Evaluation of cloud convection and tracer transport in a three-dimensional chemical transport model by W. Feng et al.

We thank the 3 reviewers for their valued comments. These comments are given below (in *italics*) followed by our responses (in **bold**).

Reply to *Anonymous Referee #1*

This manuscript presents an investigation of the convective scheme included in the TOMCAT global CTM. Mass fluxes and precipitation rates are compared for several different model versions to investigate the effects of model resolution, aspects of the convective scheme, use of different forcing data, and the use of archived mass fluxes vs. diagnosed convection. The study is interesting, and documents the first step of a re-evaluation of transport in TOMCAT. There are however a few minor improvements which could be made in the presentation, in particular the section comparing model results with radon.

We thank the reviewer for his/her positive feedback on the paper. We appreciate the comments and very useful suggestions to improve the quality of the paper.

Specific points: Most of the analysis involves zonal mean plots. A useful addition to the paper would be maps of mass flux, and a comparison with archived mass flux or observation based precipitation data sets. Do the various changes to the model cause longitudinal changes in the convection distribution, do the changes have different effects over land and sea?

Thank you for the suggestion. In the revised paper, we will have a look at the longitudinal changes in the convection distribution and different effects over land and sea from our results.

Pg 22959, line 18-21. Please comment briefly on how you created the higher resolution data set - was it a simple interpolation of the coarser one? Or was data from another source included?

The high resolution evaporation flux dataset is an interpolation of the UGAMP T42 GCM data. However, when it is used in the model the higher resolution dataset is averaged onto the lower resolution grid, rather than an interpolation. This is a more conservative way of mapping a high resolution gridded dataset to a low resolution. We will make this clear in the revised paper.

Pg 22960, line 27. "Run P_det is the same as E_Elnewconv ...." This description is not entirely accurate, according to Table 1 it also uses a lower vertical resolution and the resolution of the evaporation flux data set is lower. There needs to be a specific discussion of how these changes affect the comparison, if at all.

That was a typo. We will change "E_Elnewconv" to "K_L31" here.

Pg 22962 line 25. Please find a better reference for the difficulties of estimating precipitation from cloud-top temperature.

We will add "Morrissey and Greene (1993)" as a better reference.
Morrissey, M.L. and Greene, J.S.: Comparison of two satellite-based rainfall algorithms using Pacific atoll raingage data, *J. Appl. Meteor.*, *32*, 411-425, 1993.

Pg 22963, line 17. The word "likely" can be removed.

OK.

Pg 22964, line 24. Remove "a"

OK.

Pg 22965, line 2. The inclusion of downdrafts and mid level convection evens out much of the asymmetry in the updraft contours between the northern and southern hemispheres, at least for the weaker updraft contours (<1E-5). A comment on the reasons for this would be good.

This is partly due to the criteria for the midlevel convection used in the Tiedkte (1989) scheme, which occurs when the large scale ascent and an environmental relative humidity of more than 90% are needed for midlevel convection. We will explain this in the revised paper.

Pg 22966. The effect of changing only the vertical resolution should be discussed (i.e. the difference between A_E40 and K_L31) and K_L31 should be included in Figure 5.

We will add A_E40 and K_L31 in Figure 5 and discuss the effect of changing only the vertical resolution in the model.

Pg 22966. The different horizontal resolutions are compared. Run E_Elnewvap uses ERAinterim data while the others use operational data. Although this will not change the results, it should be pointed out to the reader.

We will add "Please note that E_Elnewvap uses ERA-Interim analyses while the others use operational forcing files".

Pg 22967, line 11. "span" - as only the ERA-interim mass fluxes are plotted, "are similar to" may be a better comparison.

OK, we will change "span" to "are similar to".

Pg 22967, line 10. Remove the second instance of "that"

OK.

Pg 22967, line 14. The sharp fall off of runs N_1991 and P_det should be mentioned, as it's not so different to that of the other runs, it just occurs at 200 instead of 300 hPa.

We will emphasise this by adding "While the sharp fall in the modelled convection from runs "N_1991" and "P_det" occurs at/above 200 hPa.

Pg 22967, line 25. does -> do

OK.

Pg 22969, line 14. I don't understand the sentence beginning "In the tropics.." Which run underestimated precipitation in the extra tropics? Both B_EI and A_E40 look to me like they have higher extra tropical values than GPI and CMAP. On the other hand if you meant A_E40 gives lower values than ERA-40 then I wouldn't imply that the ERA-40 data is correct by calling the model results an underestimation.

Sorry for the confusion - that sentence was not clear. Both runs underestimate precipitation in the extratropics (midlatitudes). This can be clearly seen also in Figure 10 which extends the higher latitudes than Figure 9. We will rephrase this.

Line 21-24. Several runs are said to overestimate the observations. This is only true in parts of the tropics. North or south of 25-30^o, all the runs dramatically underestimate the observations. This needs to be commented on and explained.

Sorry for confusion again. We will add "in the tropics" after "slightly overestimate the observations" in line 21 and "the observed precipitation rates" in line 23.

Pg 22970, line 6. What are the reasons for this overestimation?

TOMCAT gives a better simulation of radon-222 for the oceanic sites (e.g., Amsterdam Island and Bermuda) since these sites are mainly affected by large-scale transport (see Zhang et al., 2008). However, the largest discrepancy occurs at Hohenpeissenberg.
Zhang et al. (2008) pointed out that "Hohenpeissenberg in Germany is a challenging site for GCMs to simulate because of the orography.". This will also be the case for CTMs like TOMCAT.
There are various possible reasons for this overestimation. Zhang et al. (2008) mentioned two main reasons from their GCM simulations: 1) The observed surface Rn-222 strongly depends on the boundary layer. 2) The horizontal resolution in their GCM is coarse (~300 km) which is similar to our experiments Run A_E40 and B_EI.
Another possible reason is that the Rn-222 flux we used in the model from Jacob et al. (1997) may overestimate the local emissions. For example, *Conen and Robertson* (2002) reported that the direct of ^{222}Rn flux measurement in Geremany is 0.75-0.88 atom cm^{-2}s^{-1}, while *Zegvary et al.*, 2009) also reported even lower radon flux in European.
Conen, F. and L.B. Robertson, Latitudinal distribution of radon-222 flux from continents, *Tellus B*, *54*, 127-133, 2002.
Zegvary, T., F. Conen, and P. Ciais, European ^{222}Rn inventory for applied atmospheric studies, *Atmos. Eviron.*, *43*, 1536-1539, 2009.
We will add these points in the revised paper.

Pg 22970. Please provide references for the radon data you use in Fig 11/12. Information on the measurement technique and other background data would also be useful for all of the measurements discussed. Figure 12, panels C and D are not discussed. Discussion of Figure 12: how many profiles are included in the averages for the measurements? How do the model results compare to the different profiles? Are there some regions where the models do not match the observations, or do the errors average out to produce the reasonable comparison?

The radon data we used in Figures 11 and 12 are the same as Zhang et al. (2008). We will add "*Zhang et al., 2008*" and other references in Section 3.4.
Zhang et al. (2008) give some detailed information on the measurement technique and background data in their paper. Therefore, we will add "For more details about the Rn-222 data see *Zhang et al.* (2008)." in Section 3.4.
Liu et al. (1984) mentioned that their winter/summer Rn-222 observations were calculated from individual aircraft measurements at different continental locations from 1950 to 1972 (7 profiles for the winter and 23 profiles for the summer). The climatological

mid-latitude (30-60N) profiles have been widely used for the evaluation of the tracer transport in global models (e.g., Stevenson et al., 1998; Zhang et al.,2008). We will add this information in the revised paper.
For the model results, based on the land-sea mask information, we averaged the radon output between 30-60N among the land regions. We will mention this.
We apologise for not discussing panels C and D in Figure 12. These panels emphasise the effect of convection in summer and winter. We will add "The convection is very significant in summer but small in winter for the northern hemisphere mid-latitudes as expected." in the revised paper.

Pg 22970. It should be mentioned somewhere that vertical transport is not the only factor that affects the radon profile - horizontal mixing, particularly near coastal or snow covered areas may play a role, as there is a strong spatial gradient in radon emissions here.

OK. We will add this in the revised paper.

Pg 22970, line 21. You should define NARE somewhere.

OK, we will explain NARE (North Atlantic Regional Experiment) in the revised paper.

Pg 22970, discussion of Figure 13. There were huge day to day variations, and I don't think the data constrains the model results. Do the models reproduce the variations?

We will check if the model is able to reproduce the day-to-day variations and comment in the revised paper.

Pg 22972, line 19. Another significant disagreement is the latitudinal distribution (at mid-high latitudes the modelled mass fluxes and precipitation are way too low).

Yes, there is also significant disagreement in the latitudinal distribution. Normally people will be more interested in the tropical/mid-latitude convection, but we will also mention this here.

Pg 22973, line 4. Probably better to write something like "convection parameters in the Tiedtke scheme"

OK, we will change this.

L7 - Do you mean "Moreover it is not clear IF the changes....."?

Yes, we will add "if" here.

Figures: For most figures the axis and legend labels and numbering are way too small. Please also consider increasing the line thickness, in line plots as it is hard to tell the difference between the line colours.

We will replot the figures to make them more readable.

Reply to *Anonymous Referee #2*

This paper evaluates the simulation quality of a convective tracer, Radon-222, corresponding to various source and treatment of meteorological data to force convention within the TOMCAT transport model framework. The simulation results are compared with observations available at a few locations. Detailed comparisons are made for some diagnostic parameters available from ECMWF analyses as well as remote sensing products. The results can overall serve as useful guidance for the chemistry transport model developers, and could be published in Atmos. Chem. Phys. after accounting for the concerns listed below.

We thank the reviewer's positive feedback for the paper. We will revise the paper based on the reviewer's comments and very useful suggestions to improve the quality of the paper.

I understand the authors have made great effort for testing so many of the meteorological data products for convection parameterization schemes in TOMCAT. However, less effort is probably made to reveal the highlight of their results. For example, the lines 19-23 of the abstract suggest increasing resolution do not change much in TOMCAT simulations. While Fig. 1,2,5 have so many panels hiding the essential differences between the model versions. I suggest the authors to choose a fewer panels where differences are clearer and state the results from other simulations in the Fig. caption or text briefly. I am also a bit surprised why Radon simulation was not done for 1×1 degree model run, particularly for the J_T106 case.

This paper is to investigate the performance of cloud convection and tracer transport based on different simulations and sensitivity experiments from TOMCAT. The configuration of Run K_L31 is generally used as the standard full chemistry TOMCAT version. We will highlights our results in the revised paper. We are currently running the radon simulations for Runs G_5.6 and J_T106 cases and will include them in the revised paper.

The problem with the rain rate (Section 4.3) is that do we really expect simulate the rain rate in an offline model correctly? I suspect this because the model do not account evaporation/evapotranspiration, surface roughness etc. processes realistically, which has been also shown in your paper.

This is true and the rain rate calculation in the offline model is crude. We include this because it is a meteorological parameter generated by the convection scheme which can be compared to observations. This will at least give more information that the convection is occurring in the correct locations. Moreover, the wet deposition is also a key process for some trace gases and aerosols in the troposphere. Therefore, the discussion about the rain rate will provide some useful information especially for the CTM users who are using the different ECMWF analyses.

Section 4.4: I am a bit skeptic of such comparisons. How are the model results samples for Fig. 12. Unless the model results are sampled at the same time and location of measurements and then averaged, we cannot really assess the model simulation quality. The differences seen in Fig. 12 can easily be due to model sampling error, if we take some lessons from Fig. 11 & 13.

There is limited radon-222 vertical profile data from measurements. The vertical data by Liu et al. (1984) is an averaged profile at different locations and time (i.e., only 7 profiles for the winter and 23 profiles for the summer). The data has been widely used

to compare with GCMs tracer transport in global models.

For the model results, based on the land-sea mask information, we averaged the radon output between 30-60^{o}N among the land regions. The model output is every six hours based on the available ECMWF (re)analyses. We will sample the model output at the same location as measurements by interpolation (please note the model has a coarse horizontal resolution) then average and check if there are any significant difference here.

Past research have shown that model resolution do play significant role in simulation quality under the GCM or GCM-like frameworks [Wild et al. ACP, 2008; Patra et al., GBC, 2008; Rind et al., JGR, 2007]. Again I felt, a comparison of radon simulations using G_5.6 and J_T106 cases would have been interesting.

Thank you for providing these references. We will add them in the revised paper and discuss this issue. Note that GCMs likely show more sensitivity to resolution than a CTM. For a GCM the resolution will affect the resolved transport processes and change the model dynamics. This will feed into tracer transport etc. For an off-line CTM the large-scale meteorology is fixed whatever resolution the model is run at. We are now running the radon simulations for Runs G_5.6 and J_T106 and will include them in the revised paper. We will discuss these issues based on our results in the revised paper.

Reply to *Anonymous Referee #3*

The objective of this paper is to estimate the sensitivity of simulated cloud convection in the TOMCAT model to the model grid resolution, surface evaporation fluxes (strength and resolution), type and resolution of large-scale forcing (from ECMWF). The results are presented in a pedagogical order. The figures are clear. The paper, however, would gain a more deeper analysis of the results and detailed explanations on specific points (see below). These clarifications and analysis are essential to give the paper a broader scientific impact.

We will revise the paper based on the reviewer's comments and useful suggestions to improve the quality of the paper. We will endeavour to provide a deeper analysis of the results and detailed explanations in the revised paper.

Surface evaporation fluxes appear to be a key variable in this study. It is not clear in the present version of the paper how the high resolution evaporation fluxes (1×1) are mapped in the lower resolution model grid (2.8×2.8). Are the fluxes simply averaged over the coarser grid? How do these new fluxes compare with the default UGCM fluxes in terms of time evolution, location of maxima, etc ... ?

Data from the high resolution grid is mapped (averaged) onto the lower resolution grid in a conservation way - there is no interpolation. The code determines which high resolution cells contribute (fully or partly) to each low resolution cell and a sums over these area-weighted values. We will make this clear in the revised paper.
The origin of the 1×1 surface evaporation dataset is the same as the UGCM. The distribution is therefore similar to the default UGCM fluxes in terms of time evolution, location of maxima, etc. There are, however, still slight differences due to the interpolation from low to high horizontal resolution we used. We will make this clearer and emphasis the sensitivity of the convective mass fluxes to this in the lower levels.

What is the methodology of the (re)construction of convective mass fluxes or winds when the meteorological analysis have higher resolution than the TOMCAT model ? How the chosen methodology impact on the differences between the TOMCAT model and ERA40 or ERA-Interim analysis ?

The model winds are read in as spectral coefficients of vorticity and divergence. These are then averaged onto whatever model grid is being used as part of the spectral transform. If the forcing winds are higher resolution than the model grid then information from the higher wavenumbers is not use - the spectral coefficients are truncated.
Convective massfluxes are read in on a 1×1 grid. These are averaged onto the model grid employed (i.e. this conserves the total mass transport in the analyses - there is no interpolation.
We will give more details in the revised paper.

In section 4.3, why does the model underestimate the convective mass fluxes and the precipitation at latitudes higher than 30º? Could it be due to the boundary layer moisture convergence criteria used by the Tiedtke scheme?

As mentioned in ACPD, the model does not include midlevel convection and convective downdrafts and there is no organised entrainment of environmental air above cloud base. That is one main reason to explain the underestimation of mid-latitude convective mass fluxes in TOMCAT. As shown in Figures 1 and 2, the diagnosed upward mass flux

improves the mid-latitude convective mass fluxes when we used updated Tiedkte scheme (Run_EInewconv). However, the model still underestimates strong updraft to 100 hPa since the ECMWF analyses are too moist statically stable in the offline model (e.g., we mentioned that "This will offset the problem in off-line models of diagnosing convection with analyses that have already been convectively adjusted".

The model still underestimates the precipitation at latitudes higher than 30^o though there is some improvement at region of 35-40^o from Run_EInewconv (please see Figure R1 (below).
The criteria used by the Tiedkte scheme to determine the convective cloud occurrence is as follow: A near-surface air parcel lifted upward must become supersaturated at higher level, this air parcel must also be buoyant at the level where supersaturation is reached and there must be moisture convergence below this level. The boundary layer moisture convergence below cloud base is one factor. Furthermore, the rainfall is parameterized as occurring above a cloud depth of 1500 m, which means the model does not consider precipitation due to shallow cumulus convection. We will mention this in the revised paper.

Could the authors explain how the convective precipitation are calculated in the model?

The calculation of the convective precipitation (CP) in the TOMCAT model is crude. The CP in TOMCAT occurs above a cloud depth of 1500 m. The rainfall rate in each model level above the 1.5 km is calculated based on the following equation:

$$CP = C \times LWC \times \varrho \times dz$$

if there is sufficient updraft mass flux to maintain this, otherwise it is calculated as:

$$CP = LWC \times Mu$$

Here C is the constant for precipitation parameterization which uses the same value (2.0 $\times$ 10^{-3} s^{-1}) as Tiedkte (1989). LWC is the liquid water content, ϱ is the Environmental air density, dz is the vertical depth of the model level while Mu is the convective updraft mass flux. We will mention this in the revised paper.

In section 4.4, it would be worth adding a transport equation for the radon in the model, plus an explanation of the convective transport scheme applied to the tracer. The reader might not be fully aware of how the model works.

In the model tracers are transported by large-scale advection, convection and boundary layer mixing. We will mention this.
In TOMCAT, we use *Holtslag and Boville (1993)* vertical diffusion scheme for the boundary layer mixing and a cumulus convection scheme (e.g., Tiedkte, 1989) or using the archived mass fluxes for tracer vertical transport.

$$\frac{\partial M}{\partial z} = E - D$$

$$\frac{\partial Rn}{\partial z} = E\frac{S0e}{SMe} - D\frac{Rn}{M}$$

where M is the airmass convective flux, E is the rates of entrainment where the air of environment flows to the cloud. D is the rates of detrainment where air of clouds flows into the environment. z is the model altitude. Rn is radon-222 tracer. SM and SOe are the mass of air and radon-222 tracer in the model grid box. Subscript e denote air in the environment.
In TOMCAT, the redistribution of the radon tracer is performed by the calculation of a matrix for one vertical column based on given entrainment, detrainment and vertical diffusion coefficients matrix which relates the tracer masses after subgrid scale vertical transport to their initial values. We will add this in the revised paper.

In Figure 12, all the versions of the model underestimate the observations around 8 km in summer. Is the convective scheme systematically underestimating the detrainment in this layer ?

Yes, the model underestimate the radon profile around 8 km in summer. As noted above, the model results here are averaged between 30-60N among the land regions based on the land-sea mask information. We will sample the model output at the same location of measurements by doing some interpolation (please note the model has a coarse horizontal resolution) then averaged and check if any significant difference here.

The authors conclude that the model is not sensitive to the model grid horizontal or vertical resolution. This conclusion is to some extent in contradiction with previous works and would deserved a dedicated discussion. Why is the model insensitive to changes in resolutions? Is there a buffering effect in the model or convective parameterization which lessen this impact? For example, the authors may want to discuss their results in the light of the following references:

The impact of horizontal resolution on moist processes in the ECMWF model, Phillips JT, Corsetti LC, Grotch SL, Climate Dynamics, 11 (2): 85-102, 1995.

Effects of resolution and model physics on tracer transports in the NASA Goddard Institute for Space Studies general circulation models, Rind D, Lerner J, Jonas J, et al., JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 112, D9, D09315, 2007.

The processes governing horizontal resolution sensitivity in a climate model, Pope VD, Stratton RA, CLIMATE DYNAMICS, 19, 3-4, 211-236, 2002.

Impact of horizontal resolution on seasonal integrations, Brankovic T, Gregory D., CLIMATE DYNAMICS, 18, 1-2, 123-143, 2001.

Sensitivity of equatorial convection to horizontal resolution in aquaplanet simulations with a variable-resolution GCM, Lorant V, Royer JF, MONTHLY WEATHER REVIEW, 129, 1; 2730-2745, 2001.

Thank you for providing the further references regarding the issues of model resolution from previous studies. We will add them and have a detailed discussion in the revised paper (see also response to Reviewer 2 above).

Actually, there are some differences when using the different model horizontal/vertical resolution. We mentioned that the impact of large changes in resolution are small and do not really improve on the most significant discrepancies with the archived mass fluxes in the tropical upper troposphere and at high latitudes. As seen in Figures 1, 2 and 5, the archived mass fluxes show strong convection transport up to 100 hPa, but none of the model results using the convection scheme are able to capture this though the model run using the old operational analyses (Run_N1991) has higher convective updraft mass

fluxes. GCMs using different cumulus convection schemes are not able to reproduce this well (e.g. Tost et al. (2010) in http://www.atmos-chem-phys.org/10/1931/2010/acp-10-1931-2010-supplement.pdf). This is a challenge for CTMs/GCMs.

We are now re-running the experiment which will include Radon simulations for Runs G_5.6 and J_T106 cases and show the Radon-222 distribution to check if the resolution has impact on tracer transport. We will include some results in the revised paper.

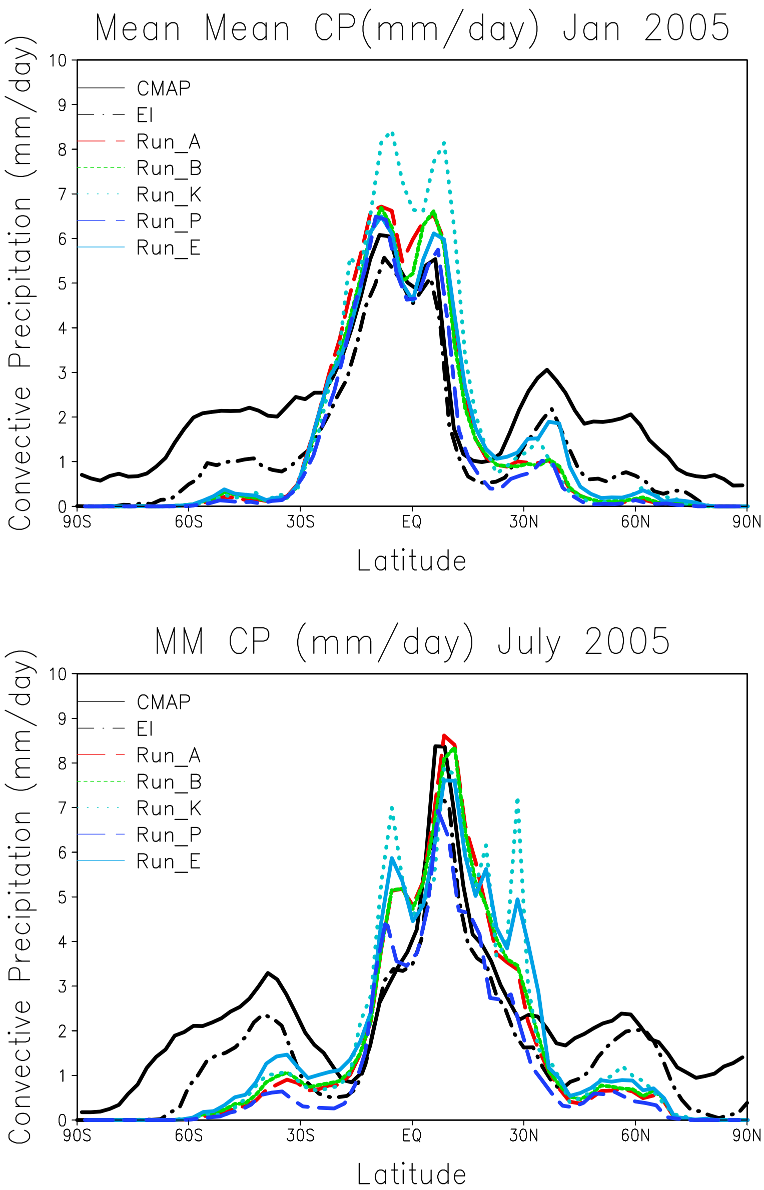

Figure 1: Figure R1. Zonal mean convective precipitation (mm/day) from CMAP data, ERA-Interim reanalyses and model runs A_E40, B_EI, K_L31, P_det and E_EInewconv for (a) January 2005, and (b) July 2005.