Annales
Geophysicae

A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements

K. Gierens[1], **U. Schumann**[1], **M. Helten**[2], **H. Smit**[2], **A. Marenco**[3]

[1] Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, D-82234 Weßling, Germany
Fax: +4918153128-1841, e-mail: klaus.gierens@der.de
[2] Forschungszentrum Jülich, Institut für Chemie der belasteten Atmosphäre (ICG2), Postfach 1913, D-52425 Jülich, Germany
[3] Laboratoire d'Aerologie, CNRS, 14 Avenue Edouard Belin, F-31400 Toulouse, France

Abstract. Data from three years of MOZAIC measurements made it possible to determine a distribution law for the relative humidity in the upper troposphere and lower stratosphere. Data amounting to 13.5% of the total were obtained in regions with ice supersaturation. Troposphere and stratosphere are distinguished by an ozone concentration of 130 ppbv as threshold. The probability of measuring a certain amount of ice supersaturation in the troposphere decreases exponentially with the degree of ice supersaturation. The probability of measuring a certain relative humidity in the stratosphere (both with respect to water and ice) decreases exponentially with the relative humidity. A stochastic model that naturally leads to the exponential distribution is provided. Mean supersaturation in the troposphere is about 15%, whereas ice nucleation requires 30% supersaturation on the average. This explains the frequency of regions in which aircraft induce persistent contrails but which are otherwise free of clouds. Ice supersaturated regions are 3–4 K colder and contain more than 50% more vapour than other regions in the upper troposphere. The stratospheric air masses sampled are dry, as expected, having mean relative humidity over water of 12% and over ice of 23%, respectively. However, 2% of the stratospheric data indicate ice supersaturation. As the MOZAIC measurements have been obtained on commercial flights mainly between Europe and North America, the data do not provide a complete global picture, but the exponential character of the distribution laws found is probably valid globally. Since water vapour is the most important greenhouse gas and since it might enhance the anthropogenic greenhouse effects via positive feedback mechanisms, it is important to represent its distribution correctly in climate models. The discovery of the distribution law of the relative humidity makes possible simple tests to show whether the hydrological cycle in climate models is represented in an adequate way or not.

1 Introduction

Water vapour is the most important carrier of energy in the atmosphere and is also the most important greenhouse gas. Its spatial and temporal distribution therefore fundamentally impacts weather and climate through a variety of processes, such as exchange of latent heat, radiative cooling and heating, cloud formation and precipitation. Lower tropospheric water vapour amplifies the predicted global warming due to CO_2 doubling (Manabe and Wetherald, 1967) with a climate sensitivity factor of about 1.6 (IPCC, 1990). The role of upper tropospheric humidity (UTH) in the climate system is much less clear. Lindzen (1990) and Pierrehumbert (1994) discuss possible mechanisms that lead to a decrease of UTH in response to global warming. This might reduce or even reverse the water vapour feedback. On the other hand, Rind et al. (1991) provide arguments for moistening of the upper troposphere as a result of global warming.

Unfortunately, the water vapour content of the upper troposphere is not well known. The radiosondes which furnish a wealth of humidity data from the lower and middle troposphere do not work properly at higher altitudes, i.e. in the upper troposphere and lower stratosphere. The current knowledge about UTH is drawn from satellite measurements. Satellites work either in nadir viewing or limb sounding mode. In nadir viewing mode the satellites receive radiation from a relatively thick layer (a few kilometres or several hundreds of hectopascals) of the upper troposphere. The water vapour at different altitudes contributes differently to the measured signal, and often the maximum contribution originates from altitudes below 300 hPa, which may already include the middle tropo-

Correspondence to: K. Gierens

sphere. On the other hand, limb sounding is able to yield high vertical resolution, however, at the expense of a rather poor horizontal resolution (typically some hundreds of kilometres). Both satellite measurement modes incorporate a retrieval process. The results should therefore be validated and gauged against in-situ data.

