We would like to thank the anonymous referee #2 for his/her very helpful review of our publication and provide here the answers to his/her questions.

1) My major concern regards the very limited size of the ensembles. I understand that this is due to unavoidable computational limitations, and that this does not necessarily hamper the general validity of the conclusions (though some results are based on rather low confidence). However, as also noted by the Authors in their conclusive statement, representativeness is a potential issue. It should be made clear that the initialization allows at least spanning substantially different ocean conditions. In a recent modeling study focusing on the same period, Zanchettin et al. (2013) showed that internal climate variability can strongly spread the simulated decadal climate response to a strong eruption, with individual realizations differing for as much as 1 K in decadal NH temperature outputs during the first two decades after the 1815 Tambora eruption. Their results also showed the relevance of internal variability for simulations-reconstructions comparisons, since individual model realizations resulted to be either incompatible or closely tracing reconstructed trajectories. I therefore strongly encourage the Authors to describe more in detail the ensemble generation method (rather than simply referencing to Anet et al., ACP), and describe succinctly how much the initial states differ concerning key climatic features (for instance ENSO and the Atlantic meridional overturning circulation).

We agree with the reviewer that the size of ensembles is rather limited and may potentially hamper the statistical significance of certain comparison. However, by choosing carefully the initial disturbance (focused on the ocean), we tried to minimize the problem of the internal variability by a maximum.

It is clear that with an Atmosphere-Ocean-Chemistry-Climate-Model, one should not simply introduce disturbances in the atmosphere by e.g. changing the CO$_2$-atmospheric concentration for the first month. This would not work as the ocean has a relatively high inertia, possibly “ignoring” the atmospheric disturbance.

Hence, the ocean was disturbed in the following way: We first analyzed the oceanic signal over the years 1775-1805 to determine the “best” year, avoiding to choose e.g. strong El Niño or La Niña years. It became apparent that the oceanic 1778, 1779, and 1780 provide a good mix between weak La Niña and El Niño years (El Niño 3.4 Indexes of -.9, +0.8, +0.6 vs. strong Indexes of +1.6 or -1.6 in adjacent years). Moreover, as can be seen in Fig. 4, the spread between the different runs is not as high that contradictory, inconsistent findings would be given.

We thus decided to add following in the manuscript on P6189L14-15:

“The simulations, identical to those described by Anet et al. (2013), were initialized from a long, transient model run covering 1600-2100 AD. The disturbances were introduced by starting the sensitivity study simulations from an ocean state being one year “older” and one year “younger” than December 1779, as the time frames of December 1778, December 1779 and December 1780 provided a good mix between weak El Niño or La Niña conditions, avoiding extreme conditions in the oceanic signal (El Niño 3.4 Indexes: -.9 for December 1778, +0.8 for December 1779 and +0.6 for December 1780 of the “mother run”).”

Moreover, the conclusions were extended by following on P6203L2:
“Also, internal variability might influence the simulated response of the Bering Sea region: Zanchettin et al. (2013) showed that internal climate variability can strongly spread the simulated decadal climate response to a strong eruption, with individual realizations differing for up to 1 K in decadal NH temperature outputs during the first two decades after the 1815 Tambora eruption. A higher number of ensembles would consolidate our findings.”

To this regard, it is also important to be sure that the control run provides a reliable characterization of internal variability. The manuscript somehow lacks information about the control run, and I would appreciate more details being reported. For instance, from the caption of Figure 3 I understand that only 60 years were considered for that analysis: if this were the full length of the control run, it would possibly be insufficient due to undersampling of internal multi-decadal and centennial variability.

The total run length of the control runs used for our analysis is 3*60 = 180 years which is long enough to avoid any undersampling. We changed the caption of Figure 3 accordingly.

I would also appreciate if some words were spent about general aspects of the climate variability simulated by SOCOL3-MPIOM, such as the spectral properties of the simulated ENSO (ENSO is known to be too strong and too regular in the similar ECHAM5/MPIOM model in T31L19/GR30 compared to observations). Such information helps putting confidence on the results and otherwise clarifies caveats to be accounted for in the conclusions.

Understanding the problem of reviewer #2, we have discussed that point thoroughly: It is not the aim of this publication to describe the model behavior of SOCOL3-MPIOM. We prefer referring to the work in of Muthers et al. (2014, in prep.), describing the entire model framework and its climatological properties. We are sure this will help putting confidence to the reader.

