparable to a climate model simulation. What should be shown here is the “big picture” of the grand solar minimum versus volcanic eruption, which is sufficiently captured by both reconstructions and simulations.

P6201 L16 ‘famines’ – this statement is highly speculative, I suggest removing it.

We agree this sound a bit alarmistic and removed “famines”.

P6202 L11 for a certain time – this statement is vague, please clarify.

We investigated this question and changed to:
“We also show that due to volcanic eruptions, the hydrological cycle can be perturbed as such to decelerate the Hadley and Ferrel cells for timescales of 1-3 years.”

Citations
Muthers et al., 2013, accepted, doi:10.1002/2013JD020138