The aim of the present study is to provide some statistical characteristics of UTH obtained from in-situ measurements. For instance we would like to know how often there is ice supersaturation in clear air and by how much, on average, these regions are supersaturated. Comparison of mean temperatures and specific humidities inside and outside supersaturated regions yields information on formation mechanisms of such regions. Ice supersaturation is a prerequisite for the formation of cirrus clouds and contrails. Short-lived contrails form even in very dry air when the ambient temperature is cold enough (Schumann, 1996), but contrails persist and grow to extended cirrus clouds only when the ambient humidity exceeds ice-saturation (Gierens, 1996). Otherwise, the formation of cirrus clouds generally requires humidities far in excess of ice saturation (e.g. Heymsfield *et al.*, 1998). The question arises of how often the critical humidities for ice nucleation are reached in ice supersaturated regions. It is evident that statistical properties of UTH are important to know for studies of high-altitude cloudiness and its impact on the Earth's radiation balance. Furthermore we are interested in statistical properties of the humidity field in the lower stratosphere. Such knowledge contributes to an observational basis for, amongst others, models of stratospheric chemistry and models that aim to explain why the stratosphere is so dry.

The only data set present today that allows these questions to be answered is the MOZAIC (measurement of ozone on airbus in-service aircraft) data set (Marenco *et al.*, 1998). This is a unique data set that comprises meteorological quantities like temperature and pressure and data about atmospheric composition, i.e. ozone and water vapour abundance. The automatic measurement instruments are installed aboard five in-service aircraft of several European airlines. This guarantees a continuous collection of data with high spatial and temporal resolution. On the other hand, the data base covers only the major flight routes and mainly the cruise altitudes of international flights. The cruise altitudes are generally above 9 km (about 300 hPa), so the data indeed refer to the upper troposphere and lower stratosphere as desired.

In the following section we give some information about the data base and evaluation procedure, including some basics about measures of water vapour for the sake of completeness. Results are presented in Sect. 3 and interpreted in terms of stochastic processes in Sect. 4. Also possible biases resulting from the particular way of data sampling using commercial aircraft as observation platforms are discussed briefly. Findings are summarised and conclusions drawn in Sect. 5.

2 The data base and measures of water vapour

2.1 The data base

We have used for the present study MOZAIC data from the period January 1995 – December 1997, i.e. three complete years. Since we are interested in properties of the upper troposphere and lower stratosphere, we have constrained the data base to data obtained in the pressure range 175–275 hPa. More than 95% of the MOZAIC flights reach the corresponding cruise altitudes. In total, we have evaluated data from 5269 flights. The four seasons contribute each about one quarter to the number of flights. Annual means thus are not dominated by a certain time of the year. The total number of data lines used for the study exceeds 1.7 million. Each data line contains 1 min averages (corresponding to a flight distance of about 15 km) of pressure p, temperature T, relative humidity with respect to liquid water RH, a validity tag and the relative uncertainty of the humidity measurement, the computed water vapour mixing ratio q (mass mixing ratio, units g/kg), and the ozone mixing ratio m_{O_3} (volume mixing ratio, units: parts per billion by volume, ppbv). The data lines contain furthermore the current flight position (longitude λ, latitude β, altitude z) and time. Most data (more than 40%) in the data base refer to flights between Europe and North America, other regions of the world are less represented or even not at all. The spatial coverage of the MOZAIC flights together with indications of percentage of flights on the major routes can be seen in Fig. 5 of the overview paper by Marenco *et al.* (1998). The considered pressure range of 175–275 hPa is divided into two layers, centred around 200 and 250 hPa, respectively. The 200 hPa layer contains about 43% of the flight data, the lower one at 250 hPa contains about 54% of the data.

