In this manuscript, the authors report on a study quantifying the effect of 3d cloud effects on satellite retrievals of tropospheric NO₂. Such effects are currently not taken into consideration but are expected to play a role at the spatial resolutions of current and future instruments. The study identifies and quantifies different effects on simple synthetic test data concluding that the cloud shadow effect is dominant. This error is then analysed systematically and different methods are developed to correct it using albedo fitting, O₂-O₂ scaling or a parametrisation approach. All of them result in a reduction of the errors at least in the clear sky part of the measurements, both in the simple cases studies and in a complex cloud scenario. Finally, two case studies are presented on real TROPOMI data indicating that the developed approaches might improve NO₂ retrievals close to cloud edges.

The manuscript is very interesting, clearly structured and written and contains a wealth of good ideas and relevant results. I, therefore, recommend publication with only minor revisions as listed below.

**Detailed comments:**

It would be good to have a short discussion of what the expected effect of aerosols is on the discussed 3d effects which here are discussed in a Rayleigh atmosphere.
Page 2, last line: This sentence is a bit unclear as spatial heterogeneity will also be relevant in clear sky scenes and several effects are addressed at the same time here. Please separate into two (or more) sentences.

Page 3, line 14 / 15: It would be nice to have a very brief indication also of what Várnai et al. found in their work.

Page 9, line 21: I think it would be good to iterate here that only one aspect of possible errors introduced by cloud correction is covered. Perfect knowledge of all parameters is assumed and in particular, the NO\textsubscript{2} profile is assumed to be the same inside and outside of the cloud.

Figure 2: I think that this display is somewhat misleading – I was tempted to see points close to the 1:1 line as “good” points while in reality, they are just points for which both cloud retrievals perform similarly. The main point of the discussion here is how large errors are and I think histograms of relative errors would be more appropriate.

Figure 10: It would be nice to have the same x-axis in both plots to allow direct comparison.

Section 4.1.1 It would be interesting to add a short discussion of what you think about the surface albedo fitting implemented in the current TROPOMI lv2 product where the surface albedo is determined from radiance in case it is lower than the climatological value for a scene.

Cases where the retrieved albedo is 0 appear to be problematic – can you discuss this a bit more? Is that because the atmosphere is illuminated less than it would without cloud which reduces the backscattered intensity but does not change the layer AMF in the same way as a small albedo?

The application to TROPOMI data is based on the assumption that NO\textsubscript{2} retrievals should yield the same column in cloudy and clear regions as well as in the cloud shadow. However, considering the reduced actinic flux in the cloud shadow (and the increased values inside the cloud), shouldn't we actually see differences?