We wish to thank the referee for his/her helpful comments and for this positive acknowledgment of the improvements of our manuscript. The full review is copied hereafter and our responses are inserted in bold.

I thank the authors for investigating the lack of sensitivity through additional computations. Those tests are useful. Of course there is a strong non-linearity as the increased NOx emissions tend often to increase OH and decrease the lifetime of NOx. There might be other chemical effects, which would require a more detailed scientific study, as the authors state in the conclusions. Besides chemical effects, the lack of sensitivity is further enhanced by 1) the (presumably small) contribution of non-anthropogenic emissions, and 2) the contribution of emissions during the preceding days. I think this should be mentioned in the manuscript.

We agree that the contribution of non-anthropogenic emissions could enhance the non-linearity during a large part of the year but their impact should be insignificant for our illustration case in February, when biogenic emissions are very small.

We have added the following sentences in section 4.2.4: "We can conclude that the strong non-linearities of the NO_x chemistry mainly explain the lack of sensitivity between NO_x emissions and satellite NO_2 columns. Besides chemical effects, the lack of sensitivity could be also partly due to the contribution of emissions during the preceding days and the assimilation window will be widened in the near future."

I can't say I agree entirely that the inversion system fails to reproduce the patterns of the observations **because** of this non-linearity. It does not help, of course, in the sense that very large emission increments are needed to overcome the negative feedbacks and match the observations. And very large increments are penalized in the cost function. In your test with anthropogenic emissions multiplied by 3, the relative emission increment was uniform, whereas in the optimization, the system is free to modify the emission distribution. In a test with infinite a priori emission errors and no spatial correlation, the system would very probably do a better job. In your setup, with correlations and conservative emission error estimates, the system finds a compromise (which is perfectly reasonable). I would appreciate if the discussion could reflect the fact that the choice of errors and correlations has a likely strong impact on the results.

We agree and we have already demonstrated the strong impact of the prior error standard deviations and of the spatial correlations in the B matrix in section 4.2.2. We have added a sentence in section 4.2.2: "To our knowledge, there are few available studies dealing with the estimates of the uncertainties in gridded bottom-up emission inventories at the 0.5°x0.5° resolution or higher. The characterization of their statistics in the inversion configuration is consequently often based on crude assumptions from the inverse modelers. **Defining the covariance matrices B and R is not an easy task, while incorrectly specifying these matrices has a very strong impact on the results of the inversion. Especially, the relative weights of B and R, and the spatial and temporal correlations in B influence the degree of freedom and the structure for the adjustments attempted by the inversion in the optimization process.** Consequently, as an example for the NO_x inversion, different sensitivity tests described in Table 3 have been performed for the construction of the **B** matrix**."**