We would like to thank the reviewer for taking the time to evaluate our manuscript and suggest changes. These comments have lead to a substantial improvement of this work. We detail the changes we have made below.

R1) This paper is presented as the second of 2 twin papers...

Russo et al. (2011) was submitted two weeks after Hoyle et al., but actually appeared in ACPD before Hoyle et al., as suggested by the reviewer. It has now been published in ACP (Atmos. Chem. Phys., 11, 2765–2786, 2011)

R1: The paper is based on simulations from 14 different models/versions of models but the number of models represented in the different plots varies from 11 in Fig. 3 to 5 in Fig. 8 without explanations. Why are there only 5 models in section 4.4 and Fig. 8? The comparisons between simulated CO and SCOUT-O3 measurements is a very interesting part of the paper but is weakened by the absence of 9 models among 14. Results from the 14 models should be discussed in each section and displayed on each figure (except global maps, tropical means or long time series for mesoscale models). The absence of some models from some part of the paper and/or some figures should be an exception and thoroughly explained and justified.

We have included pTOMCAT_tropical in more of the plots, as new data has become available, likewise for CATT-BRAMS. We have included a table after the tracer descriptions, showing which tracers were run by which models, and we have included text explaining why it was not possible to run some tracers with some of the models. We could have excluded models which could not run all the tracers, however we prefer to present as much information on the different models as possible.

Reviewer 1) I) The title of the paper is misleading and do not correspond to the results when stating “tropical deep convection and its IMPACT on composition: : :”. First, what “composition”?

The statement is really vague! Second, the paper is not dealing with the impact of deep convection on the composition of the troposphere, TTL or UTLS. There are indeed no numbers nor figures concerning the impact of deep convection on the budget/distribution of any gases such as CO, HNO3, O3 etc. In order to do so, the study should have compared results of the models with and without convection, taking into account the problem of large-scale vertical transport as mentioned in Lawrence and Salzmann (2008).

In order to fit with the content of the paper, the title should be more something like “Intercomparison of the representation of tropical deep convection in atmospheric models – Part 2: Tracer transport”.

The title has now been changed to “Representation of tropical deep convection in atmospheric models – Part 2: Tracer transport”

R1: Furthermore, the authors should add references to Doherty et al. (2005) and Lawrence et al. (2003) who are pioneer studies concerning this topic and surprisingly not referred to.

We thank the reviewer for pointing out this omission and we now briefly discuss these two studies.
R1: p20359 lines 10-23: the authors should refer to Tost et al. (2010) who have studied the impact of different parameterizations (in the same model) of deep convection on atmospheric chemistry. We have added a brief description of the findings of Tost et al. (2010).

R1: There are a lot of references to R2010 which is not accesible! P20369 lines 18-19: the boundaries of the selected regions are very important but it is detailed in R2010! P20370 line 13-15: the authors state “the meteorological analysis of convective properties in R2010 will help to attribute differences in the model’s convective transport”. Without access to R2010 I cannot fully understand the differences in the “model’s convective transport” mentioned. Are there plots of entrainment/detrainment, and vertical mass fluxes in R2010?

R2010 includes analysis of precipitation patterns, and cloud top heights/cloud distributions. The detrainment from model transport is best represented by the tracers here however. We have included more discussion of the results of R2010 in the present analysis.

R1: P20370: the convective transport of pTOMCAT-tropical has already been compared to other models over West Africa (Barret et al., 2010) with results in agreement with what is shown here. This should be mentioned when discussing the altitude of outflow of pTOMCAT-tropical relative to the other versions of pTOMCAT. Done

R1: P20372 lines 10-15 and 21-22: the discrepancy between UMUKCA-UCAM_nud and the other models over WA is very serious and the explanation is not completely convincing “this is thought to be due: : :”. Why is the representation of surface properties deficient in this particular model over this particular region? The cause of the problem should be addressed with comparisons of the fields incriminated.