The fact that the strongly-varying solar forcing produces temperature variations compatible with available reconstructions seems to be one of the most important conclusions that can be drawn from this study, since it makes the Shapiro’s SSI reconstruction plausible. This is even more so since the VOLC ensemble considerably deviates from the reconstructed trajectory (often in opposite direction that the BU ensemble does). Given the importance of the simulations-reconstructions result, I encourage the Authors to expand a bit the discussion especially on its interpretation in the light of indications from previous studies (in particular Feulner, 2011, but also the abovementioned Zanchettin et al., 2013). Certainly the outcome of such a cross-validation depends on the combination of selected forcings and selected reconstructions: for instance, here the compatibility between the all-forcing ensemble and the reconstructions is largely originated by “compensation” between solar (BU) and volcanic forcings. So, how much does the use of Gao’s aerosols matter for the overall comparison and related inferences about the validity of the Shapiro’s SSI reconstruction?

We agree that an in-depth comparison of previous studies, including a range of sensitivity studies including different solar and volcanic forcings, would be of utmost importance and interest for the scientific community. However, we fear that this would go well beyond the scope of this publication.

We would also like to refer to the answer to reviewer #1, which raised the question of the comparability of the Feulner (2011) study and ours.

Concerning the interpretation of the simulated oceanic evolutions, the Authors only briefly discuss changes occurring in the North Atlantic, although this is a key region for the climate response to volcanic and solar forcing in several models (e.g., Stenchikov et al., 2009; Swingedow et al., 2010; Otterå et al., 2012; Zanchettin et al.,
2012, 2013). I would appreciate some extended discussion about the dynamical responses in the North Atlantic Ocean or lack thereof, especially concerning the thermohaline circulation. As shown by Zanchettin et al. (2013), the interplay between oceanic heat transport in the North Atlantic and Arctic sea ice strongly determines the dynamical climate response in the continental Northern Hemisphere. So, for instance, a lack of North Atlantic response to volcanic forcing in the SOCOL3/MPIOM simulations may prevent sea-ice related feedbacks to set in in the Arctic. This could also be important for the interpretation of the all-forcing ensemble as simply resulting from “compensation” between BU and VOLC effects.

Thank you very much for this valuable input. We agree that the North Atlantic region could seem to be neglected in our analyzes. We decided to add a paragraph discussing briefly the AMOC (P6192L5):

"Moreover, as illustrated in Fig. S2, the response in the AMOC is relatively weak in BU, while a distinct increase in the AMOC is visible in VOLC. This finding agrees well to the work of Zanchettin et al. (2013), which find a significant increase of 0.6 Sv of the AMOC after Tambora, while we find a significant increase of up to 1 Sv after both volcanic eruptions (beginning of 1820ies minus pre-volcanic era). We do not find any additive effect of both eruptions, which could mean that a certain saturation effect might stop the acceleration of the AMOC. This finding however should be investigated more in detail in a future work."

As a final general comment on the writing style, I found that Section 4 (Conclusions) contained a lot of material which is actually discussion of the results. I recommend the Authors to split this lengthy section in two parts (Discussion and Conclusions), and to put only the major concluding remarks in the latter.

We considered restructuring the paper, however decided to keep the current setting.
P6182L14: isn’t the last IPCC report the AR5? Maybe it’s better to specify you refer to AR4

By the time we were writing the publication, AR5 was not yet published. We are sorry for this error and will of course change P6182L14 to AR4.

6183 L13: comparing –> compared

Changed

6184 L6-7: please check if acronyms CCM was already defined, otherwise define it here

This acronym was not yet defined, thank you.

6184 L19: “as A result”

Rephrased from “As the result” to “As a result”

6184 L25-onward: the part “significant cooling […] leading to modified patterns” is unclear to me: is “lead” to be intended in causal sense? At least in the case of NAO/AO the response is due to the downward propagation of a strengthened stratospheric polar vortex associated to in situ thermal effects of volcanic aerosols.

We do totally agree that the NAO response is due to the downward propagation of positive anomalies in the stratospheric polar vortex. Hence we rephrased to:

“Generally, a significant cooling of the surface occurs in the first weeks after major volcanic eruptions, lasting for one to two years and being associated to modified patterns of precipitation, surface pressure and the teleconnection patterns, such as the Arctic Oscillation (AO), North Atlantic Oscillation (NAO) (Shindell et al., 2000; Stenchikov et al., 2002; Fischer et al., 2007) or the El Niño Southern Oscillation (ENSO) (Robock and Mao, 1995; Adams et al., 2003), due to the downward propagation of positive anomalies in the stratospheric polar vortex strength.”

