Dear reviewer,
thank you for reviewing our paper and your helpful comments.

1 General comments: HCl discrepancy

We agree that this is a very important issue and deserves more discussion. We have now moved the discussion to a separate section in the main text and considerably expanded the discussion on the reasons of the discrepancy and its effect on the results. The results of the original uncorrected runs are now shown in a supplement. The supplement also contains comparisons to ClO from MLS and $ClONO_2$ from ACE-FTS now. We added some important references that were missing.

It is important to note that this is not a model deficiency specifically related to the ATLAS model, but that discrepancies in HCl between model and observations are a well-known problem in many stratospheric CTMs, e.g. SLIMCAT (Santee et al., 2008), SD-WACCM (Brakebusch et al., 2013, Solomon et al., 2015) or MIMOSA-CHIM (Kuttippurath et al., 2015). Interestingly, the SLIMCAT CTM shows a discrepancy of the same order of magnitude, but with opposite sign. For some models (e.g. KASIMA, CLaMS, see below) we were not able to find sufficient information in the literature. Unfortunately, discrepancies of this order of magnitude are still "state-of-the-art" in CTM modelling. This points to a gap in our understanding of the chemical and physical processes involved here.

Unfortunately, this was not expressed as clearly as it could have been in the original manuscript. We have now added some additional general discussion of the problem in the introduction and further references.

Unfortunately, discussing this issue in all the detail it deserves is out of the scope of this paper. This issue is so important that it would well deserve its own study summarizing the problems in the different models.

1 General comments: CCMs and CLaMS

We disagree with your assertion that this problem does not occur in "state-of-the-art" CCMs and CLaMS, which also seems to imply implicitely that it does not occur in other CTMs in general.

- **CCMs**: The assertion that the problem is not evident in CCMs is not correct. Fig. 6.32 in the 5th SPARC report (CCMI, Eyring et al., 2010) shows large differences in HCl compared to MLS at the 500 K level, easily exceeding 1 ppb. In May to July, many CCMs overestimate HCl, as it is the case also for ATLAS.

  Even more importantly, a comparison to CCMs not nudged to meteorological observations is not very meaningful. It is not possible to decide if differences in HCl between model and measurements are caused by e.g. temperature biases or biases in vortex strength or by problems in the chemistry and microphysics. The comparisons in Eyring et al., 2010 show large discrepancies between MLS and the models and large differences between models, but it is not clear what the reasons are.

Since it can't be decided if the problems in the CCMs come from not nudging them, we don't refer to CCMs in the discussion in the introduction.

- **CLaMS**: Your assertion that the problem does not show up in CLaMS can't be proven from the existing literature. Either the initialization of the chemical species is too late or comparisons of CLaMS model results for the polar vortex to HCl or $ClONO_2$ measurements are not shown for the relevant time period. There is just not enough information available to assess if there is the same problem in CLaMS or not.

  The initialization in the CLaMS paper that you cite here for the Arctic (Grooß et al., 2002) is too late for the problem to show up (10 February). As far as I understand the paper, Cl_y is taken from a tracer-tracer relationship, while the partitioning is taken from the Mainz 2-D model (it is however unclear to me from the text if that was done on 10 February or 8 January). That boils down to the question how well this model performs, which could easily overestimate the flux of NO_x across the vortex edge and hence could overestimate the reformation of $ClONO_2$. In addition, there is only a single comparison to HCl for a single ER-2 flight in March, which provides not enough information to assess if there is a problem in this particular run.

  We only know of 2 CLaMS publications which deal with the Antarctic winter, where the problem is most pronounced (Grooß et al., 2005, 2011). These studies don't show the initialization of the chlorine species or their development in May and June. The initialization in Grooß et al. (2005) is too late again for the problem to develop (August) and again depends on the Mainz 2-D model, but comparisons of HCl to measurements are shown. While the initialization in Grooß et al. (2011) is early enough, it shows only the HCl mixing ratio on a single, not representative trajectory and does not compare HCl to measurements. The only hint if the problem is evident in CLaMS is a reply to the reviewer in the discussion of Grooß et al. (2011), which states that CLaMS *overestimates* HCl (Reviewer 1, reply to major comment 1, C12059), which would point in the direction that CLaMS overestimates HCl as well, in contradiction to what you state in your review.

2 Specific comments

- Appendix A: "What happens if the analysis is done with the vortex criterion?"

  Unfortunately, the computational effort is much higher when comparing the modelled values to the MLS data under consideration of the vortex tracer criterion. Without the vortex tracer criterion, it is sufficient to calculate simple vortex averages. Considering the criterion requires calculating trajectories from the measurement time and location of every MLS measurement to the ATLAS output closest in time, sorting out ATLAS points with too low values of the vortex tracer and running the chemical box model forward on the trajectories then. Such point-to-point comparisons of measurements and model for the species measured by MLS were done for the 15th of every month of the simulation and are now shown in a supplement (regardless of the value of the vortex tracer), but were not shown in the original paper to limit the manuscript to a reasonable length. Examples for these plots are attached to this reply for the uncorrected run for 15 June 2006 (Figure 1), the corrected run for 15 June 2006 (Figure 2), the uncorrected run for 15 July (Figure 3) and the corrected run for 15 July 2006 (Figure 4). It can be seen that later in winter, a HCl "collar" region develops in the uncorrected model, where enough NO_x is available to replenish $ClONO_2$. This is not visible either in the MLS data or the corrected run. That means restricting the comparison to the vortex core would make the comparison between the uncorrected run and MLS even worse, with higher values for HCl for the black line in Figure 22 of the manuscript.

- Appendix A: "Is there a problem with photolysis rates for twilight conditions?"

  We have the impression that the HCl problem is not very sensitive to photolysis reactions. The main problem is the lack of $ClONO_2$ for the $HCl+ClONO_2$ reaction. In addition, if there would be a problem with the photolysis, it should show up in other species as well. Most other species compare well with measurements, however (see supplement). There is some disagreement of the modelled NO_x species to measurements of the ACE-FTS instrument, which by design measures under high solar zenith angles, but this could well be a problem of the unfavourable combination of a large satellite footprint with high gradients in mixing ratios.

  If there is a problem is very difficult to tell, since it requires comparison of short-lived species to measurements under twilight conditions. This is difficult, since there are not so many measurements of short-lived species to compare with, the mixing ratios are low in twilight, and for satellite measurements, the satellite footprint will probably cover a relatively large range of solar zenith angles under twilight conditions.

  The treatment of photolysis is pretty much standard in ATLAS compared to other models. Photolysis rates in the photolysis tables are calculated up to solar zenith angles of 100 degrees based on the photolysis coefficients given in the JPL catalogue. The spherical geometry is considered in dependence of the altitude. New solar zenith angles are calculated every 30 minutes and are then linearly interpolated to the time steps of the solver.

- Appendix A: "Is the problem related to reactions on NAT and ice and the assumption of a constant supersaturation at ECMWF grid point temperatures ignoring mountain wave effects (page 3)?"

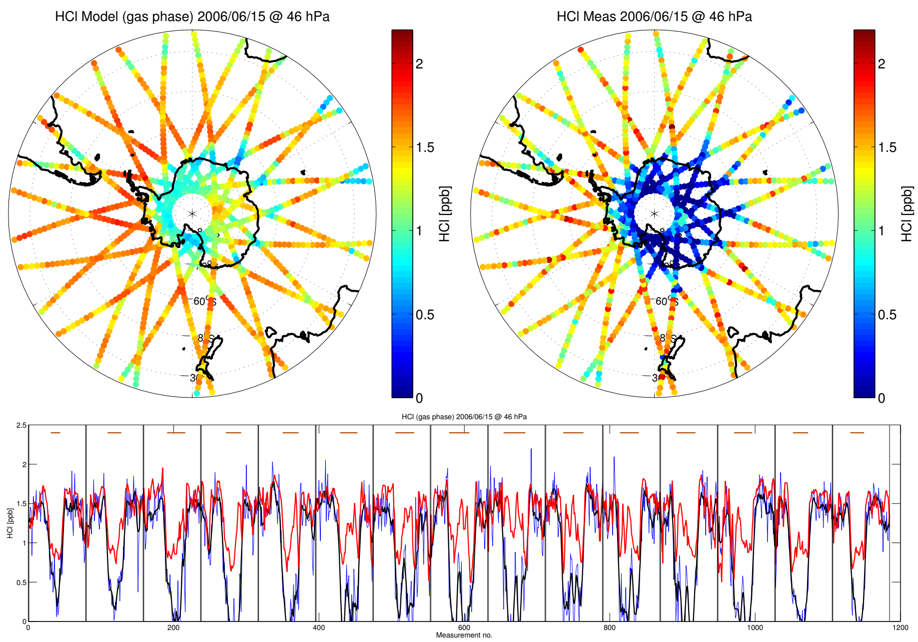

Figure 1: HCl at 46 hPa on 15 June 2006 in the uncorrected run. Left: ATLAS, Right: MLS. Bottom: Line plot of model (red) and MLS (black) along the satellite tracks

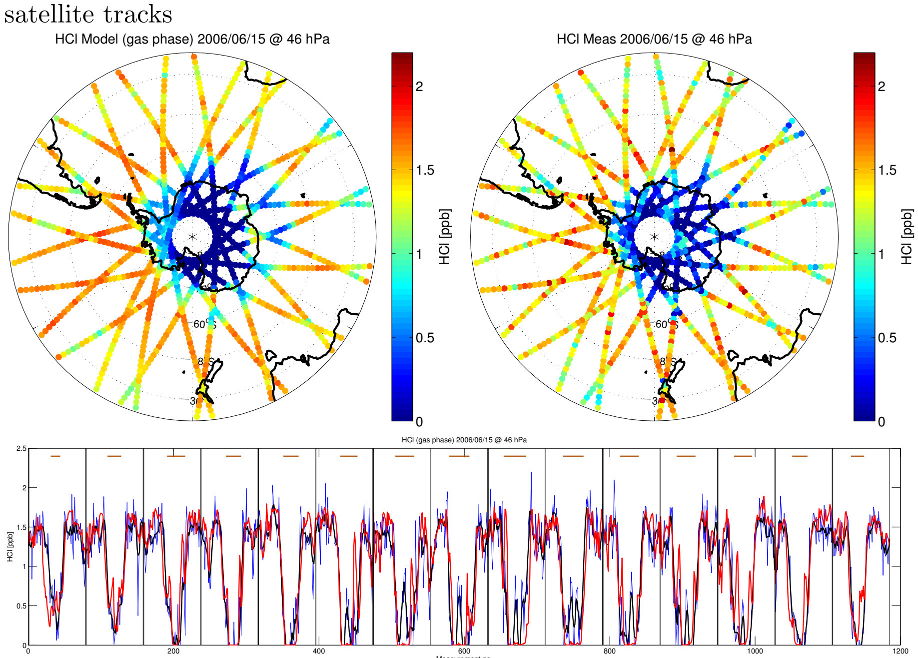

Figure 2: HCl at 46 hPa on 15 June 2006 in the corrected run. Left: ATLAS, Right: MLS. Bottom: Line plot of model (red) and MLS (black) along the satellite tracks

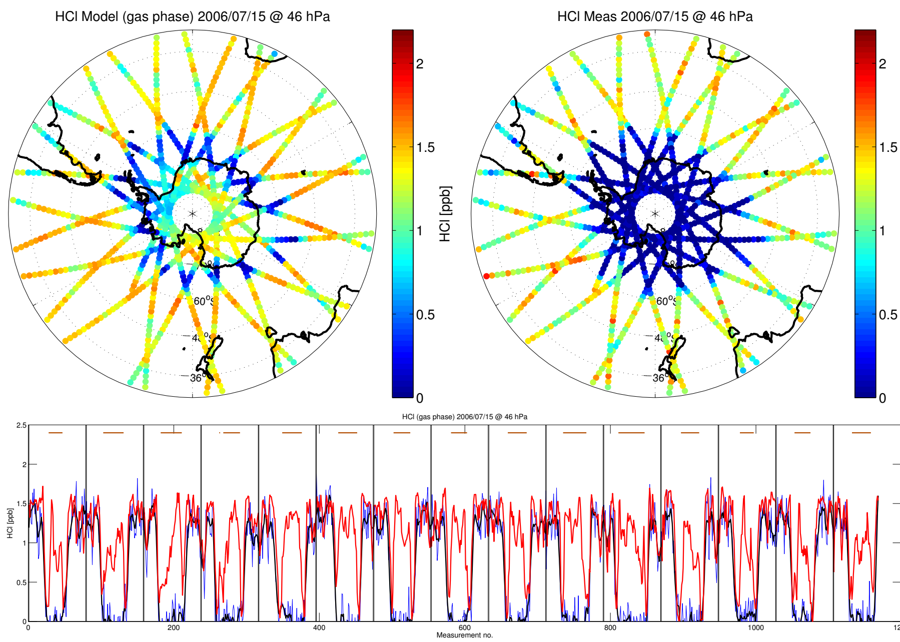

Figure 3: HCl at 46 hPa on 15 July 2006 in the uncorrected run. Left: ATLAS, Right: MLS. Bottom: Line plot of model (red) and MLS (black) along the satellite tracks