Since UMUKCA-UCAM_nud has been developed for climate studies, it uses an interactive soil scheme which allows for feedbacks between precipitation and soil moisture (this model is the only CCM and the only model using interactive soil scheme, with the exception of UM-UCAM_highres). However, for dry land areas (such as West Africa) where precipitation is the major source of soil moisture and soil moisture is a major contributor to convection initiation and precipitation, the model can get into self-feeding loops where reduced precipitation leads to dry soil which leads to reduced convection/precipitation and so on. The attached plot shows low- level (20m above the surface) moisture and precipitation from a version of the UMUKCA-UCAM model. The two plots are from identical 1-month runs for July 2005: the initial conditions for soil moisture are taken from the Willmott climatology (Willmott et al., J. Clim., 1985). The first run uses the interactive soil scheme to calculate soil moisture (similarly to UMUKCA-UCAM_nud), while in the second run the soil moisture is constrained to the Willmott climatology throughout the simulation. Even with a short run time of one month, using interactive soil scheme leads to a reduction in the low level moisture and precipitation over West Africa and this effect is expected to be even larger after a 1-year spinup (such as is used for the paper). The effect of interactive soil on low-level moisture and precipitation is almost negligible for South America. The differences between runs with and without interactive soil was found to be dependent on model horizontal resolution: the high resolution setups (similar to
UM-UCAM_highres) generally show less drying over West Africa compared to coarse resolution models. To address the referee's point in the paper we have replaced the original sentence: “this is thought to be due to poor representation of surface properties (such as surface albedo and soil moisture).” with the following: "Further investigation and additional model analysis showed that this underestimation of convection over the West Africa region is due to a problem with the interactive soil scheme used in the model. For this region, the interactive soil scheme leads to underestimation of the soil moisture compared to soil climatologies (e.g. Willmott et al., 1985), and this in turn leads to a reduction of low-level moisture and convective activity."

R1: P20374 lines 26-27: I didn't fully understand the explanation about the turbulent mixing. Is it the height of the boundary layer which is too low with the Louis scheme with convection entrainment higher than the top of the Louis boundary layer?

We have added text both to the model description section and to section 4.1 to clarify this. The height of the boundary layer is low using the Louis scheme. However, the Louis scheme does not account for entrainment at the top of PBL (the limitation by using by Louis scheme, see Stockwell and Chipperfield (1999)).

R1: The authors have chosen 3 different convective regions/season to compare the models. They should mention the Indian/Asian summer monsoon which has been shown to be responsible for the uplift of large amounts of gases emitted at the surface to the lower stratosphere where they are trapped within the Tibetan anticyclone (see e.g., Park et al., 2007 and Randel et al., 2010).

The Indian/Asian summer monsoon is obviously a very interesting phenomenon, and it can contribute to trop-strat exchange (as pointed out by the referee). It was not included in the current analysis, since this paper focuses on "Tropical" convection and the Asian Monsoon generally has different mechanism (large-scale organised convection forced by global circulation pattern) and occurs at higher latitudes outside the tropics (the referee mentions the Tibetan anticyclone which is centred around 30N). We added the following text to the Conclusion section: “Although in this pair of papers we have focussed on the representation of tropical deep convection in models, observations and modelling studies show that significant transport from the lower to the upper troposphere and the stratosphere also occurs at higher latitudes, for example during the Asian summer monsoon, where enhanced mixing ratios of lower tropospheric pollutants are found in the upper troposphere in an area over Asia and India, centred around 30N (Li et al. 2005, Randel et al. 2010, Randel et al. 2006, Park et al. 2007). The representation of these transport pathways in atmospheric models should be further evaluated in the future.”

R1: P20375 lines 20-25: the text discusses fig. 4 which shows mean annual enhancements at 90 hPa, while Ricaud et al. (2007) suggest an important convective enhancement in the UTLS only in March-April-May and not over a whole year.
We had intended this paragraph as a general introduction to section 4.2, and didn’t mean to compare the results of Ricaud et al. (2007) with the annual mean plots. We have now added a comment on the season which the Ricaud study focussed on.

R1: P20377 lines 11-12: the present study shows that above WA, convective transport the AMMA project (dedicated to the study of the WA monsoon). Studies based on spaceborne (Barret et al., 2008) or airborne (Law et al., 2010, Fierli et al., 2010) measurements have already shown that convective outflow over WA during the monsoon is maximum at around 200 hPa and is not affecting altitudes above 150 hPa. This is an excellent suggestion which improves our description of how well the models are doing. We have included these results in the paper.

