Responses to the comments of Referee #2

We thank Referee #2 for the thoughtful comments and suggestions, which certainly helped to improve the clarity of the manuscript. Please find below our detailed point-by-point reply to the comments, which we hope have addressed all satisfactorily.

1) The authors raise the issue of assimilated data biases as a factor in the poor model performance at capturing transport in the wake of the major SSW in January. But another important issue is the use of nudging restricted below 1 hPa. The mesospheric circulation is not a linear problem and is not fully determined by the dynamical boundary condition at 1 hPa. In fact, multiple solutions are possible for the mesospheric state given the same set of stratopause conditions. In other words, it is not really possible to approximate nudging of the mesospheric circulation by restricting the circulation below 1 hPa. In the case of a CTM like FinROSE, the circulation needs to be specified in the mesosphere and the quality of the ECMWF data used is not clear and likely to be deficient (for example, gravity waves become increasingly important with altitude but are treated as noise in standard data assimilation systems.)

Reply: We agree that restricting the nudging of the models to below 1 hPa leaves the mesospheric circulation largely unconstrained, thus making it more difficult to compare to observations. However, the major aim of the paper is to evaluate the capacity of climate models to simulate a realistic polar winter NOx descent. This could hardly be assessed if the mesospheric circulation was entirely constrained by nudging. The restriction of specified dynamics to the stratosphere is a compromise that is hoped to provide a realistic evolution of mesospheric meteorology by upward control, while still allowing for the assessment of self-generated tracer descent in the models.

I would think that nudging restricted to below 1 hPa is a bigger problem than the biases in the assimilated data in the stratopause region. In particular, the non-orographic gravity wave drag schemes are essentially running in free mode in the mesosphere and impose biases since there is no local nudging to offset them. Non-orographic gravity wave drag cannot be constrained to the observed regime by imposing the right winds in the stratosphere (e.g., Scinocca and McLandress, 2005). Various built-in assumptions such as a constant source spectrum will still impact the response.

Reply: Yes, we agree. But, as mentioned by the reviewer, the GWD schemes are running basically in a free-running mode and therefore one would expect a more random-like spread of model results. However, our results suggest a systematic model bias with respect to the temperature distribution and the timing of NOx descent in the lower mesosphere. How can such a systematic behaviour be explained by the lack of nudging above 1 hPa?

The basically free mode simulation of the mesosphere also affects the radiative transfer calculations. Radiatively active trace gas distributions are not likely to be ideally distributed to conform to observations. So the radiative damping impact on the evolution of the state in the wake of the major SSW will deviate from observations.

Reply: This is a good point! The potential impact of tracer distributions on the radiative cooling/heating will be mentioned in the revised version.

In addition the evolution of the highly nonlinear SSW in the mesosphere is not guaranteed to follow observations even if the wave amplitudes at the stratopause conform to observations. The progression of SSW events is rather complex (Matthewman et al., 2009) and the mesospheric component of a major SSW cannot be trivially constrained by the stratospheric part. The SSW exhibits a high degree of top-down evolution with the mesosphere being an important layer of the atmosphere for this phenomenon.

Reply: We agree that SSWs exhibit a high degree of top-down evolution. However, we'd like to mention in this context that it is presently not known if the mesosphere has any control on the stratosphere or, if the mesospheric SSW precursors are triggered from below.

It appears that relatively good performance of EMAC (Fig. 13, 14, 15, 16) and its application of assimilated data up to 0.2 hPa may not be a coincidence since every scrap of nudging in the mesosphere counts.

Reply: EMAC applies the NOx UBC down to 0.1 hPa. This represents a trivial reason for a better performance with respect to the representation of NOx in the region above this level. Note that this model significantly overestimates the descending NOx amounts in the post-SSW phase below 0.2 hPa (see Figure 17) despite the application of assimilated data in this region.

FinROSE appears to do a better job as well (Fig., 13, 14) but performs worse than EMAC apparently due to issues with the vertical upwelling.

Reply: We'd like to point out that the model performance with respect to NOx descent in the pre-SSW phase (Figures 13 and 14) is generally good (except for those models with identified issues regarding the UBC implementation or missing unresolved GWD).

CTMs use geostrophic balance to infer the upwelling from the temperature and horizontal winds, but this fails to capture the ageostrophic part which is not negligible in the mesosphere due to the amplification of gravity waves and general deviation from geostrophic balance.

Reply: We agree with this view.

In addition, several models linearly ramp nudging above 40 km so that they are not imposing as much of a constraint on the upper stratosphere and lower mesosphere dynamical state as models that fully nudge up to 1 hPa.

Reply: Also agreed. We will mention this explicitly in the revised version.

The authors mention the study of Siskind et al. (2015) which highlighted the need for nudging to much higher altitudes (92 km) to improve simulations of NO descent compared to observations. But there is no focused discussion of the limitations of the nudging approach adopted by most models in this intercomparison. Such a discussion has to be included either in the introduction or in the conclusions sections to put the results from these models in a better context.

Reply: Such a discussion will be added.

References:

N. J. Matthewman, J.G. Esler, A.J. Charlton-Perez and L.M. Polvani: A new look at stratospheric sudden warmings. Part III. Polar vortex evolution and vertical structure, J. Climate, 22, 1566-1585 (2009).

J. Scinocca, and C. McLandress: The GCM Response to Current Parameterizations of Nonorographic Gravity Wave Drag, J. Atmos. Sci., 62(7), 2394-2412, 10.1175/JAS3483.1, 2005.