**Initial text (last two paragraphs of section 3.3):**

The main difference of all model experiments to SAGEII-derived data and especially to the latest SAGE-3𝜆 composite in Fig. 7 is the presence of a large amount of aerosol mass in 1991 in the lowermost stratosphere, i.e. below approximately 60 hPa, which is not consistent with SAGE. Potentially the SAGE II-derived data can be still influenced by the known problems in observing the lower stratosphere, which became opaque for limb-occultation instruments in 1991 (Russell et al., 1996). This is partly confirmed by comparison of SAGE II-derived data with HIRS and AVHRR in previous sections. However, recently Revell et al. (2017) analysed the stratospheric warming after Pinatubo using SOCOLv3, which has the same dynamical and chemical cores as SOCOL-AER, but used prescribed aerosols from the SAGE-4𝜆 and SAGE-3𝜆 composites. They found that model simulations driven by SAGE-3𝜆 aerosols are much closer to temperature reanalyses than simulations driven by SAGE-4𝜆, which also has more aerosol mass present in the lowermost tropical stratosphere.

We added results of both experiments from Revell et al. (2017) to our Fig. 6 (dashed and dotted blue curves). Analysis of late 1991 reveals that model results driven by SAGE-3𝜆 are also biased compared to reanalysis temperature but to the opposite direction than REF. Purely radiatively, this fact suggests that the sharp cut of SAGE-3𝜆 aerosol cloud below 60 hPa in late 1991 is not realistic and there should be something in between the REF and SAGE-derived data. However, the dynamics is also highly involved in this region, since modification of warming at different levels also causes changes to the tropical upwelling and therefore adiabatic cooling of higher levels as well as aerosol redistribution causing further changes to local radiative effects. Besides this, the extra-tropical wave-breaking and thus the stratospheric residual circulation is also modified with further consequences for the tropics (e.g. Toohey et al., 2014). Comparison of REF and SAGE results in Figs. 6 and 7 for period after 1991 also suggests that relation of the vertical distribution of aerosol mass and the resulting warming is nonlinear and needs further detailed investigation separating dynamical and radiative effects.

**Modified text:**

The main difference of all SOCOL-AER experiments to the SAGEII-derived data is the presence of a large amount of aerosol mass in 1991 in the lowermost stratosphere, i.e. below approximately 60 hPa, which is not consistent with the SAGEII-based composites, in particular SAGE-3𝜆 (Fig. 7). Potentially the SAGE II-derived data can be still influenced by the known problems in observing the lower stratosphere, which became opaque for limb-occultation instruments in 1991 (Russell et al., 1996). This is partly confirmed by comparison of SAGE II-derived data with HIRS and AVHRR in previous sections. Recently, Revell et al. (2017) analysed the stratospheric warming after Pinatubo using SOCOLv3, which has the same dynamical and chemical cores as SOCOL-AER, but with prescribed aerosols from the SAGE-4𝜆 and SAGE-3𝜆 composites (dashed and dotted blue curves in Fig. 6). While our SOCOL-AER-based REF simulation overestimates the temperature response at 70 hPa in 1991, both SOCOLv3 simulations driven by the SAGEII-based data sets are low-biased compared to the reanalysis temperatures (Fig. 6 bottom). From a pure radiative point of view, neglecting dynamical feedbacks, this suggests that the correct aerosol mass loading below 60 hPa lies between our REF simulation and SAGEII-derived data.

At higher altitudes, i.e. around the level of maximum aerosol loading, all SOCOL-AER simulations as well as the SOCOLv3-SAGE4𝜆 results show a linear relation between aerosol mass and the resulting warming, while the SOCOLv3-SAGE3𝜆 results show a different behavior: despite having the largest aerosol mass among all considered cases, the simulated temperature response is smallest (Fig. 6 top). Even considering potential dynamical effects (e.g. Toohey et al., 2014), the SAGE3𝜆-based results are surprising. This encouraged Revell et al. (2017) to reassess their results, and they noticed an issue with the spectral integration of the SAGE-3 extinctions. The effect of this spectral integration issue on the comparison between SAGE-3, SAGE-4𝜆, and SOCOL-AER will be investigated and discussed further in an upcoming study.