6185 L10: Feulner … their

Was rephrased to Feulner … his

6185: it could be worth including Zanchettin et al. (2013) in the list of previous modeling works focusing of the early 19th century, especially since they demonstrate how the simulated decadal climate response to the Tambora eruption depends on the background climate state, including the set of considered forcings and the ongoing internal variability.

We agree that including the publication of Zanchettin et al. (2013) in the list would be a benefice to this work. Hence we added on P6185 in the end:

“Zanchettin et al. (2013) investigated the decadal response change of the Tambora 1815 volcanic eruption to different background climate states. They found a significant dependence to background conditions when looking at ocean dynamics, especially concerning heat transport and sea ice in the North Atlantic region.”

6186 L18: please include acronym MPIOM at this point. Also, please provide in the following also details for the configuration/resolution of MPIOM.
We modified the sentence as such on P6186L18
"The AO-CCM SOCOL3-MPIOM emerges from a modification of CCM SOCOL version 3 (Stenke et al., 2013), which has been coupled with the OASIS3 coupler (Valcke, 2013) to the Max Planck Institute ocean model (MPIOM, Marsland et al., 2003)."

and P6187L18

"The ocean is run in GR30 resolution (nominal resolution of around 3°). Its north pole is displaced to Greenland, making it possible to raise the resolution in the North Atlantic basin."

6188 L3: THE state-of-the-art (?)

We adopted this change.

6188 L21/Figure 1d: how does the volcanic forcing compare with estimates from previous modeling studies? Along this line, it would be important to also provide top-of-atmosphere radiative fluxes (solar, thermal, net).

We agree that a comparison would increase the quality of this work. However, additional top-of-atmosphere radiative fluxes here seems superfluous as the solar input is shown in a and b of Figure 1. We however decided to change P6188 L21 to:

"The globally averaged effect on incoming surface shortwave radiation is shown in Fig. 1d, and shows higher anomalies than those of Crowley (2000) or Robertson et al. (2001)."

6188 L22: is the same QBO nudged to all experiments? Can you please discuss a bit more the associated caveats?

The same QBO is nudged to all experiments, which avoids any additional disturbances. We have difficulties finding important caveats which should be discussed here. The most important feature is that QBO is represented, and that QBO does not change from sensitivity experiment to sensitivity experiment. Hence, as we focus here on the effect of solar vs. volcanic forcing, excluding any important background condition sensitivity study (unlike Zanchetti et al. 2013), the piece of information given on P6188L22 is sufficient.

6189 L13: as a side note, I found intriguing that no simulation considering full-band spectral solar irradiance changes was performed (BU+TD).

This is not particularly intriguing, as the focus on this work is to discern between the effect of bottom-up, top-down and volcanic eruptions. The control run ALL includes BU+TD, but of course also VOLC.

6191 L11: warming \rightarrow warm ?

We reworded to “warm”.

6191 L15: is there any implication from sea-ice for the warm anomaly west of the Antarctic Peninsula? To this regard, dynamics in the Antarctic region simulated by ECHAM5/MPIOM in T31/GR30 resolution suffer from an unsatisfactory representation of the southern mid-latitude westerlies. Is this the case also for SOCOL3/MPIOM? Can this be discussed a bit more?

Unhappy to discover this, we have to admit that the sea-ice argumentation vanished from the final version of the discussion paper. We hence reworded P6191L15 to:
“The western Antarctic Peninsula warming is associated with an enhanced transport of milder air masses from the subtropics, leading to a slight, but significant sea ice melting (not shown).”

Any deficiencies or unsatisfactory representations are discussed in Muthers et al. (2014, in prep.).

**6191 L19:** it may be worth discussing the proposed mechanism with the general response mechanism described by Wang et al. (2012)

We agree and reworded the subsection as such:

“The major warming over the Bering Sea originates from a strengthening of the northward surface winds inducing a positive meridional wind stress anomaly above the northwestern Pacific and the opposite – namely a weakening of the northward surface winds inducing a negative anomaly of the meridional wind stress – in the northeastern Pacific region (not shown). This facilitates ocean upwelling via the Ekman mechanism at this region, where deep water upwelling prevails (oceanic conveyor belt). The surface water of the northern Bering Sea region, cooling down during the winter season, is being replaced by deeper, older water from the thermocline region, which has no imprint of the volcanic signal yet, as indicated by a slight increase of the modeled vertical ocean mass transport in the winter season in that region. The warming signal is being so strong that it persists throughout the year. The same warm anomaly was also found by Wang et al. (2012), which explained the argued the finding by weakening surface westerly winds due to a strengthening polar vortex. Forming a positive surface pressure anomaly, net heat fluxes and ocean advection in the Northern Pacific region are modified. Although corroborative, these results should be confirmed by using a higher number of ensemble members to assure its robustness, which would go beyond the scope of this work.”