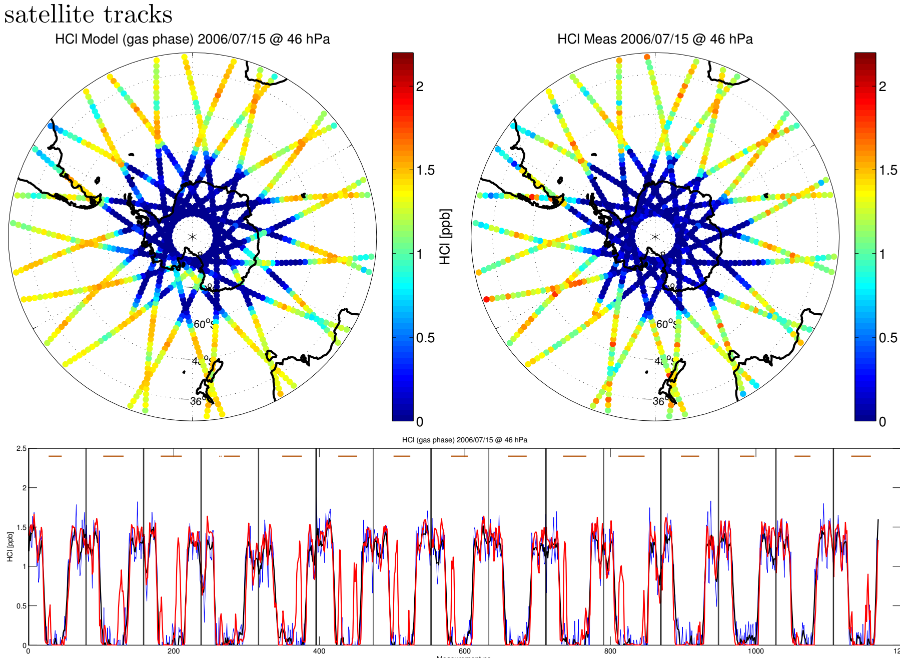

Figure 4: HCl at 46 hPa on 15 July 2006 in the corrected run. Left: ATLAS, Right: MLS. Bottom: Line plot of model (red) and MLS (black) along the satellite tracks

This probably cannot explain the discrepancies. The problem with the $HCl + ClONO_2$ reaction is not that is it not fast enough or that the area covered by the reaction is too small, it is simply that there is no $ClONO_2$ as a reaction partner. It is also very likely not related to the $HOCl + HCl$ reaction. The discrepancy develops in June in the southern hemisphere, but the $HOCl + HCl$ reaction needs sunlight and ClO_x and does not really start to be important before August. In addition, NAT and ice clouds play a minor role compared to STS clouds in these model runs.

- Page 6, line 17: Do you mean that there is a study that we could cite here or do you mean we should perform a study with MIPAS data? Please clarify. In the first case, can you please give the reference?
- Page 22, Fig. 13: We agree. It is however extremely difficult to find a model parameterization that exactly matches the time evolution and spatial characteristics of HCl from the measurements and is at the same time based on some plausible assumptions about the reason of the HCl discrepancy. Note also that this unfortunately is "state-of-the-art" and that other models show discrepancies in the same order of magnitude (Santee et al., 2008, Brakebusch et al., 2013, Solomon et al., 2015, Kuttippurath et al., 2015). I.e., this is the best one can do with the current knowledge.
- Page 37, Wegner reference: We have to admit that citing this conference abstract was a little bit unfortunate, since it is not publicly available and there are at least two published studies on the HCl discrepancy in SD-WACCM, which we could have cited. We have added two references for SD-WACCM (Brakebusch et al., 2013, Solomon et al., 2015) and we have added a reference to the Ph. D. thesis of T. Wegner, which is available online.

3 Technical corrections

- Thanks for pointing me to the incorrect year. City has been added.

References

- Brakebusch et al. (2013), J. Geophys. Res., 118, 2673–2688, doi:10.1002/jgrd.50226
- Eyring et al. (2010), SPARC Report No. 5
- Grooß et al. (2002), J. Geophys. Res., 107, doi:10.1029/2001JD000456
- Grooß et al. (2005), J. Atmos. Sci., 62, 860–870, doi:10.1175/JAS-3330.1
- Grooß et al. (2011), Atmos. Chem. Phys., 11, 12217–12226
- Kuttippurath et al. (2015), Atmos. Chem. Phys., 15, 10385–10397
- Santee et al. (2008), J. Geophys. Res., 113, doi:10.1029/2007JD009057
- Solomon et al. (2015), J. Geophys. Res., 120, 7958-7974, doi:10.1002/2015JD023365