R1: High correlations between surface CO mixing ratios in convective areas and UTLS CO mixing ratios seems rather natural. It is also quite normal that the models show such a correlation, especially when the convective areas have been selected based on results from the models themselves. Furthermore, the difference between models with high correlations (FRSGC/UCI, Oslo CTM2 and UMUKA\textsuperscript{nud}) and models with low correlations (TOMCAT and pTOMCAT) just highlight that convection is not reaching 153 hPa with the low correlation models. This result has been discussed previously and illustrated with Fig. 1. Therefore, unless the authors find something more to state about these correlations, I think that section 4.3, Fig. 7 and Table 3 should be removed from the paper. This section has been removed, as recommended.

R1: P20379 line 15: “On the 16th November air in the outflow of a hector was sampled”. Why are there such big differences between the different flights on November 16 as can be seen in fig. 8? It seems that one flight was not “convective”. There was only one Geophysica flight on November the 16th, but during part of the time convection was sampled, and during part of the time the non-convectively perturbed troposphere was sampled. That’s probably what looks like two different flights on the 16th. We have added text to clarify this.

R1: P20379 lines 25-28: “On the 30th of November: : : lower down, however there are substantial over-estimation of the measured values by all models”. Concerning the lowermost layers, this is almost true for the 4 days, with measured values of about 80 ppbv and modeled values mostly higher than 100 ppbv. The models produce a vertical CO sharp gradient which is not observed with the aircraft measurements. Nevertheless, the plots start at 600 hPa which is very high. Why? How do the modeled and aircraft profiles behave lower down? The part below would give an indication of the boundary layer height as seen by the aircraft and of the behavior of the model to represent this layer (this is interesting regarding what has been discussed before). As it is, it seems that the models are mixing polluted air masses above the real boundary layer. Concerning the free troposphere, it is true that the highest discrepancies between models and measurements are displayed on November 30. The plots start at 600 hPa, because there is no aircraft data available below this point. pTOMCAT\_tropical CO data has become available, and is now included here too. Further, lower level data has been included from the Egrett and Dornier aircraft measurements, and the model/measurement comparison in the boundary layer is discussed.
R1: P20380 line 10: “a lifetime of around 3 months”. Can you give the reference for such a long lifetime? The lifetime of CO in the troposphere is usually ranging from 1 to 2 months depending on location, altitude and season.

The reviewer is correct, this should of course have read “1-3 months” similar to the statement we made in the tracer description table. We have corrected this now to “with a lifetime of between 1 and 3 months, depending on location and season (the lifetime is closer to 1 month in the tropics)”

R1: P20380 lines 14-25: TES CO measurements should be described in the measurement section together with the SCOUT-O3 measurements.
Done

R1: P 20381: this second part of the section is really weak and limited to a crude description of Fig. 9! First, the results are absolutely not discussed and linked to what has been shown and discussed before. Second, satellites and models are producing maps. Why are there no maps of comparison? Third, TES measures about 2 independent pieces of information in the troposphere for CO. Why a comparison of free tropospheric CO is not given? This would help the discussion concerning the behavior of the models in the mid-troposphere.
This section now includes maps, and the line plots have been removed. The maps provide a far better opportunity for analysis than the previous plots. The discussion has been extended considerably.

R1: P20382 line 25: “all of the models” is not right. As stated above only 5 out of 14 models are discussed in section 4.4 and displayed in Fig. 8 and 9!
This sentence was not clearly phrased. We have now changed it to specify that all of the models which were shown in the TES comparison plots reproduce the general structure of the observed changes.

R1: P20382 line 27/P20383 line 4: as mentioned above, I find section 4.3 rather weak and I think the same about the summary given here.
This has now been removed.

R1: P20383 lines 5-24: the statements are very general about transport modeling and the discussion should be more focused on the particular results provided and on the models evaluated in the present study.
This section has been made more focussed by referring to the figures which support the conclusions.

Figures:

R1: Fig 1: the top altitude should be changed to 50 hPa in order to improve the visibility of the many profiles on the plot.
Done.

R1: Fig. 3: the log scale on the x-axis makes the differences between models almost unreadable. A linear scale from 0.1 to 1 would make this plot much more useful.
The difference at lower mixing ratio/higher is already present at 0.1 ppt and can be discussed in the text.

We left figure 3 as it is. One of the points we made in the text is that the difference in the stratosphere continues to increase between some models, between others it gets less. Therefore we need the low mixing ratios to be shown on the plot. Also, by maintaining the same axis as figure 1, we can show how the lifetime of the tracer affects the modelled differences.

R1: In general there are some typing and syntax errors that have to be corrected.

Done.
Modelled low-level moisture and precip with UMUKCA-UCAM