**6192 L21:** basing -> based

Adapted.

**6192 L26:** please be more accurate, like: “A subsequent surface warm anomaly at high latitudes is the consequence”

We agree this sentence may be not clear enough. We reworded to: “There, surface warm anomalies are the consequence”.

**6193 L11:** “mid-latitude westerlies”

We adopted the introduced change.

**6193 L22:** I would avoid the formulation “normal condition”. Is normality the mean unperturbed state or does it refer to some range around the mean?

We agree that “normal” is a useless term. We reformulated as such: “After the temperature minimum following the Tambora eruption (1815), the modeled temperatures show a slight recovery, but do not completely reach unperturbed conditions.”

**6194 L7:** as well as
**6194 L15:** temperature increases

Adopted.
as a side note, a similar delayed warming over continental regions after strong volcanic eruptions has been detected in the ECHAM5/MPIOM-based “Millennium” simulation ensemble (Zanchettin et al., 2012).

The work of Zanchettin et al. 2012 was already known to us. We think it is worth citing it in this paragraph as such:

“The temperature increase in the NH from 1813 to 1820 back to unperturbed temperature levels – and even positive anomalies in the 1820s – represents a supercompensation-like feature simulated by our model after each strong volcanic eruption. As it will be shown later, this warm anomaly pattern is caused by oceanic influence – a finding which agrees to a similar study of Zanchettin et al. (2012).”

“persistence […] is more constant” reads strange, maybe rephrase?

We agree this sounds strange. We reformulated to:

“On the other hand, the BU-signal among all layers is more persistent than the volcanic imprint due to the lack of “peaks” of activity.”

please report whether all these reconstructions are consistently ocean+land (or land only) estimates.

All reconstructions are Northern hemispheric temperature reconstructions, partly basing on RCS as ice cores, tree rings and coral proxies are unevenly spaced.

warming -> recovery from the cold anomaly?

We prefer the notation of reviewer #2 and reformulated to:

“Finally, the model simulation shows a recovery from the cold anomaly after 1836, which can be explained by a general increase in solar irradiance at the end of the DM as well as dilution and removal of volcanic aerosols in the stratosphere.”

isn’t it more over the Pacific warm pool region?

It is correct that the pacific warm pool region describes better the region concerned. We reformulated to:

“Furthermore, a sharp decrease in precipitation both during the boreal summer and winter is modeled over the Pacific warm pool region, eastern Central America, and the maritime continent.”

change in SST: I guess this is cooling, can you explicit it?

We clarified this by changing to:

“In contrast, the precipitation anomaly over the western Pacific is related to a decrease of sea surface temperatures in the El Niño 3 region, which is consistent to a reduced evaporation, a modified circulation and a significant change in the ENSO signal, impacting – via the atmosphere – also the precipitation patterns in the Indopacific region – corresponding to the mechanism presented in McGregor and Timmermann (2011).”

is the “both” at the right place?

Yes, both is at the right place.

significantly weakened: yes but only in its upper branch, right?

This is correct. We specified to:
“During the boreal summer season (JJA), the upper branch of the Hadley cell is significantly weakened (Fig. 8a) – most probably due to the volcanic eruptions (Fig. 8c).”

6201 L1-2: is the mentioned cool period referred to reconstructions? If so, I do not see it in Figure 6, where the gray shading stays mostly above the zero line during the 1820s. This seems also to contradict a later statement (6201 L19-23).

We totally agree that this sentence is misleading. We changed the sentence as following: “On the other hand, our volcanic eruptions experiment exaggerates the recovery of the surface temperatures in the 1820s.”

Figure 4: the temperature response to the 1809 and Tambora eruptions seems very similar, despite the different aerosol loads. Any comment on this?

First seemingly contradictory, this finding is due to the differing aerosol load in the northern hemisphere than in the southern hemisphere. As the timing of the volcanic eruptions was different (January vs. April), the transport of the aerosols differed and was proportionally higher for the northern hemisphere during the 1809 eruption.

Figure 6 caption: shouldn’t be Figure S5?

No, why? Figure 6 shows the reconstructions as an envelope. Figure S5 shows all reconstructions one by one.

References: