

An evaluation of the performance of chemistry transport models by comparison with research aircraft observations. Part 1: Concepts and overall model performance

D. Brunner¹, J. Staehelin¹, H. L. Rogers², M. O. Köhler², J. A. Pyle²,
D. Hauglustaine³, L. Jourdain⁴, T. K. Berntsen⁵, M. Gauss⁵, I. S. A. Isaksen⁵,
E. Meijer⁶, P. van Velthoven⁶, G. Pitari⁷, E. Mancini⁷, V. Grewe⁸, and R. Sausen⁸

¹Institute for Atmospheric and Climate Science, ETH Zürich, Switzerland

²Centre for Atmospheric Science, Cambridge University, UK

³Laboratoire des Sciences du Climat et de L'Environnement, Gif-sur-Yvette, France

⁴Service d'Aéronomie, Paris, France

⁵Department of Geophysics, University of Oslo, Norway

⁶Section of Atmospheric Composition, Royal Netherlands Meteorological Institute, The Netherlands

⁷Dipartimento di Fisica, Università L'Aquila, Italy

⁸Institut für Physik der Atmosphäre, DLR, Germany

Received: 11 February 2003 – Accepted: 23 April 2003 – Published: 15 May 2003

Correspondence to: D. Brunner (dominik.brunner@iac.umnw.ethz.ch)

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

Abstract

A rigorous evaluation of five global Chemistry-Transport and two Chemistry-Climate Models operated by several different groups in Europe was performed by comparing the models with trace gas observations from a number of research aircraft measurement campaigns. Whenever possible the models were run over the four-year period 1995–1998 and at each simulation time step the instantaneous tracer fields were interpolated to all coinciding observation points. This approach allows for a very close comparison with observations and fully accounts for the specific meteorological conditions during the measurement flights, which is important considering the often limited availability and representativity of such trace gas measurements. A new extensive database including all major research aircraft and commercial airliner measurements between 1995 and 1998 as well as ozone soundings was established specifically to support this type of direct comparison. Quantitative methods were applied to judge model performance including the calculation of average concentration biases and the visualization of correlations and RMS errors in the form of so-called Taylor diagrams. We present the general concepts applied, the structure and content of the database, and an overall analysis of model skills over four distinct regions. These regions were selected to represent various degrees and types of pollution and to cover large geographical domains with sufficient availability of observations. Comparison of model results with the observations revealed specific problems for each individual model. This study suggests what further improvements are needed and can serve as a benchmark for re-evaluations of such improvements. In general all models show deficiencies with respect to both mean concentrations and vertical gradients of the important trace gases ozone, CO and NO_x at the tropopause. Too strong two-way mixing across the tropopause is suggested to be the main reason for differences between simulated and observed CO and ozone values. The generally poor correlations between simulated and measured NO_x values suggest that in particular the NO_x input by lightning and the convective transport from the polluted boundary layer are still not well described by

CTM model evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

current parameterizations, which may lead to significant differences in the spatial and seasonal distribution of NO_x in the models. Simulated OH concentrations, on the other hand, were found to be in surprisingly good agreement with measured values.

1. Introduction

5 Global chemistry transport models (CTMs) and chemistry general circulation models (C-GCMs) are today becoming standard tools for estimating the contribution of individual pollutant sources to the distribution of trace gases on continental and global scales. One of the most relevant applications is the description of changes in radiative forcing due to changes in ozone caused by anthropogenic activities. In this context
10 the upper troposphere/lower stratosphere region (UT/LS) is particularly important because of the sensitivity of radiative forcing to the vertical distribution of ozone which has a maximum around the tropopause (Lacis et al., 1990). A large fraction of aircraft emissions is deposited at these sensitive altitudes and therefore the impact of emissions, in particular of nitrogen oxides, by the current and future air traffic was studied extensively in recent years and reviewed in several assessment reports (Penner et al.,
15 1999; Brasseur et al., 1998; NASA, 1999). CTMs and C-GCMS playing an important role in these assessments often showed significantly differing results both in terms of background concentrations of relevant species such as NO_x and in terms of perturbations caused by aircraft emissions. These differences underline the fact that the adequate simulation of the various chemical and dynamical processes in the upper
20 troposphere/lower stratosphere (UT/LS) region is a particularly difficult task. A high vertical model resolution is needed to adequately represent the steep concentration gradients across the tropopause and to describe mixing between the stratosphere and the troposphere. Convective processes, which are difficult to simulate, strongly affect
25 the photochemistry of the UT region by rapidly lifting upwards pollutants emitted at the surface (Berntsen and Isaksen, 1999). In addition, lightning activity in deep convective clouds is an important source of nitrogen oxides to this region but its overall strength is

CTM model evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I◀

▶I

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

still poorly quantified (Hauglustaine et al., 2001; Jourdain et al., 2001; Stockwell et al., 1999; Bernsten and Isaksen, 1999). Primary production of OH radicals in the UT/LS region depends not only on O₃ and H₂O levels but also on carbonyls and peroxides whose concentrations are again strongly linked to vertical transport and mixing (Jaeglé et al., 2001; Brühl et al., 2000; Prather and Jacob, 1997). Finally, due to the relatively long lifetimes and steep vertical gradients of many compounds in the UT/LS region small inaccuracies in the advection scheme may significantly affect their concentration levels as pointed out by Bregman et al. (2001). The downward flux of ozone from the stratosphere to the troposphere is particularly sensitive to the formulation of transport and to upper (stratospheric) boundary conditions, and recent estimates of the global annual mean flux obtained in a number of model studies varied by at least a factor of three (Houghton et al., 2001). All these issues add up to a substantial uncertainty in ozone budget estimates in the UT/LS region and careful analysis of model results by comparison with observations is therefore essential.

In the framework of the EU project TRADEOFF (Aircraft emissions: Contributions of various climate compounds to changes in composition and radiative forcing – tradeoff to reduce atmospheric impact) an extensive model evaluation study was undertaken involving five state-of-the-art CTMs and two C-GCMs with particular emphasis on the UT/LS region. The project TRADEOFF aims at a better understanding of the impact of aircraft emissions on the state of the atmosphere and the climate through a sequence of complementary modelling studies. Estimates of the tradeoffs of flying at different altitudes or latitudes as well as projections of future changes until the year 2050 are essential elements of the project.

The models involved in TRADEOFF contributed to numerous previous studies on the impact of aircraft emissions (Penner et al., 1999; Brasseur et al., 1998; Grewe et al., 2001; Meijer et al., 2000; Rogers et al., 2000; Pitari and Mancini, 2001) and participated in previous evaluation studies (Bregman et al., 2001; Law et al., 2000; Rogers et al., 2000; Penner et al., 1999; Brasseur et al., 1998; Houghton et al., 2001). Recently, Law et al. (2000) compared monthly mean ozone fields of five CTMs with data from the

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

MOZAIC program, in which ozone is measured continuously from passenger aircraft (Marengo et al., 1998). Another model intercomparison using data from MOZAIC and ozone soundings evaluating seasonal mean profiles in the lower stratosphere was presented by Bregman et al. (2001). Grewe et al. (2001) used the extensive data set of NO_x observations of the NOXAR project obtained from a Swissair B-747 airliner over the course of more than one year to evaluate the NO_x distributions calculated by two GCMs (E39/C and ULAQ GCM) and to investigate the importance of different sources. Further studies by Wang et al. (1998), Levy II et al. (1999), Emmons et al. (2000), and Bey et al. (2001) investigated the overall ability of CTMs to simulate tropospheric photochemistry, which is also the main focus of the present study. In these latter studies the models were compared with measurements of many different trace gases related to ozone photochemistry which were obtained from a number of surface stations and aircraft measurement campaigns.

The "classical" approach for evaluating a model followed by the above studies was to aggregate the observations over specific domains, altitude ranges, and time periods, and then to compare statistical quantities such as mean or median values and standard deviations for these aggregates with corresponding model data. However, the model fields were usually not sampled at exactly the same times and positions as the measurements but rather averages over entire time periods (e.g. monthly means) and domains (e.g. over a range of grid cells) were calculated because such quantities can easily be derived from standard model output. Furthermore, model results and observations were often taken from different years assuming that interannual variations were not strongly affecting the main results. Here we adopt a much more direct approach by comparing each measured data point with its temporally and spatially interpolated model counterpart. However, such a "point-by-point" analysis requires to simulate exactly the same time periods when the measurements were obtained which is not easily feasible with climate models or with CTMs that take their meteorological input from climate simulations. For two TRADEOFF models, namely ECHAM4.L39(DLR)/CHEM (subsequently named E39/C) and ULAQ, a direct comparison was therefore not ap-

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

plicable and hence some analysis based on the "classical" approach was included. Considering the often poor coverage and representativity of in-situ trace gas observations in the UT/LS region, the point-by-point strategy offers clear advantages as it fully accounts for the specific meteorological situation during the measurements, and hence for the specific transport and photochemical histories of the air masses encountered. Furthermore, no averaging or resampling is required which could reduce the quality of agreement between models and observations. Finally, the method largely simplifies quantitative analysis of model performance since two data sets of equal size (observations versus model output) are generated that can easily be compared applying statistical methods. More details on the implementation of this method are given in Sect. 2.1 and 2.2.

A new extensive observation database was established specifically for this study, which is described in Sect. 2.2. This was necessary because previous collections as the one by Emmons et al. (2000) did not fully support our preferred point-to-point strategy which requires to use the original data files rather than gridded composites. A huge amount of model output and analysis products was generated in the course of this evaluation exercise, most of which were made accessible to the individual modelling groups through a dedicated web site (see <http://www.iac.ethz.ch/tradeoff/database>). Most of the observation data sets used in this study are publicly accessible through that web site. Here we can only present examples of the methods applied and highlight the main results and overall tendencies. We will focus on measurement campaigns using research aircraft which typically provide information on many different species but which have only a limited coverage in both time and space. A comparison with climatologies derived from commercial airliner measurements and ozone soundings will be presented in a complementary study by Köhler et al. (2003), which will also inter-compare monthly mean trace gas distributions calculated by the models. Sect. 2.1 briefly describes the models and the experimental setup. Sections 2.2 and 2.3 provide an overview of the structure and content of the observation database and introduce the general concepts applied for evaluating the models. Finally, an analysis of the

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

overall model performance over four distinct regions is given in Sect. 3. In a second paper (Brunner et al., 2003) we will present a detailed comparison with two selected aircraft measurement campaigns, that is PEM-Tropics A and POLINAT/SONEX, evaluating both time-series and vertical profiles at various locations. That paper will also analyze in more detail how well the most relevant physical and chemical processes determining the distribution of different trace species are represented in the models.

2. Models, data and methods

2.1. Description of the models and simulations

Table 1 presents an overview of the key features of the models involved emphasizing their differences in terms of transport, chemistry, and model resolution. A description of the individual models will be given in the companion paper by Köhler et al. (2003). Five chemistry-transport (CTMs) and two chemistry general circulation models (C-GCMs) were used in TRADEOFF. The ULAQ CTM is driven by meteorological fields from the ULAQ climate-chemistry coupled model whereas the other CTMs are driven by European Center for Medium Range Weather Forecasts (ECMWF) analyses. The LMDz-INCA GCM was run in a special nudged-mode in which winds are relaxed to ECMWF analyses. Thus, except for E39/C and ULAQ the models were able to simulate the real weather conditions during individual measurement campaigns which is a prerequisite for using the point-by-point approach. The CTM2, TM3, TOMCAT and LMDz-INCA models only included tropospheric chemistry whereas both tropospheric and stratospheric chemistry were considered in CTM2-Gauss (a special version of CTM2 extending higher up into the stratosphere), E39/C, and ULAQ. Results of the SLIMCAT model, which is a stratospheric model formulated on isentropic surfaces, are not discussed here.

The following two types of model output were generated and analyzed with respect to observed trace gas distributions:

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

1) Point-by-point output (interpolated along-flight-path output) of the components O₃, CO, OH, HO₂, H₂O₂, H₂O, NO, NO₂, HNO₃, PAN, and Rn222. Models generating this type of output were run over the period 1995–1998 and at each simulation time step (typically of the order of 30 min) the instantaneous tracer fields were linearly interpolated to the positions of coinciding observations. The positions and times of these measurement points were provided to the modeling groups in separate tables (see Sect. 2.2 for further details).

2) Gridded monthly mean fields and standard deviations at 5° × 5° horizontal resolution and at 30 equally (1km) spaced vertical levels of the same trace gases as above, and in addition of net ozone production rates P(O₃), lightning NO emissions, and wet HNO₃ deposition. This type of output was only used if no point-by-point data was available. Another useful application of this output is to compare it with the point-by-point data which allows to analyze the representativity of the point-by-point data and hence of the observations for a given month.

Three out of the five models with meteorological input from ECMWF, namely TM3, TOMCAT, and LMDz-INCA performed simulations for the entire time period of 1995 to 1998. The selection of a four year period was a compromise between including as many measurements as possible while keeping computation time and costs within reasonable bounds. The setup of the CTM2 model only allowed to simulate 1996 with the tropospheric version, and only 1997 with the CTM2-Gauss version of the model, respectively. E39/C performed a multi-year simulation with year 1990 climate conditions (greenhouse gas concentrations and sea surface temperatures as of year 1990) and reported monthly output of four consecutive years. ULAQ reported monthly output fields from a single year based on a year 1990 climate simulation of the driving ULAQ-GCM. In order to facilitate interpretation of the results it was recommended to adapt emissions as listed in Table 2 which are based on the recent IPCC OxComp intercomparison exercise (Houghton et al., 2001) with some updates. These recommendations were followed to a great deal with the following exceptions: In TM3 the individual NMHC contributions were somewhat different but the annual total was as recommended. In

**CTM model
evaluation – Part 1**D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

LMDz-INCA the total NO_x emissions were somewhat lower mainly due to lower fossil and bio fuel emissions (27.8 instead of 31.8 Tg(N)/yr). CO emissions due to isoprene oxidation were somewhat higher as recommended (270 instead of 165 Tg/yr). Note that no NMHCs were simulated in this version of LMDz-INCA. In the E39/C simulations generally different emissions were used, representing rather year 1990 than 2000 conditions: Total NO_x emissions were lower (38.7 instead 50.1 Tg(N)/yr) and CO concentrations were fixed at the surface according to [Hein et al. \(2001\)](#). As in LMDz-INCA no NMHC were simulated. Each model applied its specific implementation of a lightning NO source parameterization but the values were scaled uniformly to a global total of 5 Tg (N)/yr.

2.2. A new database for model evaluation purposes

As part of this study a new database of in-situ measurements from both aircraft and ozonesondes focusing on the UT/LS region was established. Table 3 presents an overview of the contents. Several campaigns not available in the collection of [Emmons et al. \(2000\)](#) were included in this new data set, that is measurements of the STREAM, JAL, CARIBIC, ACE-2, ACSOE, STRAT and POLARIS campaigns. Also, the complete set of MOZAIC data of four consecutive years (1995–1998) was included. An important fraction is made up by the commercial aircraft measurement programs NOXAR, JAL, CARIBIC, and in particular by MOZAIC. Ozonesonde measurements obtained from the World Ozone and UV Data Center (WOUDC) and the NADIR data center at the Norwegian Institute for Air Research (NILU) contribute another important fraction. Research aircraft measurement campaigns contribute only about 2% of all data records but for many compounds this is the only reliable source of information currently available in the UT/LS region.

Figure 1 shows the coverage of commercial aircraft measurements and ozone soundings (a and b). However, in this study we only analyze measurements from research aircraft campaigns conducted between 1995 and 1998 (Figs. 1c–f). CO was measured on nearly every research aircraft mission whereas for other species the

availability is often strongly reduced. With respect to research aircraft observations good coverage exists only for a small proportion of the globe whereas other important areas such as the African and Asian continents or the South Atlantic were virtually unexplored during the selected period. Nevertheless, the data set allows evaluating the models over several distinct regions under strongly differing conditions in terms of meteorology and pollutant sources (cf. Fig. 5). Figures 1e and 1f show that in the northern hemisphere the bulk of data was obtained in the UT/LS region between about 9 and 12 km and in midlatitudes between 30° N and 60° N. Measurements from the high-altitude aircraft ER-2 during the STRAT and POLARIS campaigns extend well into the lower stratosphere up to about 21 km. Observations in the southern hemisphere are provided mainly by the ACE-1 and PEM-Tropics A campaigns which concentrated on the Pacific region.

To exemplify Fig. 2 describes the processing of the original PEM-Tropics A campaign data files for inclusion in the database. The original files were first processed for each campaign and platform separately by averaging or rearranging all measured parameters to a common timeline and by then merging these data into a single intermediate data file per flight (labeled M1 to Mn in the figure). 6-min averaging intervals were selected for aircraft measurements and 500 m altitude intervals for the soundings. The only exceptions are PEM-Tropics A and SONEX (5-min averages), and MOZAIC (12-min averages for cruise data and 450 m altitude intervals for profiles during take-off and landing). Depending on the speed of the aircraft a 6-min time interval corresponds to a horizontal flight distance of about 50–100 km type. This is still significantly smaller than the typical resolution of a global CTM which is of the order of 200–500 km. The individual flights were then combined into monthly campaign data tables which constitute the main part of the TRADEOFF database. Finally, the times and positions of observations from all campaigns were merged to obtain a chronologically ordered "TIMEPOS" table for each individual month, constituting the second type of tables in the database. These TIMEPOS tables were used during the model simulations to generate the output for the point-by-point comparison. Since multiple observations from different campaigns are

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)
[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

usually available for a given time, a link between the entries in the TIMEPOS tables and the monthly campaign data tables is accomplished by an identification number. This number is chosen to be unique for a given campaign, platform, and flight, and is preserved in the point-by-point output allowing to link the model results unambiguously with the corresponding observations.

2.3. Quantitative analysis of model performance using Taylor-diagrams

Recently, Taylor (2001) presented a new type of diagram that can concisely summarize the degree of correspondence between simulated and observed fields (see Sect. 3.3 for an application of "Taylor diagrams" in this study). On this diagram the correlation coefficient R and root-mean-square (RMS) error E' between a test field f (model) and a reference field r (observations), along with the ratio of the standard deviations (σ_f and σ_r) of the two patterns are all indicated by a single point in a two-dimensional plot. Figure 3 illustrates the geometric relationship between these quantities in the diagram (see Taylor (2001) for details). The pattern RMS E' is defined as

$$E' = \sqrt{\left\{ \frac{1}{N} \sum_{n=1}^N [(f_n - \bar{f}) - (r_n - \bar{r})]^2 \right\}} \quad (1)$$

and is thus the RMS difference between the deviations of the test and reference fields from their respective mean values \bar{f} and \bar{r} . In the diagram the correlation coefficient is simply the cosine of the angle between the x-axis and the test point F , and the RMS difference is the linear distance between the test point and the reference point Ref . Hence, the point of a well performing test field would appear near point Ref . In this paper we will always use standard deviations of the model field that are normalized by the observed standard deviation, denoted as $\hat{\sigma}_f$.

The correlation coefficient and the RMS difference provide complementary aspects of model performance. If, for instance, peaks of high NO concentrations caused by lightning were well represented in a model but the amplitudes of these signals were

**CTM model
evaluation – Part 1**

 D. Brunner et al.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Print Version](#)
[Interactive Discussion](#)

© EGU 2003

strongly underestimated, the model would exhibit a high correlation coefficient but at the same time a poor performance in terms of RMS error. On the other hand, for a given RMS value it is impossible to determine how much of the error is due to a difference in structure and phase and how much is simply due to errors in the amplitude of variations. Judging the overall skill of a model must therefore take into account both aspects. Taylor (2001) presented several formulations for an overall skill score. Here we adopt a formulation that puts more weight on a good correlation than on a small RMS error. A good correlation suggests that fundamental processes are adequately represented in a model. We subjectively judge this here to be more important than for instance an accurate representation of the strength of emission sources which would additionally improve the skill in terms of RMS difference. We define a skill score as

$$S = \frac{4(1 + R)^2}{(\hat{\sigma}_f + 1/\hat{\sigma}_f)^2 (1 + R_0)^2} \quad (2)$$

where R_0 is the maximum attainable correlation. The maximum attainable correlation is limited for instance by the fact that the model fields can not fully resolve all features of the 6-min averaged observation data. To estimate this influence we used the correlation between modelled and measured temperature. Instrument noise further reduces the maximum attainable correlation. To estimate this effect we added an artificial Gaussian noise to the point-by-point output of a particular model according to the stated instrument precision. The correlation between the original and the noisy model output then provides an estimate of the influence of instrument noise on R_0 . It is important to note that the above definition of a skill score does not account for the overall bias between simulated and observed concentrations. Therefore, Taylor diagrams of model performance are presented in this study along with tables of the overall bias. In some situations the overall bias may be more indicative of a model's capabilities than a correlation coefficient. As an example, concentration time series over remote areas often resemble constants with some noisy pattern, and the correlation coefficient

may be low despite the fact the simulated concentrations lie within a few percent of the measured ones.

3. Evaluation of overall model performance in the UT/LS region

To demonstrate the capacity of the point-by-point approach Fig. 4 shows exemplarily time series of ozone and temperature along the path of 5 consecutive flights of the NASA DC-8 aircraft during the POLINAT/SONEX campaign in October/November 1997 (Singh et al., 1999; Thompson et al., 1999). Measured values are shown in black and interpolated model fields in colors. Differences between the model temperatures are very small since except for the LMDz-INCA GCM the temperature fields are taken from the driving ECMWF model. In the LMDz-INCA model only winds were relaxed to ECMWF analyses by the nudging procedure, but not temperature. LMDz-INCA has a cold bias of about 5°C at the tropopause in mid-latitudes which was corrected for in the figure. The comparison with temperature allows to get a rough estimate of the limitations that differing spatial resolutions between measurements and models impose on the maximum agreement that can be achieved. The excellent agreement indicates that for the selected 5-min averages along the flight track most of the observed variability can be resolved by the models.

To a large extent the variability in ozone is due to variations in tropopause altitudes resulting in alternating flight stretches in the upper troposphere (O₃ below 100 ppbv) and in the lowermost stratosphere (O₃ above 100 ppbv). Variations in ozone are not as well reproduced by the models as variations in temperature, indicating that other factors apart from model resolution such as numerical diffusion must play an important role. Occasionally, also the ECMWF analysis can severely be in error. As an example, towards the end of flight F9 the model temperatures differ from the observations by about 5 to 7°C and at the same time simulated ozone concentrations are substantially lower than observed. Hence, errors in the driving meteorological fields additionally put a limitation on the maximum achievable agreement. This limitation is likely to be most

CTM model evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

important for longlived species such as ozone and CO since their concentrations are determined by the history of air parcels over long time periods during which errors in the analyzed meteorological fields may accumulate. However, we can only speculate to what extent this may influence the quality of the simulated trace gas fields since so far there have been no detailed investigations to this problem.

The kind of data shown in Fig. 4 is used in the following sections that present an evaluation of the overall model performance over four distinct regions. Grouping the results by different geographical regions provides a test for the model's capabilities in simulating a range of trace species having different sources and lifetimes under a range of conditions differing in terms of meteorology and remoteness from pollutant sources. The coverage of research aircraft measurements in the four regions is displayed in Fig. 5 with flights from different campaigns shown in different colors. Sect. 3.1 presents a few selected scatter plots of measured versus modelled trace gases for the North America region, and Sects. 3.2 and 3.3 present an analysis of average concentration biases and overall model performance in all four regions, respectively. Finally, meridional distributions over the Pacific ocean where the measurements covered a broad range of latitudes from about 45° S to 45° N are evaluated in Sect. 3.4.

3.1. Scatter plots

Figures 6 and 7 are scatter plots between measured and simulated ozone and CO from all research aircraft measurements obtained over North America. Only the results of the LMDz-INCA, TM3, and TOMCAT models are fully comparable since output of the CTM2-Gauss and CTM2 models was only available for campaigns in 1997 (POLINAT/SONEX) and in 1996 (SUCCESS, TOTE/VOTE), respectively. The comparison with ozone shows high correlation coefficients of about 0.85 for the TM3 and TOMCAT models but they both overestimate the increase in ozone when changing from the UT into the LS by roughly a factor of two. Particularly in the range of observed concentrations between 70 to 150 ppbv the two models quite strongly overestimate ozone suggesting too strong downward mixing from the stratosphere. In contrast, lower strato-

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

spheric ozone is underestimated in the CTM2 and LMDz-INCA models. In both models the agreement is best for the TOTE/VOTE campaign which was carried out in winter. Similar to TM3 and TOMCAT the CTM2 model tends to overestimate ozone in the upper troposphere, possibly also due to too strong downward mixing. The vertical model resolution appears not to be the dominant factor in determining the ozone levels in the UT/LS region. The LMDz-INCA and TM3 models, for instance, have the same vertical resolution but quite a different behavior with respect to ozone. The results for TOMCAT, on the other hand, are similar to TM3 despite its substantially higher resolution (see Table 1). The CO data reveal similar problems. In all models the slope in the CO-CO correlation is too low meaning that the models tend to underestimate CO when the observed concentrations are high and vice versa. Only the CTM2-Gauss model agrees well with the observed trend but the results are limited to the POLINAT/SONEX data set which covers only a relatively small range of CO concentrations. Particularly at very low concentrations (< 50 ppbv) observed in the lowermost stratosphere the models quite strongly overestimate CO. Thus, the models tend not only to mix down too much ozone from the LS into the UT, but also too much tropospheric pollutants in the opposite direction. It remains inconclusive to what extent these problems are related simply to model resolution and how much is due to numerical diffusion and other factors. Since high CO concentrations are mainly observed at low altitudes (blue to green symbols in Fig. 7) and generally lower values aloft, a too flat CO:CO correlation as seen in most models could indicate too intense upward transport and mixing from the boundary layer. A sensitivity test with two different boundary layer schemes in the TM3 model indeed showed a reduction in upper tropospheric CO by a few percent when using a less diffusive scheme (Ernst Meijer, personal communication).

The SUCCESS measurements (represented by “+” symbols in the figures) exhibit some very high CO concentrations observed at about 10 km altitude. These high values are most likely due to rapid convective transport from the polluted boundary layer. The fact that none of the models was able to reproduce these values suggests that the convective transport during this particular event was not correctly simulated.

3.2. Model biases

Average model biases were calculated separately for each measurement campaign and grouped by the four regions (Fig. 8). The Pacific Ocean was further separated into a North Pacific and a South Pacific domain. Figure 8a) presents the results for the altitude range of 300–170 hPa (approx. 9–13 km) and Fig. 8b) for lower altitudes of 510–350 hPa (approx. 5.5–8 km), respectively. The upper range covers measurements in both the UT and the LS whereas the lower range basically represents free tropospheric air with little direct influence from the stratosphere. Some measurement campaigns only covered the lower or the upper levels but not both. Missing or insufficient data are represented by crosses in the figure. Model biases are expressed as $(\text{mean}_{\text{model}} - \text{mean}_{\text{obs}}) / \text{mean}_{\text{obs}}$ and are shown as colored boxes with blue colors indicating negative and yellow to red colors indicating positive model biases, respectively. Results were obtained from the point-by-point output whenever possible. Since the CTM2 and CTM2-GAUSS models only covered a single year, results for these models were derived additionally from gridded monthly output fields by selecting the same months and by averaging over the same geographical domains covered by the measurement campaigns. Results for the E39/C and ULAQ models were derived in the same way from monthly output fields.

The two figures reveal tendencies for individual models and often for entire groups of models. The TM3 model clearly simulates too high O₃ and CO concentrations at both levels and over all regions except for the lower altitude range over the South Pacific and Tasmania. The LMDz-INCA model tends to underestimate ozone in the 9–13 km altitude range but is generally in close agreement with observations at the lower levels. CO is consistently too low at both levels and over all regions. The TOMCAT model is also generally low in CO but it is closer to observed values over the Pacific ocean than at northern mid-latitudes. Remarkably, TOMCAT overestimates O₃ concentrations over North America in both altitude ranges but is close to or even below the observed values over the other regions. As seen in the previous section TOMCAT simulated

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

too high ozone in the lowermost stratosphere. Most probably the high values over North America are due to the fact that a higher fraction of the flights took place in the lowermost stratosphere than over the other regions.

As noted before, results for the two versions of the CTM2 model were derived both from the point-by-point data and from monthly output fields. Average biases calculated in these two different ways are broadly consistent suggesting that also the comparison with monthly mean fields provides valuable information on deviations between models and observations. This is important with respect to the evaluation of the E39/C and ULAQ models for which only monthly output data is available. The CTM2-Gauss model displays a general tendency towards a positive ozone bias over all regions and at both levels whereas deviations are typically small for CTM2, in particular when only considering the more accurate point-by-point data. Considering the monthly mean data the two models behave very similar with the "Gauss version" always showing somewhat higher ozone. Also with respect to CO the two model versions behave quite similar. Deviations from observations are rather small and most of the time within 30% of the measured means.

Deviations from the observed mean O₃ and CO concentrations are similar in the E39/C and ULAQ models. However, a notable difference is evident at the lower altitudes over the North Pacific ocean where E39/C quite significantly overestimates O₃ whereas ULAQ is in close agreement with the observations. CO values tend to be too low in both models over all regions and at both levels.

All models tend to overestimate ozone and underestimate CO at the upper altitude range over Tasmania when compared to PEM-Tropics A measurements. However, it should be noted that this is only a small data sample. Too low CO concentrations are also seen in the models when compared with ACE-1 measurements, which is a much more representative data set for this region.

NO_x is the sum of NO plus NO₂. Since only few NO₂ measurements are available and the quality of the NO₂ measurements is a subject of debate (Crawford et al., 1996), we have calculated "observed NO_x" by using only measured NO concentrations and

CTM model
evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

scaling these values by the NO_x to NO ratio predicted by the TM3 model. The advantage of using NO_x instead of NO alone is that NO_x concentrations are insensitive to the solar zenith angle (which only influences the partitioning between NO and NO_2) and therefore a comparison with monthly mean values is more easily possible. NO virtually vanishes at night due to titration by ozone and therefore nighttime measurements were excluded here.

The models tend to underestimate NO_x over North America at both altitude ranges and over the North Atlantic at the higher levels. Mainly elevated values are seen in the models at the lower levels over the North Atlantic when compared with POLINAT 2 and over the North Pacific when compared with PEM-Tropics A (DC-8) measurements. The TM3 model significantly overestimates NO_x at 9–13 km over the remote South Pacific. A detailed comparison with PEM-Tropics A measurements (Brunner et al., 2003) suggests a too strong lightning contribution in TM3 over this region. The LMDz-INCA model is often biased high at the lower altitudes (5.5–8 km) compared to both the measurements and the other models, in particular over the North Atlantic. This model also tends to be high in HNO_3 at those levels. Hence, excessive NO production through HNO_3 photolysis is a likely explanation for the elevated NO_x values. In a recently revised version of the LMDz-INCA model washout of HNO_3 has been increased which would likely bring the model in better agreement with the observations. As a positive side-effect, OH concentrations, which were also found to be high in the model, were also significantly reduced (D. Hauglustaine, personal communication). Several models strongly overestimate NO over the North Pacific Ocean when compared with TOTE/VOTE measurements. Since most of the TOTE/VOTE flights were performed at night during winter, only few samples could be used to calculate NO_x concentrations and hence the representativity of this data is quite poor. However, a similar tendency is seen in the comparison with the much more representative PEM-Tropics A measurements from the DC-8 aircraft in fall. HNO_3 is much higher in all models than observed over the South Pacific. The same discrepancy was reported for other models also and several hypotheses were formulated (Wang et al., 1998; Bey et al., 2001), including in-

CTM model
evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

sufficient wash-out of HNO_3 (Wang et al., 1998), missing heterogeneous conversion of HNO_3 to NO_x on sulfate aerosols (Chatfield, 1994) or on soot (Hauglustaine et al., 1996), overestimate of N_2O_5 hydrolysis which is suppressed if aerosols were mostly dry (Schultz et al., 2000), and removal of HNO_3 due to gravitational settling of cirrus ice crystals which is missing in the models (Lawrence and Crutzen, 1998). Our detailed comparison with PEM-Tropics A measurements (Brunner et al., 2003) also suggest too strong downward transport of HNO_3 from the lowermost stratosphere which is confirmed by similarly elevated O_3 concentrations in most models over this region.

Except for the TOMCAT model PAN is often substantially overestimated by the models. A clear exception is the PEM-Tropics A measurements over the South Pacific and Tasmania. As reported by Fuelberg et al. (1999) these measurements were strongly affected by biomass burning activity. Biomass burning is an important source of PAN in the atmosphere which appears to be underestimated by the models. The closest agreement with observations is generally achieved by the ULAQ model.

Results for hydrogen peroxide reveal remarkable differences between more polluted areas such as North America and the North Atlantic and more remote regions such as the Pacific Ocean and Tasmania. H_2O_2 tends to be underestimated by all models over remote regions whereas some models overestimate its concentrations over the more polluted areas. These differences are particularly evident for the models LMDz-INCA, TOMCAT, CTM2, and ULAQ. The high H_2O_2 concentrations in the LMDz-INCA model at northern midlatitudes can be explained by the excessive water vapor simulated by this GCM. With respect to the other models the mainly too low NO concentrations over North America and the North Atlantic may contribute to the elevated H_2O_2 over these regions. The reaction of peroxy radicals with NO is in competition with the production of H_2O_2 and hence too low NO leads to excessive H_2O_2 . However, differences between individual model results are inconclusive in that respect.

3.3. Taylor diagrams

For a quantitative analysis of model performance with respect to correlation and RMS error the two altitude ranges studied separately in the previous section were combined to a single altitude range of 510–170 hPa (approx. 5.5–13 km). The combined point-by-point model output from all campaigns available over a given region was evaluated in the form of Taylor diagrams (Fig. 9). No results can be presented here for the ULAQ and E39/C models and results for the CTM2 and CTM2-Gauss models are restricted to the years 1996 and 1997, respectively.

As expected, by far the highest scores are reached for temperature (dark blue labels) with the nudged LMDz-INCA model being almost equivalent to TM3 and CTM2-Gauss. With respect to the trace gases the highest scores are usually obtained for ozone (purple). The ozone labels for TM3 and TOMCAT often appear at normalized standard deviations greater than one meaning that these models tend to overestimate the variability. However, they usually perform the best in terms of correlation resulting in a good overall skill score. CTM2 and LMDz-INCA tend to underestimate the variability in particular over regions where the stratospheric influence is strongest (as seen by high average ozone concentrations, not shown), resulting in a reduced overall skill. The same grouping of the models with respect to ozone is also apparent in Fig. 8a) showing that the TOMCAT and TM3 models tend to overestimate ozone in the UT/LS region whereas CTM2 and LMDz-INCA tend to slightly underestimate ozone there. All models exhibit a poor performance in terms of ozone over Tasmania.

With respect to CO (light blue) the models are grouped together closely, and, somewhat surprisingly, the skill scores are rather low over all regions. In particular over remote regions, time-series of CO often resemble constants with some noisy pattern. Therefore, the overall bias may sometimes be more indicative of the quality of the simulation than the skill score as used here. The highest scores are reached over the North Pacific and the lowest scores over the South Pacific. The unsatisfactory results for the South Pacific are most likely related to the representation of biomass burning

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

emissions in the models. Our detailed comparison with PEM-Tropics A measurements (Brunner et al., 2003) suggests that the climatological distribution of biomass burning emissions used for these simulations is not representative for the true emissions that occurred over southern Africa, South America and Australia during fall 1996. In agreement with the results for CO the variability in PAN, another important tracer of biomass burning activity, is strongly underestimated in the models over the South Pacific.

Model performance with respect to NO is often low which points out the fact that NO is one of the most difficult species to simulate in the UT/LS region. All models tend to significantly underestimate NO variability over North America and the North Atlantic region whereas TM3 quite strongly overestimates the variability over the North and South Pacific ocean. The TM3 model also showed a significant positive bias in NO over those regions (see Fig. 8a) suggesting a too strong source of nitrogen oxides over the remote South Pacific, probably due to lightning. The TM3 model has a unique lightning parameterization which is coupled to the intensity of convective precipitation predicted by the ECMWF model (Meijer et al., 2001), which may overestimate lightning activity over the oceans (Ernst Meijer, personal communication). Over the North Pacific also LMDz-INCA and CTM2 overstate the variability in NO along with a positive average bias. The models generally perform better with respect to HNO₃ as compared to NO over North America and the North Atlantic but they perform worse with respect to HNO₃ over the more remote North and South Pacific. This may reflect problems in simulating correctly the wash-out of HNO₃ which is likely a major factor in determining HNO₃ variability over remote regions.

In terms of H₂O₂ (red) the models perform much better over the North Atlantic than over North America. Convective activity is probably an important source of variability in H₂O₂ and OH concentrations in the upper troposphere over the continents (Prather and Jacob, 1997; Jaeglé et al., 2001). The reduced performance over North America may reflect difficulties in simulating correctly the strength, positioning and timing of convection. In agreement with this, also OH shows a strongly reduced performance over North America.

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

The TOMCAT model quite strongly underestimates the concentrations of PAN (not shown) and consequently its variability, resulting in a low skill score. Results of the other models for PAN are not conclusive since differences in skill scores are very large between the different regions. Obviously, the adequate simulation of PAN is a demanding task for CTMs.

3.4. Meridional distributions over the Pacific Ocean

The measurement campaigns ACE-1, PEM-TropicsA and TOTE/VOTE covered a broad range of latitudes over the Pacific Ocean from approximately 45° S to 45° N allowing to study meridional trace gas profiles over this region as shown in Fig. 10. Since we focus on the UT/LS region our analysis is restricted to the 300–170 hPa range (approx. 9–13 km). In the upper panels (Fig. 10a) the latitudinal model distributions were derived from point-by-point output. Results of the CTM2 model were only comparable with measurements in 1996, that is with PEM-TropicsA and a fraction of the TOTE/VOTE campaign. In the lower panels (Fig. 10b) the model distributions were derived from gridded monthly output fields by selecting the same months of the year covered by the observations and by selecting the three grid levels located within the 300–170 hPa range. Both types of profiles are shown for the TM3 model in order to emphasize any differences between them.

Significant differences are seen for instance at 35–45° N where the point-by-point data show a larger contribution of stratospheric air as expressed by a strong increase in O₃ and a drop in CO concentrations. This likely explains why the climatological distributions of CTM2-Gauss, E39/C and ULAQ do not follow the observed drop in CO at these latitudes. Significant differences also exist near the lowest latitudes between 30° S and 40° S where only very few measurements are available. With respect to NO_x the two distributions differ quite strongly over the North Pacific with significantly higher NO_x values in the "climatological" monthly mean distribution comparable to the concentrations over the South Pacific. The observed distribution is much better reproduced in the point-by-point output. This underlines the importance of sampling the model

data at exactly the same times and positions and hence under the same meteorological conditions, or to compare long-term means, which reflect the whole meteorological variability.

The shapes of both the O_3 and CO profiles are well reproduced by the models for which point-by-point output is available (Fig. 10a). Between 35–45° N where an important stratospheric contribution is apparent, O_3 is overestimated by TM3 and TOMCAT and underestimated by CTM2 and LMDz-INCA. This concurs with the results presented in the scatter plots of ozone over North America (Fig. 6) where TM3 and TOMCAT were found to overstate O_3 in the lowermost stratosphere. Although the measured increase in O_3 at these latitudes is well reproduced by the models, the corresponding drop in CO is clearly underestimated, in particular by TM3 and LMDz-INCA. Since these models were run at a lower vertical resolution than TOMCAT and CTM2 this may be an issue of vertical resolution.

The measured O_3 profile shows a flat minimum around the equator which in general is well represented by the models. In the reduced data set for CTM2 both the measured and modelled minimum are shifted northward by about 10°. The E39/C model simulates too high O_3 in the tropical regions (about 45 ppbv compared to 30 ppbv) whereas the ULAQ model simulates a too strong O_3 minimum here (a minimum of about 15 ppbv).

The measurements show significantly higher CO concentrations in the Northern as compared to the Southern Hemisphere (apart from the drop at 35–45° N). This difference is also seen in the model results but the amplitude is too small in TOMCAT, E39/C, and ULAQ, suggesting a too small source or a too large sink of CO in the Northern Hemisphere. The TM3 model generally overestimates CO whereas LMDz-INCA generally underestimates the concentrations. However, the shape of the meridional profile and also the increased variability in CO in the Northern Hemisphere is well reproduced by these models. The ULAQ model finds a significant increase in CO near the equator not seen in the observations. Since at the same time O_3 is too low this likely reflects excessively strong vertical mixing in the equatorial regions which would

**CTM model
evaluation – Part 1**D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[I◀](#)[▶I](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

transport upwards elevated CO from continental sources and low ozone depleted in the moist tropical (marine) boundary layer (Kley et al., 1996).

The TM3/TOMCAT model overestimates/underestimates NO_x over the Pacific in both the Southern and Northern Hemisphere, in agreement with the findings presented in Fig. 8a). However, apart from this bias the main observed features are well represented. This is also true for LMDz-INCA and CTM2 which are in good agreement with the observations also in absolute terms. However, the CTM2 model simulates strongly enhanced NO_x at high latitudes not seen in the measurements. This is most likely due to a missing sink in the model at high latitudes during winter.

With respect to NO_x the shapes of the meridional profiles of CTM2-Gauss, E39/C and ULAQ (Fig. 10b, bottom) differ quite significantly from each other, in particular in the tropics. However, the low representativity of the measured distributions precludes the possibility of judging which profile is more realistic than the other.

4. Conclusions

We have performed a rigorous and quantitative evaluation of the performance of a number of global CTMs and C-GCMs in the UT/LS region by mainly comparing observed concentrations with model fields sampled at exactly the same times and positions as the observations. This approach is different from evaluations performed in previous studies in which typically spatially and temporally averaged fields were compared. It offers important advantages as it fully accounts for the specific meteorological conditions during the measurements. This is particularly important when comparisons are made with spatially and temporally sparse data sets such as those typically obtained from research aircraft campaigns. A new extensive observation database with better support for this type of evaluation than provided by preexisting collections has been established. We have applied a quantitative method for judging complementary aspects of model performance such as correlation coefficients and root-mean-square (RMS) errors, and we have combined this analysis with information on average biases between

measurements and models. Results were analyzed separately for four different regions to test the models under varying conditions in terms of meteorology and remoteness from pollutant sources.

The main conclusions that can be drawn from for the individual models are as follows:

5 TM3: Ozone concentrations simulated by this model are well correlated with the observations. However, TM3 significantly overestimates ozone in the lowermost stratosphere. As a consequence of this and probably also of too strong mixing across the tropopause, upper tropospheric ozone in the extratropics is also enhanced. Excessive mixing across the tropopause is also evidenced by elevated CO concentrations in the
10 lowermost stratosphere. CO concentrations in the UT/LS region are generally biased high in the model irrespective of the geographical domain. One factor contributing to this is a deficit in OH in the UT when compared with observations. Additionally, too rapid venting of the boundary layer may transport too much CO to these altitudes. However, a sensitivity test with a less diffusive boundary layer scheme showed a
15 reduction in upper tropospheric CO by only a few percent. Average NO_x concentrations are in good agreement with observations over North America and the North Atlantic, but are significantly too high over the Pacific. The unique lightning parameterization of TM3 which links lightning NO production with convective precipitation may release too much NO over the tropical ocean. Along with this nitric acid is too high in the model
20 over the Pacific, a problem common to most CTMs in this study and also to other CTMs as reported by Wang et al. (1998) and Bey et al. (2001).

25 TOMCAT: Despite its higher vertical resolution TOMCAT behaves similar to TM3 with respect to ozone and CO at the tropopause, with too much O₃ and CO in the lowermost stratosphere. Different from TM3 the TOMCAT model is low in CO in the UT. This deficiency is more pronounced at northern mid-latitudes than in the southern hemisphere. An overly strong sink is unlikely since OH radical concentrations are in good agreement with observations at northern midlatitudes. A likely explanation is that TOMCAT tends to underestimate convective transport and occurring convection is biased towards the tropics (M. Köhler, personal communication). NO_x tends to be low in the

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[I◀](#)[▶I](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

TOMCAT model in the UT/LS region. In the TM3 and LMDz-INCA models the vertical distribution of the lightning NO source follows the recommendations of Pickering et al. (1998) which are based on simulations with a cloud-resolving model. This is not the case in TOMCAT which likely contributes to the somewhat lower concentrations. PAN is strongly underestimated over all regions suggesting missing direct sources or missing conversion from NO_x or both. As in all other models, hydrogen peroxide concentrations are underestimated over the North and South Pacific ocean. The reason for this discrepancy is not clear. Excessive wet removal is a possible explanation.

CTM2: In contrast to TM3 and TOMCAT, CTM2 underestimates ozone in the lowermost stratosphere. In the upper troposphere, however, ozone is often overestimated suggesting too strong downward mixing across the tropopause similar to TM3 and TOMCAT. Average CO concentrations over the different regions and also the meridional distribution over the Pacific ocean are in very good agreement with observations. The performance with respect to CO in terms of correlation coefficients and pattern RMS error, however, is similarly poor as in the other models. CO is a rather longlived species in the upper troposphere and hence time series of CO often resemble constants with some noisy pattern. The mean bias may therefore be more indicative of the quality of agreement with observations than the correlation coefficient. NO_x concentrations are typically 10–50% too low over North America and the North Atlantic. At high latitudes during late fall and winter, however, CTM2 strongly overestimates NO_x most probably due to missing or insufficient N₂O₅ hydrolysis on aerosols during nighttime.

CTM2-Gauss: In many respects the behavior of CTM2-Gauss (a version with stratospheric chemistry included and extending higher up in the atmosphere) resembles that of CTM2. Yet, since output from these two models was not available for the same year, a direct comparison is quite difficult. Average biases of CTM2-Gauss derived from monthly means fields versus observations show the same pattern of overestimation/underestimation as average biases of CTM2. However, O₃ is generally somewhat higher in CTM2-Gauss in the UT/LS and NO_x somewhat lower compared to CTM2. CTM2-Gauss appears to have a similar problem of excessive NO_x at high latitudes

during winter.

LMDz-INCA: LMDz-INCA is a GCM that was run in nudged mode in which winds were relaxed to ECMWF analyses. In this way it was possible to simulate the meteorology of the time period 1995 to 1998. In the extratropical UT/LS region the model is somewhat too cold and clearly too wet. Apart from this cold bias temperature fluctuations in the model are in excellent agreement with observations (and with temperatures in the ECMWF model). The extratropical tropopause is too high in the model resulting in somewhat too low ozone at aircraft cruising altitudes. In wintertime a better agreement with observations is achieved. Scatter plots of simulated versus measured CO suggest too strong mixing across the tropopause as seen also in the other models. LMDz-INCA shows the highest levels in OH and H₂O₂ at northern midlatitudes of all models, which is most probably related to the wet bias. Since OH concentrations determine the lifetime of many trace species including CO the concentrations of CO are consistently too low. In addition, due to missing NMHC chemistry secondary production of CO from NMHCs is not well represented in this model version. Nitrogen oxide concentrations are in fairly good agreement with observations at 9 to 13 km but tend to be too high in the free troposphere at lower altitudes. In a recently revised version of the model wet deposition of HNO₃ has been increased. It is expected that this will bring simulated HNO₃ and NO_x in better alignment with observed concentrations. As a positive side-effect OH concentrations are also reduced by approximately 15% in that model version (Didier Hauglustaine, personal communication).

ULAQ: The ULAQ model is run on a coarse grid due to the rather detailed description of microphysical processes and heterogeneous chemistry on aerosols and PSCs. The model is driven by meteorological fields from a GCM and therefore a simulation of the 1995 to 1998 period was not feasible. Hence, only a rather coarse analysis of model performance based on average biases and meridional profiles over the Pacific was possible. The ULAQ model shows no clear tendency of a positive or negative bias with respect to O₃. The agreement with observed O₃ concentrations is generally quite good and follows a pattern very similar to CTM2-Gauss. CO is mainly too low in the model

CTM model
evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

over all regions but deviations from observed values are mostly within 30%. H_2O_2 is mainly too low in the UT/LS region. The ULAQ model shows the best agreement in terms of average PAN concentrations of all models. The analysis of meridional trace gas distributions over the Pacific suggests a somewhat too strong vertical mixing in the model in the tropics. Consequently, CO is too high and O_3 , which is rapidly depleted in the tropical (marine) boundary layer, is too low in the upper tropical troposphere.

E39/C: Similar to ULAQ no point-by-point analysis was feasible since E39/C is a GCM and no nudging was applied. Both with respect to CO and with respect to O_3 E39/C shows a similar behavior as the ULAQ model. Hence, no evidence is found for either a positive or negative bias with respect to O_3 . Yet, CO tends to be generally too low in the model. NO_x concentrations are among the highest when compared with other models and are often somewhat higher than observed. Over North America during spring and summer, however, NO_x concentrations tend to be too low as in most other models. The comparison of meridional trace gas distributions over the Pacific reveals that E39/C simulates a too small contrast in CO between the Southern and Northern Hemisphere. Probably northern hemispheric CO sources are underestimated in the model. Furthermore, ozone tends to be too high in the tropical UT and too low at mid-latitudes. The too small increase in average O_3 concentrations at 9–13 km altitude when changing from the subtropics to midlatitudes is probably due to a positive bias in the tropopause altitude in the extratropics.

General conclusions:

Background ozone concentrations in the lowermost stratosphere differed by more than a factor of two between individual models indicating that substantial improvements at these altitudes are needed. Stratospheric boundary conditions as well as vertical transport and diffusion likely have a strong impact on the results and need to be analyzed carefully. Scatter plots of measured versus simulated ozone and CO indicate too strong two-way mixing across the tropopause which may significantly impact the chemical environment in which aircraft emissions are released. In several models this appears to be responsible for elevated ozone concentrations in the upper tropo-

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

sphere which in turn results in enhanced OH radical production. The abundance of HO_x radicals is a crucial factor in determining the sensitivity of ozone production to the additional release of NO_x from air traffic into the UT (Jaeglé et al., 1999). However, despite the problems in background ozone levels OH concentrations were found to be generally in surprisingly good agreement with observed values.

Vertical model resolution could not be identified as the only major factor being responsible for excessive cross-tropopause transport. The LMDz-INCA and TM3 models, for instance, were run at the same vertical resolution but they show quite different ozone profiles across the tropopause. The results for TOMCAT, on the other hand, are similar to TM3 despite its substantially higher vertical resolution. With respect to the rapid drop in CO concentrations observed above the tropopause, however, models run at higher resolution revealed a better performance, although this drop was clearly underestimated even by these models. Too high concentrations of CO (and other trace gases of tropospheric origin) in the lowermost stratosphere may pose a problem as this likely results in a too high sensitivity of local ozone production to aircraft NO_x emissions.

Results for NO_x differed quite significantly between the models but even more strongly between individual campaigns. Time series that contain elevated concentrations due to fresh lightning emissions (as observed for instance on individual flights of the STREAM98 and POLINAT/SONEX campaigns, (Lange et al., 2001; Thompson et al., 1999)) cannot be reproduced by global models since their lightning parameterizations are only able to calculate grid-cell averaged emissions. This is, perhaps to a lesser extent, also true for convective transport of NO_x . However, not only individual NO_x plumes frequently observed in the upper troposphere (Brunner et al., 1998) were often missed by the models, but also campaign averaged biases revealed a larger underestimation of mean NO_x concentrations for campaigns during which a substantial contribution by lightning was experienced. Even the TM3 model, which simulates the largest lightning contribution in the upper troposphere at mid-latitudes of all models due to its specific parameterization (Meijer et al., 2001), still underestimates NO_x in these circumstances.

Acknowledgements. We would like to thank the many research and commercial aircraft measurement project leaders and managers for allowing us to use their data sets, in particular H. Schlager and U. Schumann, J.-P. Cammas and A. Marenco, C. Brenninkmeijer, H. Matsueda, S. Gaines, A. Thompson, H. Singh, B. Bregman and F. Raes. We would also like to thank Louisa Emmons for her pioneering work in this field and for the many hints on where to obtain such data sets. Finally, we highly appreciated the comments on a draft version by U. Schumann and J. Hoell. This work has been supported by the European Community grant through the project TRADEOFF (contract EVK2-CT-1999-00030).

References

- 10 Berntsen, T.K. and Isaksen, I. S. A.: Effects of lightning and convection on changes in tropospheric ozone due to NO_x emissions from aircraft, *Tellus (B)*, 51, 766–788, 1999. [2501](#), [2502](#)
- Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modelling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *Journal of Geophysical Research*, 106, 23 073–23 096, 2001. [2503](#), [2516](#), [2523](#)
- 15 Brasseur, G. P., Cox, R. A., Hauglustaine, D., Isaksen, I., Lelieveld, J., Lister, D. H., Sausen, R., Schumann, U., Wahner, A., and Wiesen, P.: European scientific assessment of the effects of aircraft emissions, *Atmospheric Environment*, 32, 2329–2418, 1998. [2501](#), [2502](#)
- 20 Bregman, A., Krol, M. C., Teyssère, H., Norton, W. A., Iwi, A., Chipperfield, M., Pitari, G., Sundet, J. K., and Lelieveld, J.: Chemistry-transport model comparison with ozone observations in the midlatitude lowermost stratosphere, *Journal of Geophysical Research*, 106, 17 479–17 496, 2001. [2502](#), [2503](#)
- Brühl, C., Pöschl, U., Crutzen, P. J., et al.: Acetone and PAN in the upper troposphere: Impact on ozone production from aircraft emissions, *Atmospheric Environment*, 34, 3931–3938, 2000. [2502](#)
- 25 Brunner, D., Staehelin, J., and Jeker, D.: Large-scale nitrogen oxide plumes in the tropopause region and implications for ozone, *Science*, 282, 1305–1309, 1998. [2527](#)
- Brunner, D. Staehlin, J., Rogers, H. L., et al.: An evaluation of the performance of chemistry transport models by comparison with scientific aircraft observations, detailed comparison
- 30

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

with the PEM-Tropics A and SONEX campaigns, in preparation, 2003. [2505](#), [2516](#), [2517](#), [2519](#)

Chatfield, R. B.: Anomalous HNO_3/NO_x ratio of remote tropospheric air: Conversion of nitric acid to formic acid and NO_x ?, *Geophysical Research Letters*, 21, 2705–2708, 1994. [2517](#)

Crawford, J., Davis, D., Chen, G., et al.: Photostationary state analysis of the NO_2 -NO system based on airborne observations from the western and central north pacific, *Journal of Geophysical Research*, 101, 2052–2072, 1996. [2515](#)

Emmons, L. K., Hauglustaine, D. A., Müller, J.-F., Carroll, M. A., Brasseur, G. P., Brunner, D., Staehelin, J., Thouret, V., and Marenco, A.: Data composites of airborne observations of tropospheric ozone and its precursors, *Journal of Geophysical Research*, 105, 20 497–20 538, 2000. [2503](#), [2504](#), [2507](#)

Fuelberg, H. E., Newell, R. E., Longmore, S. P., Zhu, Y., Westberg, D. J., Browell, E. V., Blake, D. R., Gregory, G. L., and Sachse, G. W.: A meteorological overview of the Pacific Exploratory Mission (PEM) Tropics period, *Journal of Geophysical Research*, 104, 5585–5622, 1999. [2517](#)

Grewe, V., Brunner, D., Dameris, M., Grenfell, J. L., Hein, R., Shindell, D., and Staehelin, J.: Origin and variability of upper tropospheric nitrogen oxides and ozone at northern mid-latitudes, *Atmospheric Environment*, 35, 3421–3433, 2001. [2502](#), [2503](#)

Hauglustaine, D. A., Ridley, B. A., Solomon, S., Hess, P. G., and Madronich, S.: HNO_3/NO_x ratio in the remote troposphere during MLOPEX2: Evidence for nitric acid reduction on carbonaceous aerosols, *Geophysical Research Letters*, 23, 2609–2612, 1996. [2517](#)

Hauglustaine, D. A., Emmons, L., Newchurch, M., Brasseur, G., Takao, T., Matsubara, K., Johnson, J., Ridley, B., Stith, J., and Dye, J.: On the role of lightning NO_x in the formation of tropospheric ozone plumes: A global model perspective, *Journal of Atmospheric Chemistry*, 38, 277–294, 2001. [2502](#)

Hein, R., Dameris, M., Schnadt, C., Land, C., Grewe, V., Köhler, I., Ponater, M., Sausen, R., Steil, B., Landgraf, J., and Brühl, C.: Results of an interactively coupled atmospheric chemistry-general circulation model: Comparison with observations, *Ann. Geophys.*, 19, 435–457, 2001. [2507](#)

Houghton, J. T., Ding, Y., Griggs, D. J., Noguera, M., van der Linden, P. J., and Xiaosu, D. (eds.): IPCC Third Assessment Report: Climate Change 2001: The Scientific Basis, Cambridge University Press, UK, 2001. [2502](#), [2506](#)

Jaeglé, L., Jacob, D. J., Brune, W. H., Faloon, I. C., Tan, D., Kondo, Y., Sachse, G. W., Ander-

CTM model
evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

- son, B., Gregory, G. L., Vay, S., Singh, H. B., Blake, D. R., and Shetter, R.: Ozone production in the upper troposphere and the influence of aircraft during SONEX: Approach to NO_x saturated conditions, *Geophysical Research Letters*, 26, 3081–3084, 1999. [2527](#)
- Jaeglé, L., Jacob, D. J., Brune, W. H., and Wennberg, P. O.: Chemistry of HO_x radicals in the upper troposphere, *Atmospheric Environment*, 35, 469–489, 2001. [2502](#), [2519](#)
- Jourdain, L., Hauglustaine, D. F., and Hourdin, F.: The global distribution of lightning NO_x simulated on-line in a general circulation model, *J. Phys. Chem. Earth (C)*, 26, 585–591, 2001. [2502](#)
- Kley, D., Crutzen, P. J., Smit, H. G. J., Vmel, H., Oltmans, S. J., Grassl, H., and Ramanathan, V.: Observations of near-zero ozone concentrations over the convective pacific: Effects on air chemistry, *Science*, 274, 230–233, 1996. [2522](#)
- Köhler, M. O. et al.: Model intercomparison and validation with observations made aboard commercial aircraft, *Atmospheric Environment*, in preparation, 2003. [2504](#), [2505](#)
- Lacis, A. A., Wuebbles, D. J., and Logan, J. A.: Radiative forcing of climate by changes in the vertical distribution of ozone, *Journal of Geophysical Research*, 95, 9971–9981, 1990. [2501](#)
- Lange, L., Hoor, P., Helas, G., Fischer, H., Brunner, D., Scheeren, B., Williams, J., Wong, S., Wohlfrom, K.-H., Arnold, F., Ström, J., Krejci, R., Lelieveld, J., and Andreae, M. O.: Detection of lightning-produced NO in the midlatitude upper troposphere during STREAM 98, *Journal of Geophysical Research*, 106, 27 777–27 785, 2001. [2527](#)
- Law, K. S., Plantevin, P. H., Thouret, V., et al.: Comparison between global chemistry transport model results and measurements of ozone and water vapor by airbus in-service aircraft (MOZAIC) data, *Journal of Geophysical Research*, 105, 1503–1525, 2000. [2502](#)
- Lawrence, M. G. and Crutzen, P. J.: The impact of cloud particle gravitational settling on soluble trace gas distributions, *Tellus, Ser. B*, 50, 263–289, 1998. [2517](#)
- Levy II, H., Moxim, W. J., Klonecki, A. A., and Kasibhatla, P. S.: Simulated tropospheric NO_x: Its evaluation, global distribution and individual source contributions, *Journal of Geophysical Research*, 104, 26 269–26 306, 1999. [2503](#)
- Marenco, A., Thouret, V., Nedelec, P., et al.: Measurement of ozone and water vapor by airbus in-service aircraft: The MOZAIC airborne program, an overview, *Journal of Geophysical Research*, 103, 25 631–25 642, 1998. [2503](#)
- Meijer, E. W., van Velthoven, P. F. J., Thompson, A. M., Pfister, L., Schlager, H., Schulte, P., and Kelder, H.: Model calculations of the impact of NO_x from air traffic, lightning, and surface emissions, compared with measurements, *Journal of Geophysical Research*, 105, 3833–

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

3850, 2000. [2502](#)

Meijer, E. W., van Velthoven, P. F. J., Brunner, D., and Kelder, H.: Improvement and evaluation of the parameterisation of nitrogen oxide production by lightning, *J. Phys. Chem. Earth (C)*, 26, 577–583, 2001. [2519](#), [2527](#)

5 NASA: Atmospheric Effects of Aviation, A Review of NASA's Subsonic Assessment Project, National Academy Press, Washington, D.C., 1999. [2501](#)

Penner, J. E., Lister, D., Griggs, D., Docken, D., and MacFarland, M., (eds.): IPCC Special Report on Aviation and the Global Atmosphere, Cambridge University Press, New York, 1999. [2501](#), [2502](#)

10 Pickering, K., Wang, Y., Wei-Kuo, T., Price, C., and Müller, J.-F.: Vertical distributions of lightning NO_x for use in regional and global chemical transport models, *Journal of Geophysical Research*, 103, 31 203–31 216, 1998. [2524](#)

Pitari, G. and Mancini, E.: Climatic impact of future supersonic aircraft: Role of water vapour and ozone feedback on circulation, *J. Phys. Chem. Earth (C)*, 26, 571–576, 2001. [2502](#)

15 Prather, M. J. and Jacob, D. J.: A persistent imbalance in HO_x and NO_x photochemistry of the upper troposphere driven by deep tropical convection, *Geophysical Research Letters*, 24, 3189–3192, 1997. [2502](#), [2519](#)

Rogers, H. L., Chipperfield, M. P., Bekki, S., and Pyle, J. A.: The effects of future supersonic aircraft on stratospheric chemistry modeled with varying meteorology, *Journal of Geophysical Research*, 105, 29 359–29 369, 2000. [2502](#)

Schultz, M. G., Jacob, D. J., Bradshaw, J. D., Sandholm, S. T., Dibb, J. E., Talbot, R. W., and Singh, H. B.: Chemical NO_x budget in the upper troposphere over the tropical south pacific, *Journal of Geophysical Research*, 105, 6669–6679, 2000. [2517](#)

Singh, H. B., Thompson, A. M., and Schlager, H.: The 1997 SONEX aircraft campaign and coordinated POLINAT 2 activity: Overview and accomplishments, *Geophysical Research Letters*, 26, 3053–3056, 1999. [2511](#)

25 Stockwell, D., Giannakopoulos, C., Plantevin, P.-H., Carver, G. D., Chipperfield, M. P., Law, K. S., Pyle, J. A., Shallcross, D. E., and Wang, K.-Y.: Modelling NO_x from lightning and its impact on global chemical fields, *Atmospheric Environment*, 33, 4477–4493, 1999. [2502](#)

30 Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research*, 106, 7183–7192, 2001. [2509](#), [2510](#)

Thompson, A. M., Sparling, L. C., Kondo, Y., Anderson, B. E., Gregory, G. L., and Sachse, G. W.: Perspectives on NO , NO_y , and fine aerosol sources and variability during SONEX, *Geophys-*

ical Research Letters, 26, 3073–3076, 1999. [2511](#), [2527](#)
Wang, Y., Logan, J. A., and Jacob, D. J.: Global simulation of tropospheric O₃-NO_x-hydrocarbon chemistry – 2. model evaluation and global ozone budget, Journal of Geophysical Research, 103, 10 727–10 755, 1998. [2503](#), [2516](#), [2517](#), [2523](#)

ACPD

3, 2499–2545, 2003

**CTM model
evaluation – Part 1**

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

CTM model
evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

Table 1. Model properties

Model	TM3	CTM2	CTM2- Gauss	TOMCAT	LMDz/ INCA	ULAQ	E39/C
Operated by	KNMI	Univ. Oslo	Univ. Oslo	Cambridge	IPSL	Univ. Aquila	DLR
Model type	CTM	CTM	CTM	CTM	GCM	CTM	GCM
Meteorology	ECMWF	ECMWF	ECMWF	ECMWF	ECMWF	ULAQ-GCM	GCM
Lat x lon resolution	3.75° x 5°	T21	T21	T21	2.5° x 3.75°	10° x 20°	T30
Vertical levels	19 hybrid	19 hybrid	40 hybrid	31 hybrid	19	26 log-p	39 hybrid
Model top (hPa)	10 hPa	10 hPa	10 hPa	10 hPa	4 hPa	0.04 hPa	10 hPa
Transport scheme	slopes ⁽¹⁾	2nd order moments ⁽²⁾	2nd order moments ⁽²⁾	2nd order moments ⁽²⁾	Van Leer ⁽³⁾	Eulerian fully explicit	semi-lagr. ⁽⁴⁾
Vertical velocities	hor. wind divergence	hor. wind divergence	hor. wind divergence	hor. wind divergence	hor. wind divergence	radiation scheme ⁽⁵⁾	hor. wind divergence
Convection	Tiedtke ⁽⁶⁾	Tiedtke ⁽⁶⁾	Tiedtke ⁽⁶⁾	Tiedtke ⁽⁶⁾	Tiedtke ⁽⁶⁾	Müller and Brasseur ⁽⁷⁾	Tiedtke ⁽⁶⁾
Lightning param.	Meijer ⁽⁸⁾	Price ⁽⁹⁾	Price ⁽⁹⁾	Stockwell ⁽¹⁰⁾	Jourdain ⁽¹¹⁾	Grewe ⁽¹²⁾	Grewe ⁽¹²⁾
Dynamical/chemical timestep (min)	120/120	60/60	60/60	30/15	15/30	60/60	30/30
Transported species	24	42	76	27	27	40 ⁽¹⁴⁾	13
Total species	39	52	98	49	45	70 ⁽¹⁴⁾	37
Gas phase + photolytic reactions	67 + 24	69 + 18	163 + 50	101 + 27	78 + 28	131 (40)	107
Heter. reactions	4 (aqueous)	2	7	0	4	10	4
Strat. chemistry	no	no	yes	no	no Cl or Br	yes	yes
NMHC chemistry	yes	yes	yes	Ethane/Propane	no	yes	no
Acetone chemistry	no	yes	yes	no	no	no	no
References	Meijer et al. (2000)	Sundet (1997); Kraabol et al. (2003)	Kraabol Rum-mukainen et al. (1999)	Law et al. (2000)	Jourdain (2000)	Jourdain et al. (2001); Pitari et al. (2002)	Hein et al. (2001)

⁽¹⁾Russel and Lerner (1981)⁽²⁾Prather (1986)⁽³⁾van Leer (1979)⁽⁴⁾Williamson and Rasch (1994)⁽⁵⁾Pitari (1993)⁽⁶⁾Tiedtke (1989)⁽⁷⁾Müller and Brasseur (1995)⁽⁸⁾Meijer et al. (2001)⁽⁹⁾Price et al. (1997), but zonally redistributed according to convection⁽¹⁰⁾Stockwell et al. (1999)⁽¹¹⁾Jourdain et al. (2001)⁽¹²⁾Grewe et al. (2001)⁽¹³⁾Grewe et al. (2002)⁽¹⁴⁾plus 44 aerosol bins

**CTM model
evaluation – Part 1**

D. Brunner et al.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Print Version](#)
[Interactive Discussion](#)

© EGU 2003

Table 2. Year-2000 emissions recommended for the TRADEOFF model runs

Species	Emission source	Strength (Tg/yr)	
NO _x	fossil & bio fuels (~30+1.8)	31.8 (N)	
	savannah & ag-waste burning/deforestation (3.2+1.2+2.7)	7.1	
	aircraft (IPCC-TAR year 2000)	0.68	
	soils	5.5	
	lightning	5.0	
		Total 50.08	
CO	fossil fuel/domestic burning	650	
	deforestation/savannah & waste burning	700	
	vegetation(150)+oceans(50)	200	
			Sub-Total 1550
	CH ₄ oxidation	~800	
	isoprene oxidation ⁽¹⁾	~165	
	terpene oxidation ⁽¹⁾ (included in vegetation)		
	biomass burning NMHC oxidation ⁽¹⁾	~30	
	acetone oxidation ⁽¹⁾	~20	
			Total ~2675
NMHC	(if considered)		
	fossil fuel/domestic burning	161 (C)	
	deforestation/savannah & waste burning	34	
	isoprene	220	
	terpene	127	
	acetone	30	
		Total 572	

⁽¹⁾ use only if not treated separately as NMHC emission.

Table 3. Measurement campaigns and programs included in the TRADEOFF database (see <http://www.iac.ethz.ch/tradeoff/database> for further details)

Campaign	Periods	Trace gases	Records
MOZAIC (Marenco et al., 1998)	1995 - 1998	O ₃ , H ₂ O	379501
NOXAR (Brunner et al., 2001)	05 May 1995 - 13 May 1996, 12 Aug 1997 - 13 Nov 1997	NO, NO ₂ , O ₃	44522
JAL (Matsueda et al., 1998)	1995 - 1998	CO, CH ₄ , CO ₂	1115
O ₃ soundings	1995 - 1998	O ₃ , H ₂ O	332521
CARIBIC (Brenninkmeijer et al., 1999)	Nov 1997 - Dec 1998	O ₃ , CO	1143
STREAM (Bregman et al., 1995)	09 - 14 Feb 1995 23 Nov - 03 Dec 1995 22 May - 01 Jun 1996 09 Mar - 25 Mar 1997 27 Jun - 24 Jul 1998	1995: O ₃ , CO, HNO ₃ , NO _y , acetone, N ₂ O 1996: O ₃ , H ₂ O, HNO ₃ , NO _y , acetone, N ₂ O 1997: O ₃ , CO, H ₂ O, HNO ₃ , NO _y , acet., N ₂ O 1998: O ₃ , CO, H ₂ O, NO, HNO ₃ , NO _y , acetone, H ₂ SO ₄ , CH ₃ CN, CO ₂ , N ₂ O (twice)	1398
STRAT	01 May - 18 May 1995 19 Oct - 09 Nov 1995 22 Jan - 05 Feb 1996 15 Jul - 10 Aug 1996 13 - 23 Sep 1996 02 - 20 Dec 1996	O ₃ , CO, H ₂ O, OH, HO ₂ , NO, NO _y , ethane, ethyne, N ₂ O, CO ₂ , CH ₄	2331
POLINAT (Schumann et al., 2000)	21 Jun - 05 Jul 1995 19 Sep - 25 Oct 1997	1995: O ₃ , H ₂ O, NO, HNO ₃ , acetone 1996: O ₃ , CO, H ₂ O, NO, NO ₂ , NO _y , HNO ₃ , acetone, J _{NO2}	737
ACE-1 (Bates et al., 1998)	31 Oct - 22 Dec 1995	O ₃ , CO, OH, H ₂ O ₂ , H ₂ O, NO	2757
TOTE/VOTE	06 Dec 1995 - 19 Feb 1996	O ₃ , CO, H ₂ O, NO, NO _y , N ₂ O	1354
SUCCESS (Toon and Mlake-Lye, 1998)	10 Apr - 16 May 1996	O ₃ , CO, H ₂ O, OH, HO ₂ , NO, NO _y	988
PEM-Tropics A (Hoeil et al., 1999)	30 Aug - 05 Oct 1996	DC-8: O ₃ , CO, H ₂ O, H ₂ O ₂ , NO, NO ₂ , HNO ₃ , PAN, ethane, J _{NO2} ; P3-B: O ₃ , CO, H ₂ O, H ₂ O ₂ , OH, NO, HNO ₃ , ethane, J _{NO2}	3222
ACSOE ⁽¹⁾	09 - 19 Sep 1996 02 - 10 Apr 1997 11 Aug - 23 Sep 1997	1996: O ₃ , H ₂ O, H ₂ O ₂ 1997: O ₃ , CO, H ₂ O ₂ , NO, NO _y , HCHO, J _{NO2}	1098
POLARIS	16 Apr - 15 May 1997 20 Jun - 12 Jul 1997 02 - 25 Sep 1997	O ₃ , CO, H ₂ O, OH, HO ₂ , NO, NO ₂ , NO _y , ethane, ethyne, N ₂ O, CO ₂ , CH ₄	1899
ACE-2 (Johnson et al., 2000)	16 Jun - 25 Jul 1997	ARAT: O ₃ , H ₂ O; C-130: O ₃ , CO, H ₂ O ₂	1847
SONEX (Thompson et al., 1999)	07 Oct - 12 Nov 1997	O ₃ , CO, OH, HO ₂ , H ₂ O ₂ , H ₂ O, NO, NO _y , HNO ₃ , PAN, acetone, ethane, ethene, J _{NO2}	1325

⁽¹⁾ ACSOE consisted of several sub-campaigns, including OXICOA (9–19 Sep 1996), TACIA (2–10 Apr 1997) and NARE (11 Aug–23 Sep 1997). Only data from the C-130 aircraft are included. See <http://www.uea.ac.uk/~acsoe/report.html>

CTM model evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

CTM model
evaluation – Part 1

D. Brunner et al.

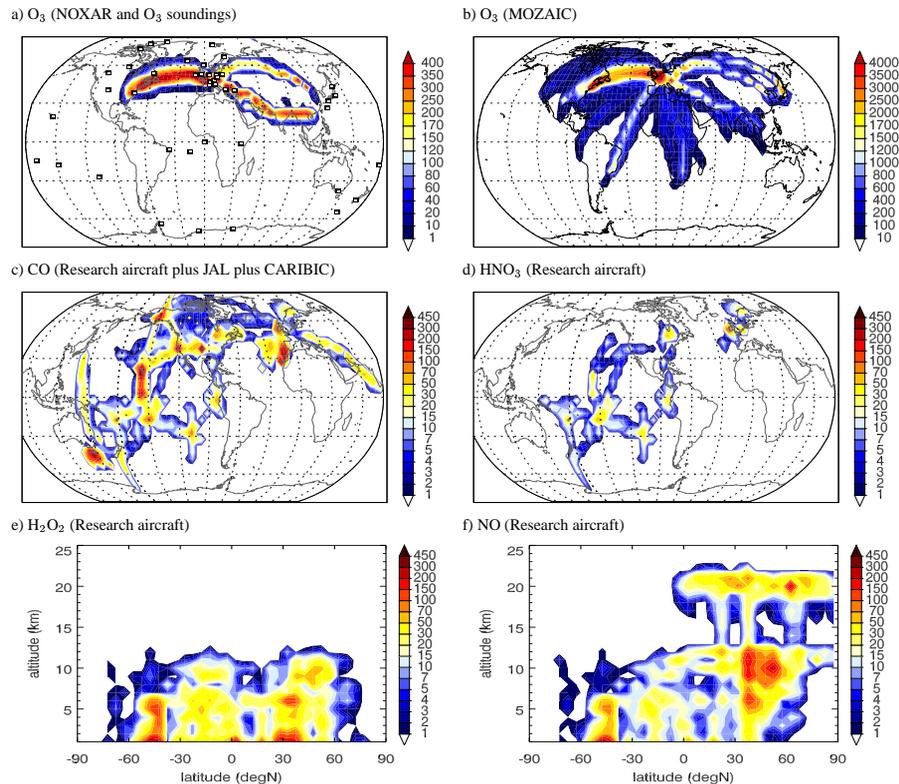


Fig. 1. Data density distributions of selected tracers in units of samples per $5^\circ \times 5^\circ$ grid box (vertically integrated, panels a–d) and samples per 5° latitude \times 1 km altitude grid box (zonally integrated, panels e–f), respectively. The rectangles in (a) are the positions of the 45 ozone sounding stations included in the database. (c)–(f) are distributions of selected tracers as provided by research aircraft campaigns. Also included in (c) are measurements from the two commercial aircraft measurement programs JAL (between Japan and Australia) and CARIBIC (between Germany and India).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

CTM model evaluation – Part 1

D. Brunner et al.

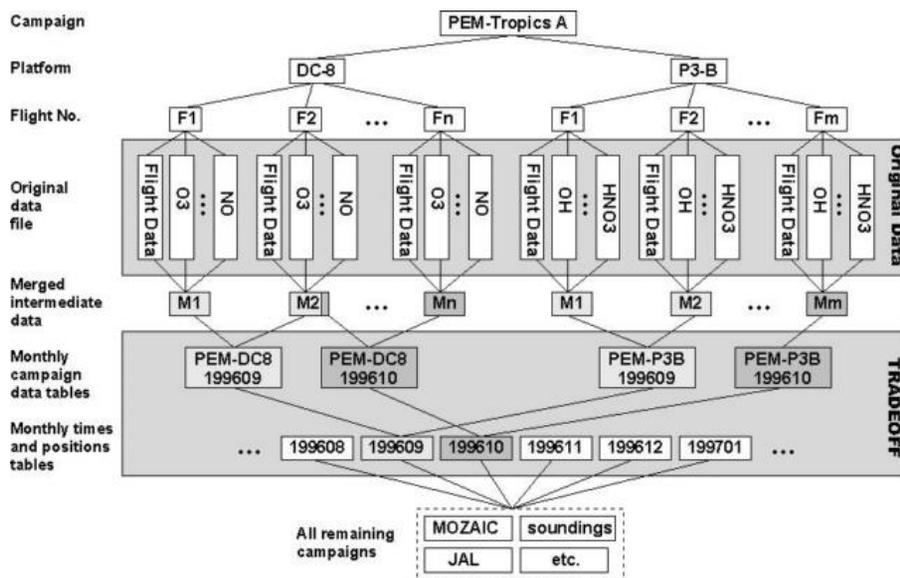


Fig. 2. Data file processing steps and structure of the TRADEOFF database. See text for further details.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Print Version

Interactive Discussion

CTM model
evaluation – Part 1

D. Brunner et al.

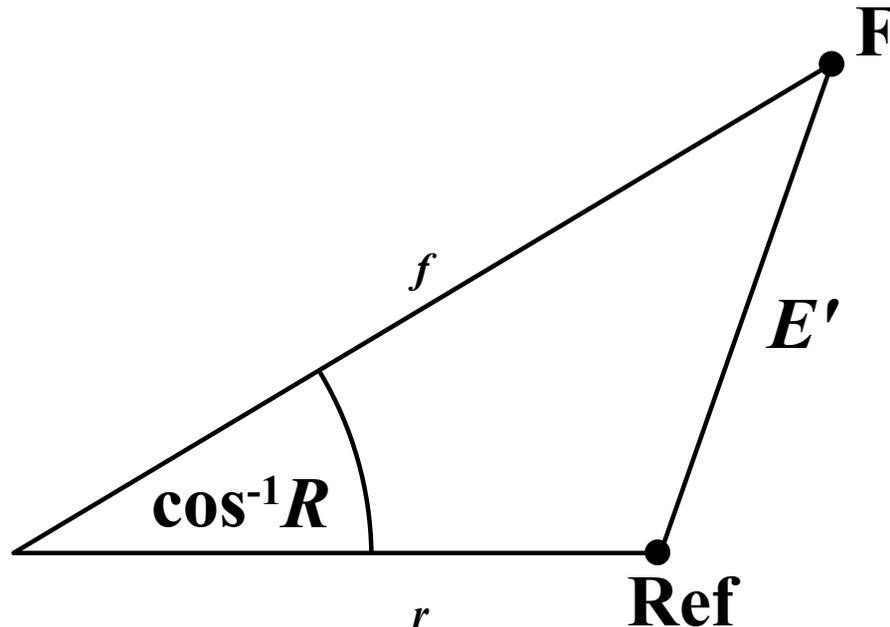


Fig. 3. Geometric relationship in a Taylor diagram between the correlation coefficient R , the root mean square (RMS) error E' and the standard deviations of the test field σ_f and reference field σ_r , respectively.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

CTM model
evaluation – Part 1

D. Brunner et al.

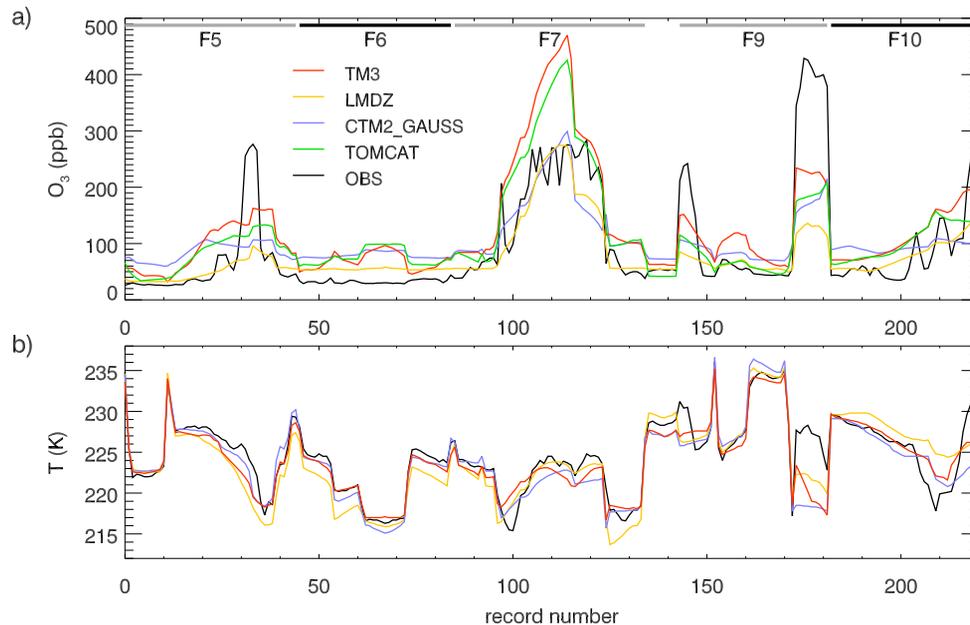


Fig. 4. Time series of (a) ozone and (b) temperature in the UT/LS region ($p < 300$ hPa) along the flight tracks of five consecutive flights of the NASA DC-8 aircraft during the POLINAT/SONEX campaign in Oct/Nov 1997. Measured values are shown in black and interpolated model fields in color. Temperatures in the LMDz-INCA GCM have a cold bias at these altitudes. The line was therefore shifted by 5° C to match the ECMWF temperatures. Temperature is not available in the output of the TOMCAT model.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)

© EGU 2003

CTM model
evaluation – Part 1

D. Brunner et al.

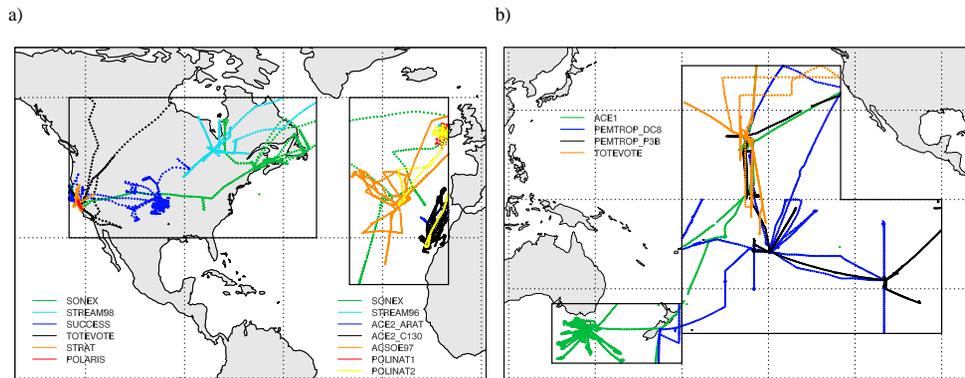


Fig. 5. The four separate domains used for the overall evaluation: North America, 30° to 60° N, 60° to 125° W; North Atlantic: 20° to 60° N, 10° to 40° W; Pacific 45° S to $^{\circ}$ N, 125° to 180° W (North Pacific) or 90° to 180° W (South Pacific); Tasmania: 35° to 55° S, 135° to 180° E. Overlaid are the measurement flights in different colors for the different campaigns.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

CTM model evaluation – Part 1

D. Brunner et al.

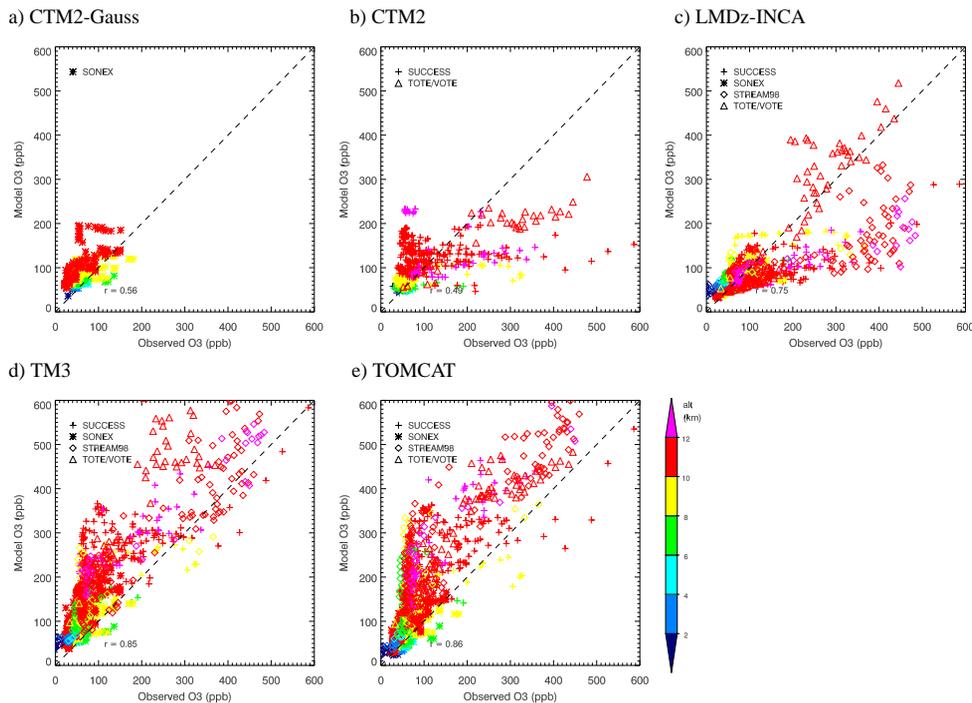


Fig. 6. Scatter plots of modelled versus measured ozone concentrations over North America. **(a)** CTM2-Gauss, **(b)** CTM2, **(c)** LMDz-INCA, **(d)** TM3, **(e)** TOMCAT. Different campaigns are represented by different symbols. Symbols are colored according to the altitude range (see the color bar). The dashed line is the 1:1 ratio.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

CTM model
evaluation – Part 1

D. Brunner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003

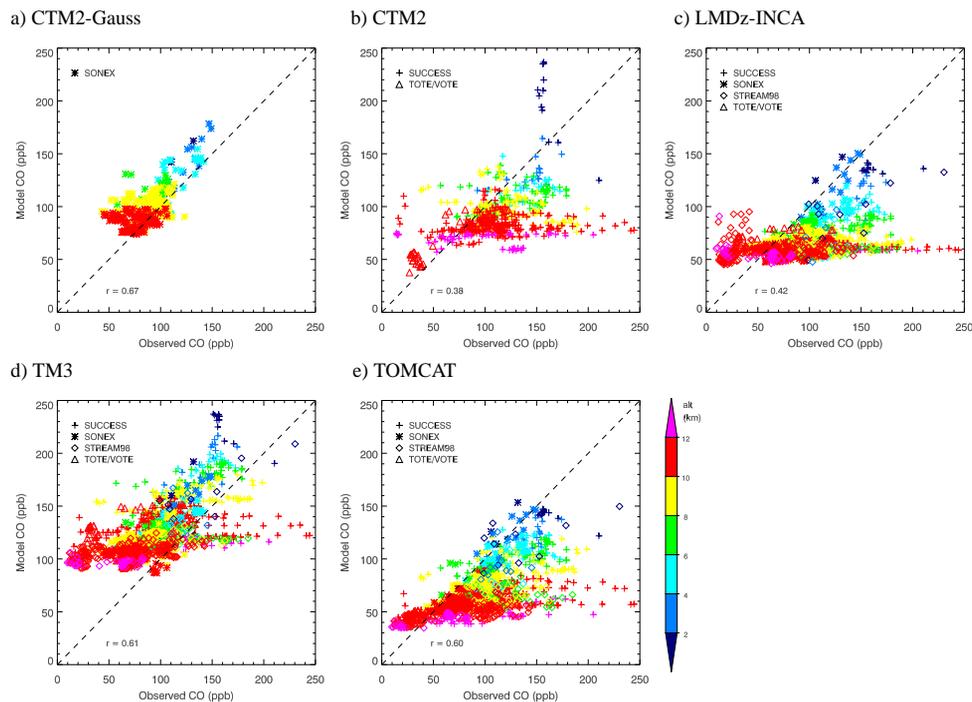


Fig. 7. Scatter plots of modelled versus measured CO concentrations over North America. See Fig. 6 for further details.

CTM model
evaluation – Part 1

D. Brunner et al.

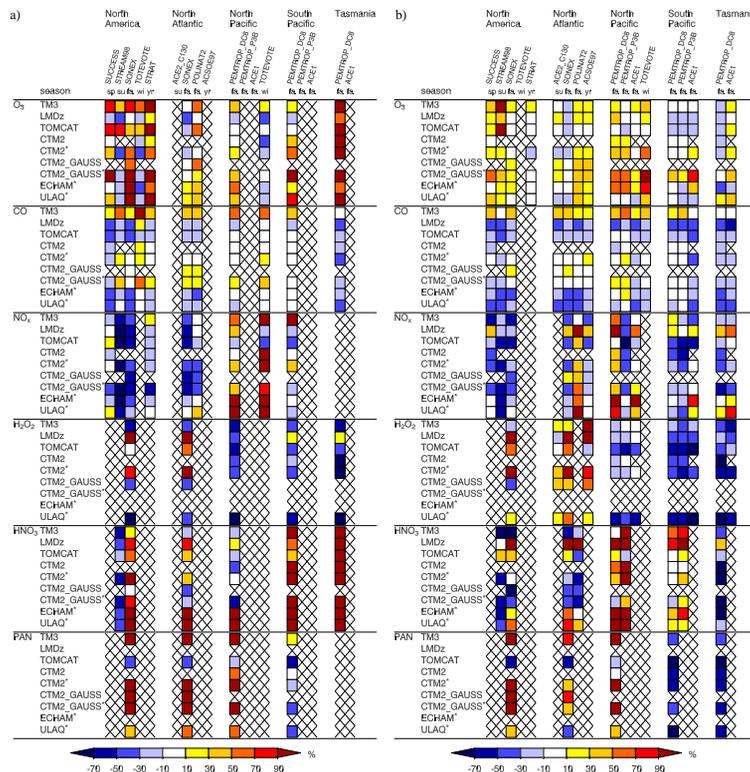


Fig. 8. Model biases $(\text{model-meas})/\text{meas} \times 100\%$ in the altitude range of **(a)** 300 to 170 hPa (approx. 9–13 km), and **(b)** 510 to 350 hPa (approx. 5.5 to 8 km). Biases are shown for each measurement campaign separately and grouped by the four regions. Campaign names and the corresponding season are indicated at the top (sp = spring (MAM), su = summer (JJA), fa = fall (SON), wi = winter (DJF), yr = measurements in several different seasons). The different model versions are listed in the left column. The star (*) behind a model name indicates that the values were obtained from monthly mean fields instead of point-by-point output.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

CTM model evaluation – Part 1

D. Brunner et al.

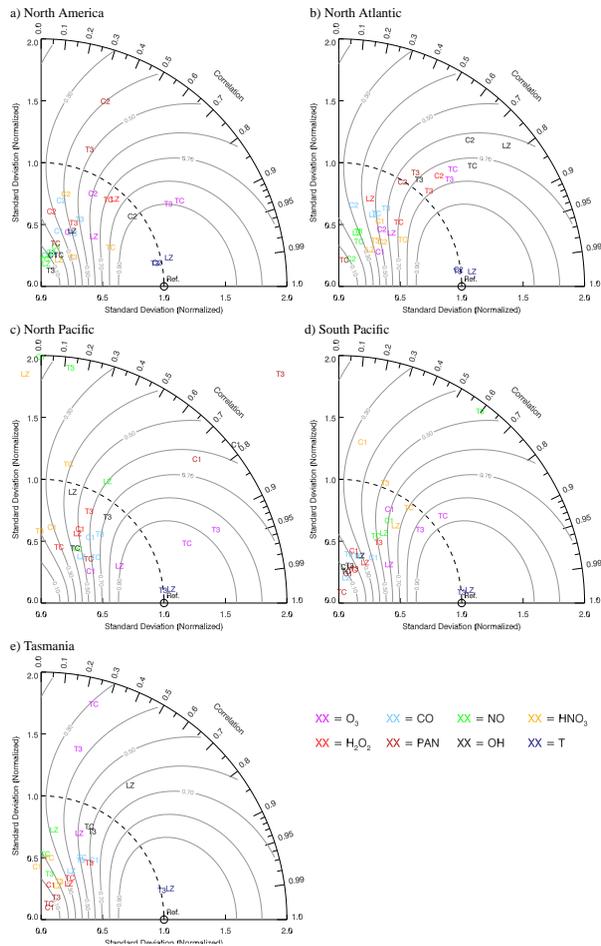


Fig. 9. Taylor diagrams for the different regions and for the altitude range of 170 hPa < p < 510 hPa (approx. 5.5–13 km).

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Print Version

Interactive Discussion

CTM model
evaluation – Part 1

D. Brunner et al.

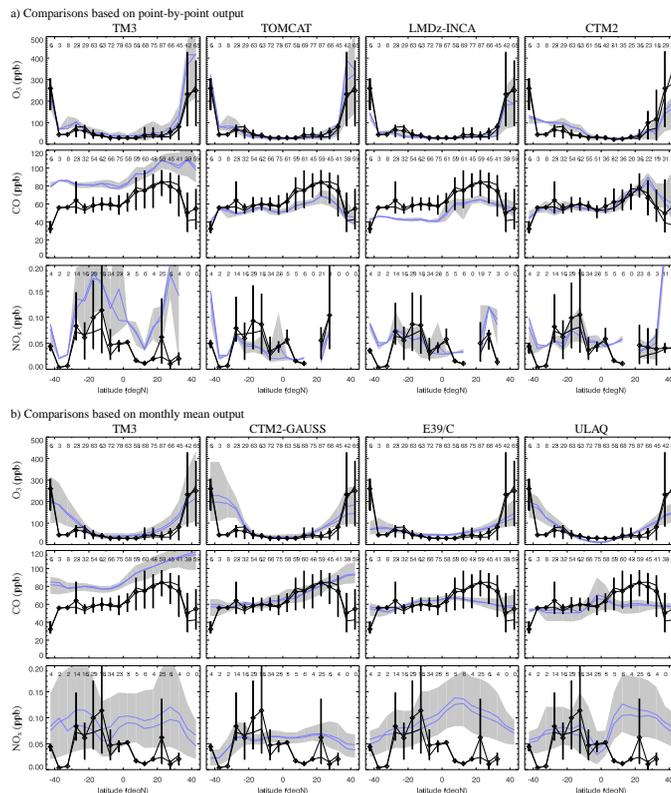


Fig. 10. Meridional distributions over the Pacific Ocean at 9–13 km altitude. **(a)** Derived from point-by-point output, **(b)** Derived from monthly mean output. The range of model values is shown as light grey shading and mean/median values as thick/thin blue lines. The range of measured values is indicated by vertical bars and mean/median values are connected by thick/thin black lines. The number of observations per 5° latitude bin is shown at the top of each panel.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Print Version

Interactive Discussion

© EGU 2003