The humidity measurements on the data base are tagged with a validity indicator, which is an integer number between 0 and 3. Validity "0" means "device not working"; there are no such data on the data base. Measurements labelled with validity "3" are qualified "unreliable" and are excluded from our analysis. Only data with a validity label "1" (reliable measurement) or "2" (humidity lower than detection threshold of the hygrometer) have been included in the evaluation. Those were about 77% of all data. Out of these, 89.5% bear a validity tag "1", the remaining 10.5% accordingly have validity "2". The relative uncertainty is also quoted in the database, even for RH measurements above the detection limit of 5% which are tagged with "1". It is deduced by the law of error propagation from the uncertainties of the individual sensor calibration, the differences of pre- and post-flight calibration coefficients, and the uncertainties of the parameters like total recovery temperature, total air temperature, and static air temperature used to compute the ambient relative humidity RH from the sensor reading inside the sensor housing (Helten *et al.*, 1998). Figure 1 shows the absolute uncertainty as a mean value over the RH

measurements of 1995–1997 in the pressure range 175–275 hPa as function of the measured *RH* value. For *RH* values below the detection limit (5%) an uncertainty cannot be given. The precision characterises the reproducibility of the *RH* measurement with the same sensor. The curve shown in Fig. 1 is a mean value estimated from three independent calibrations of three different sensors. The precision is included in the uncertainty and more than a factor of seven smaller. The main contribution to the uncertainty is that through the differences of the pre- and post-flight calibration coefficients.

The MOZAIC ozone measurements allow us to characterise the data as tropospheric or stratospheric. We identify the tropopause with an ozone mixing ratio of 130 ppbv (Bethan *et al.*, 1997; Duhnke *et al.*, 1998), i.e. we consider data as tropospheric for $m_{O_3} <$ 130 ppbv, whereas $m_{O_3} >$ 130 ppbv implies data from stratospheric air. According to Duhnke *et al.* (1998) the 130 ppbv threshold is a mean O_3-concentration at the thermal tropopause; however, the standard variation of this concentration is quite large, i.e. ±92 ppbv, because of the sharp O_3-gradient at the tropopause. Fortunately, the choice of the O_3-tropopause is not a critical parameter for our investigation. Using the 130 ppbv ozone-tropopause we find that 61.8% of the data included in the evaluation are tropospheric and the remaining 38.2% are stratospheric.

2.2 Measures of water vapour

The MOZAIC data base provides two quantities that give information on the water vapour content of the atmosphere around the current flight position. These are the relative humidity with respect to liquid water, *RH*, which is a measure that depends on ambient temperature, and the water vapour mixing ratio, *q*, which is the vapour mass per unit dry air mass. Both measures are related to each other to a very good approximation as follows:

$$RH \approx \frac{qp}{\varepsilon e_s(T)} \quad , \tag{1}$$

where $e_s(T)$ is the saturation partial pressure of water vapour over a plane of liquid water at temperature *T*, $\varepsilon = 0.622$ is the ratio of the molar masses of water and dry air.

Since we are dealing with the cold atmospheric layers around the tropopause, the relative humidity with respect to liquid water is not necessarily the best measure to quantify the water vapour concentration. A measure that is more relevant in these layers, e.g. for cloud formation processes, is the relative humidity with respect to ice, *RHi*. Therefore, we will focus our analysis mainly upon this latter quantity. *RH* and *RHi* are related in a simple way, but their ratio *RH*/*RHi* varies with temperature:
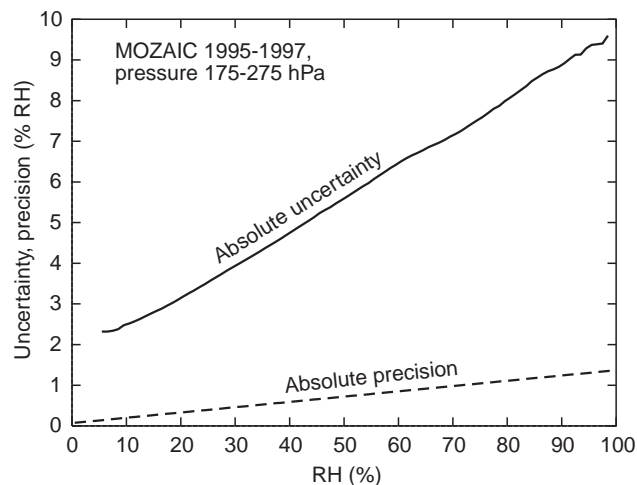
$$RHi = RH \cdot \frac{e_s(T)}{e_{si}(T)} \quad , \tag{2}$$

where $e_{si}(T)$ is the saturation partial pressure of water vapour over a plane of ice at temperature *T*. For the present study we have used the empirical formulae given by Sonntag (1994) for the two saturation pressure curves.

The formation of cirrus clouds is not yet completely understood. However, there is clear evidence that the nucleation of ice crystals requires relative humidities much larger than ice saturation. Heymsfield *et al.* (1998) have determined from many in-situ measurements in orographic wave clouds a critical relative humidity necessary for ice nucleation, $RH_{nuc}(T)$. This function decreases linearly from water saturation at temperatures higher than −39 °C to 75% at −55 °C. We assume that $RH_{nuc}(T)$ stays at the latter value at lower temperatures, although this is not evident from the measurements of Heymsfield *et al.* (1998) Thus we use:

$$RH_{nuc}(T) = 75\% + (T + 55\,°C) \cdot (25\%/16\,°C) \tag{3}$$

for −55 °C ≤ *T* ≤ −39 °C, and $RH_{nuc}(T) = 75\%$ for *T* < −55 °C.

3 Results

3.1 Statistical properties of relative humidity in the tropopause region

Figure 2 shows (non-normalised) frequency distributions of the relative humidity over liquid water, $N(RH)$, as measured during the 5962 MOZAIC flights. Two pairs of curves are displayed, one pair refers to measurements in the troposphere, the other one is based on stratospheric measurements. Both pairs show the respective distributions for each considered pressure range (i.e. 175–225 hPa and 225–275 hPa). The single *RH* values are binned into 100 1% wide classes (numbered from 0–99), where the last class (99) contains all data with *RH* ≥ 99%. The first nine bins are not shown, since the hygrometers do not work properly at the corresponding low vapour concentrations (validity
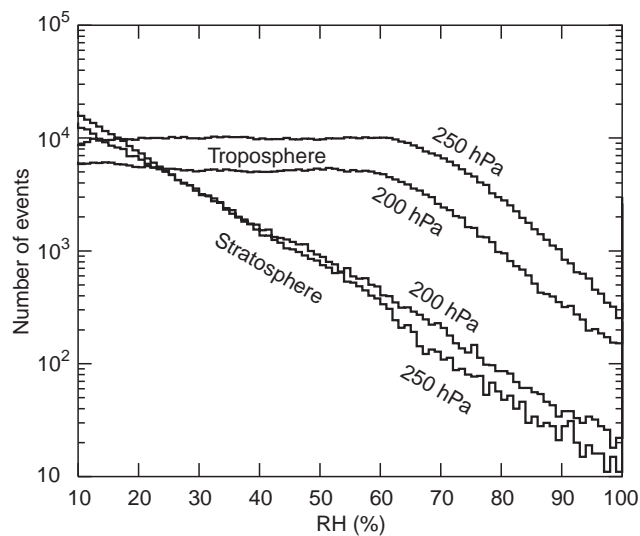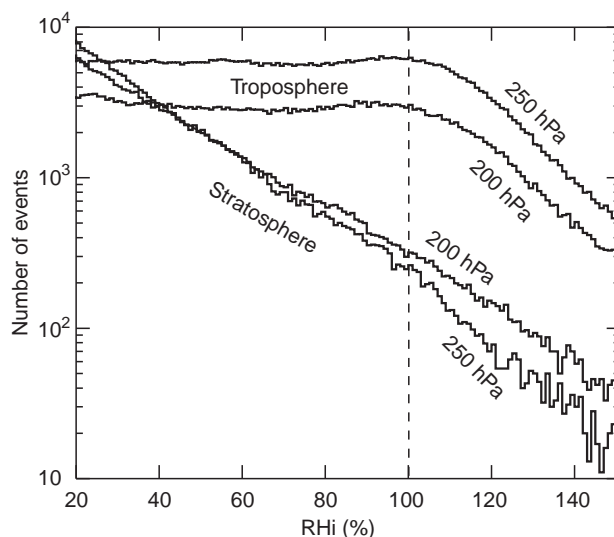

Fig. 1. Absolute uncertainty and precision of the MOZAIC humidity data as functions of relative humidity with respect to water

Fig. 2. Frequency distribution of relative humidity with respect to liquid water, *RH*, for stratospheric and tropospheric air in two layers centred around 200 and 250 hPa. The frequency distribution depicts the number of events per 1% bin

Fig. 3. Frequency distribution (events per 1% *RHi* bin) of relative humidity with respect to ice, *RHi*, for stratospheric and tropospheric air in two layers centred around 200 and 250 hPa

tag "2", see Sect. 1) and the distributions are very noisy there. The figure shows that the distributions of relative humidity *RH* possess a simple structure, in particular in the stratosphere where both distributions are modified geometric (i.e. $P_{RH}(k) = q^k(1-q); k \in \{0, 1, 2, \ldots\}$). (The geometric distribution refers to integer variables, e.g. bin numbers; if we consider the relative humidity a real variable, an exponential distribution will be appropriate.) The nature of the tropospheric humidity distribution is simple, too; but it will become clearer when we consider the corresponding distributions of the relative humidity with respect to ice. The tropospheric humidity distribution will be discussed in Sect. 4.2.

Figure 3 shows (non-normalised) frequency distributions of the relative humidity over ice, $N(RHi)$. The single *RHi* values are binned into 150 1% wide classes (numbers 0–149), where the last class (149) contains all data with $RHi \geq 149\%$. The first 19 bins are not shown. As before for *RH*, all distributions of *RHi* possess a simple structure. In the troposphere the distribution is approximately uniform below ice saturation, above ice saturation the distribution decreases geometrically. In the stratosphere the distributions are geometric throughout the range shown in the figures, but the distribution coefficient (q) seems to change at about ice saturation for both pressure levels. The exponents $b := -\ln q$ for distributions of both *RH* and *RHi* have been computed for several fit ranges using a straight line fit to the function $\ln[N(k)] = a - b \cdot k$. The resulting numbers are compiled in Tables 1 and 2. For the two stratospheric distributions, which can be fitted over the whole ranges from $10\% \leq RH \leq 99\%$ and $20\% \leq RHi \leq 149\%$, we can use the computed exponents b for the calculation of an average relative humidity over water and ice knowing that the expectation value of a modified geometric distribution is $q/(1-q) = e^{-b}/(1 - e^{-b})$. We find for the 200 hPa level mean values of $\langle RH \rangle = 13\%$ and

Table 1. Exponents b for the frequency distribution law $\ln[N(k)] \propto a - b \cdot k$ as obtained from straight line fitting to the distributions of relative humidity with respect to liquid water, *RH*, shown in Fig. 2. The standard deviations σ_b for the determined values of b are given in brackets. All these numbers have been multiplied by 100 for easier reading

Stratosphere/ Troposphere	p (hPa)	Fit range		
		75–98	10–60	10–98
Troposphere	200	10.0 (0.2)	0.3 (0.04)[a]	
Troposphere	250	11.8 (0.2)	−0.1 (0.03)[a]	
Stratosphere	200	7.4 (0.3)	7.1 (0.07)	7.4 (0.04)
Stratosphere	250	7.7 (0.4)	7.2 (0.04)	8.1 (0.06)

[a] Approximately uniform distribution with $b \approx 0$

Table 2. Exponents b for the frequency distribution law $\ln[N(k)] \propto a - b \cdot k$ as obtained from straight line fitting to the distributions of relative humidity with respect to ice, *RHi*, shown in Fig. 3. The standard deviations σ_b for the determined values of b are given in brackets. All these numbers have been multiplied by 100 for easier reading

Stratosphere/ Troposphere	p (hPa)	Fit range		
		119–148	20–100	20–148
Troposphere	200	5.8 (0.1)	0.1 (0.03)[a]	
Troposphere	250	6.2 (0.1)	0.0 (0.02)[a]	
Stratosphere	200	4.7 (0.2)	3.9 (0.3)	4.0 (0.2)
Stratosphere	250	5.0 (0.5)	4.1 (0.3)	4.6 (0.4)

[a] Approximately uniform distribution with $b \approx 0$

$\langle RHi \rangle = 25\%$. The corresponding numbers for the 250 hPa level are $\langle RH \rangle = 12\%$ and $\langle RHi \rangle = 21\%$. The standard deviations that correspond to the given mean values are of similar size. Thus they are considerably larger than the local instantaneous fluctuations of relative humidities around instantaneous mean values

within areas of about 300×300 km^2 that we have derived in our previous paper (Gierens *et al.*, 1997).

We consider the determined form of the humidity distribution law (i.e. the geometric or exponential distribution) for the stratosphere and for the ice-supersaturated regions in the troposphere the most important result of the present study. Although the actual values of the computed exponents might depend on the choice of the ozone-tropopause, the form of the distribution law for the ice-supersaturated regions is invariant, and that for the stratosphere retains the exponential form as long as the chosen threshold is not too low. The nearly uniform humidity distribution found for the ice-subsaturated regions in the troposphere is probably an artifact, see Sect. 4.2.

Overall, we find that 13.5% of the analysed data imply ice supersaturation, 0.3% even imply water supersaturation. Ice supersaturation is most frequent in autumn (21.8% of the data), frequent in winter (14.8%), and less frequent in spring and summer (8.7%). This seasonal distribution is probably influenced by the vertical shift of the mean tropopause during the year: in autumn there are the fewest flights occurring in the stratosphere (32.1%) whereas in spring more than half the data (52.5%) have been obtained in the stratosphere. (The actual fractions of stratospheric flights depends on our choice of the tropopause, but the sense of their annual cycle does not.) Since flight levels do not have a seasonal variation, it should be the tropopause height, that varies seasonally. As is well known, the mean tropopause height is larger in autumn than in spring at northern mid-latitudes (Hoinka, 1998).

It is interesting to note that there are ice supersaturated situations also in the stratosphere which usually is considered a very dry region (compare the mean values of *RH* and *RHi* given already). This is consistent with findings by Murphy *et al.* (1990) who found indications of ice supersaturation above but near the tropopause in the atmosphere above Stavanger (Norway, 59°N). Their result was obtained by comparison of frequency distributions of water vapour mixing ratio and saturation mixing ratio with respect to ice. These distributions overlapped, giving an indirect indication of ice supersaturation in the lower-most layers of the stratosphere. Here we present direct evidence of ice supersaturation in the stratosphere. Murphy *et al.* (1990) did not find ice supersaturation in the stratosphere over Punta Arenas (Chile, 53°S). Over Moffett Field (USA, 38°N) they found ice supersaturation only in winter, not in summer. Therefore it might be that the fraction of observations of ice supersaturation in stratospheric air in the MOZAIC data (about 2%) is an overestimation since the MOZAIC flight routes are mainly located in the northern latitudes where the results of Murphy *et al.* (1990) indicate a higher probability for ice supersaturation above the tropopause.

The derived fraction of ice supersaturation in the lowermost stratosphere might depend also on the choice of the ozone tropopause. However, at least for the longest distance flown in a stratospheric ice-supersaturated region the ozone tropopause criterion is consistent

with the meteorological situation: The longest distance in a stratospheric ice-supersaturated region was flown on a flight from Osaka (Japan) to Vienna (Austria) via the polar route on March 21, 1997. The 1400 km long path through the ice-supersaturated region was flown on the 205 hPa level along the coast of northern Siberia between about 80°E and 60°E (close to Novaya Zemlya island). The 200 hPa meteorological chart (not shown) shows that there is an extended low pressure region over the siberian part of the north Polar Ocean. The tropopause pressure along the flight path varies from 250 to 375 hPa. This shows that the large ice-supersaturated region was indeed in the stratosphere, consistent with the result obtained using the ozone tropopause criterion.

3.2. Properties of ice supersaturated regions in the troposphere

The results of the previous subsection can also be used to infer the mean ice supersaturation in supersaturated regions in the troposphere. In the 200 hPa layer this number amounts to 17%, whereas in the 250 hPa layer it is slightly lower, 16%.

We have determined from the MOZAIC data frequency distributions of temperature under various conditions, in particular in ice supersaturated regions. The resulting mean values, $\langle T \rangle$, and standard deviations, σ_T, are listed in Table 3. The results are similar for the whole dataset (i.e. stratosphere plus troposphere) and for tropospheric data alone. However, mean temperatures in ice supersaturated regions are 3 to 4 K lower than outside such regions. A statistical test (Bosch, 1996, ch 9.2) has proved this result to be highly significant with an error probability of less than 10^{-4}. In contrast to that, the standard deviations of T are similar in sub- and supersaturated regions. The colder air in supersaturated regions suggests that uplifting of air-masses is one major source of ice supersaturation in the tropopause region. Assuming a moist adiabatic lapse rate of 6 K/km, the uplift must occur over a distance of at least half a kilometre to result in a temperature depression of 3 to 4 K. Using the mean temperatures in tropospheric ice supersaturated regions from Table 3 and applying the

Table 3. Mean temperatures, $\langle T \rangle$, and standard deviations, σ_T, obtained from MOZAIC data under certain conditions. *IC* is an indicator variable that indicates whether the air is sub- ($IC = 0$) or supersaturated ($IC = 1$) with respect to ice

Stratosphere/ Troposphere	p (hPa)	*IC*	$\langle T \rangle$ (°C)	σ_T (°C)
S + T	200	0	−53.3	5.7
S + T	250	0	−49.0	6.8
S + T	200	1	−57.5	5.8
S + T	250	1	−52.1	6.9
T	200	0	−53.7	5.6
T	250	0	−48.0	7.1
T	200	1	−57.2	5.9
T	250	1	−51.9	6.9

formula of Heymsfield *et al.* (1998) for the calculation of RH_{nuc} (see Sect. 2) we find corresponding ice supersaturation for nucleation of 29% in the 200 hPa layer, and 31% in the 250 hPa layer. Both these values exceed the mean supersaturation in those regions by about 15%. This explains that an ice supersaturated region is not necessarily a region of cirrus cloud formation and it explains furthermore that one can often find persistent contrails (indicating ice supersaturation) where no other clouds are present.

Another possible origin of ice supersaturation in the upper troposphere is advection of moisture leading to variations of absolute vapour concentration. We have determined mean values of water vapour mixing ratio, $\langle q \rangle$, and their corresponding standard deviations, σ_q, from the MOZAIC data. The results are displayed in Table 4. We see that standard deviations are similar in size to the mean values. These relatively large fluctuations seem to be characteristic for the humidity field. For both types of frequency distributions used in the previous subsection, i.e. for uniform and geometric distributions, mean and standard deviations are of similar size; thus it is not surprising to find the same result here. The derived mean values show that the vapour mixing ratio is considerably larger (more than 50%) in ice supersaturated regions than outside. Although the "error-bars" (σ_q) are large and overlap substantially, a statistical test (Bosch, 1996, ch 9.2) gave the highly significant (error probability less than 10^{-4}) result that $\langle q \rangle_{IC=1} > \langle q \rangle_{IC=0}$. Obviously, moisture advection is another important origin of ice supersaturated regions.

4 Discussion

4.1 Possible causes and implications of the geometric humidity distribution

The finding of a distribution law for a certain quantity naturally leads to the question of what processes lead to the observed law. Since the data have been collected over a three year period at many different locations and under a multitude of various synoptic situations, the answer to the question can only be a system theoretical one. This means that we can find some general characteristics of the processes that cause the geometric humidity distribution, but we cannot determine details of these processes. The latter would require case studies of cases with similar synoptic situations. In spite of this

constraint, we nevertheless believe that knowledge of some general characteristics of the underlying processes is worthwhile.

We have devised a simple stochastic model that naturally leads to the observed geometric humidity distribution. The mathematical details can be found in the Appendix. Here we only present the essentials of the model and summarise the results.

We consider a physical system (e.g. an ice supersaturated region in the upper troposphere) that has reached a steady state and is not in a formation or termination phase. In the system there act various processes that can change the water vapour concentration. These processes consist of many single random events that may be assumed to be independent of each other (e.g. water molecules released from or condensed onto many different droplets or ice crystals). If the rates of the processes that enhance the vapour concentration are proportional to the rates of corresponding processes that diminish the vapour concentration (which implies that both types of processes depend in a similar way on the vapour concentration) and if the processes that remove water molecules from the vapour phase are slightly faster than the processes that supply molecules to the vapour phase (otherwise a steady state is not possible), then a geometric humidity distribution results. Diffusion of water molecules into and out of the system under consideration is an archetypical example of an appropriate process. Condensation and evaporation of droplets, as long it is diffusion-like, is another example. These are processes that likely act in supersaturation regions. In contrast, there is no geometric humidity distribution in tropospheric subsaturated regions probably because they do not reach a steady state. The lowermost stratosphere, which is the only part of the stratosphere encountered by MOZAIC flights, is affected by eddy transport effects, i.e. by turbulent diffusion (Holton *et al.*, 1995). This may explain the finding of a geometric humidity distribution in that part of the stratosphere.

When we devised the stochastic model we expected that the slope of the geometric distribution could give an indication of the residence time of water molecules in the systems under consideration. Our belief was soon realised to be a disillusion. Whereas the slope of the distribution depends on the ratio of the process rates, the residence time depends on their difference. In other words, there is no unique relationship between the residence time and the distribution of relative humidity in a system.

Table 4. Mean water vapour mixing ratios, $\langle q \rangle$, and standard deviations, σ_q, obtained from MOZAIC data from the troposphere. *IC* is an indicator variable that indicates whether the air is sub- ($IC = 0$) or supersaturated ($IC = 1$) with respect to ice

p (hPa)	*IC*	$\langle q \rangle$ (g/kg)	σ_q (g/kg)
200	0	0.044	0.037
250	0	0.081	0.081
200	1	0.072	0.059
250	1	0.125	0.124

4.2 The humidity distribution in subsaturated regions of the upper troposphere

The relative humidity in ice subsaturated regions in the troposphere turned out to follow approximately a uniform distribution. Relative humidities with respect to water up to 60% occur frequently whereas towards higher humidities the number of events decreases exponentially. The turnover point of 60% is obviously the mean *RH* at which ice saturation is reached in the

upper troposphere. This value (i.e. 60%) is consistent with results from SAGE II (Chiou *et al.*, 1997) and from Meteosat 6.3 μm measurements (van de Berg *et al.*, 1991) which both give a range of $5\% \leq RH \leq 60\%$ as typical for the upper-tropospheric humidity at clear sky. But the uniformity of the distribution for $RH < 60\%$ found from the MOZAIC data seems to be artificial and it is not consistent with the SAGE II results. If the data are discriminated according to whether they are of tropical/sub-tropical or mid-latitude/polar origin (with a boundary at 30°N), the uniformity of the distributions vanishes. Figure 4 shows that dryer cases dominate in the tropics and sub-tropics (compare Spencer and Braswell, 1997), whereas in the mid-latitudes and polar air masses moister cases dominate. This relation is also apparent from the SAGE II data (see Fig. 3a, b of Chiou *et al.*, 1997). The relative numbers of flights north and south of 30°N are such (about 2:1) that the addition of all these data result in an about uniform distribution of RH. However, we found that the geometric (exponential) character of the distributions in ice supersaturated regions is present on both sides of the 30°N demarcation line.

4.3 Possible biases

The MOZAIC instruments are installed on commercial airliners, not on research aircraft. The main task of commercial airlines is of course to transport people and freight, not to perform measurements in the atmosphere. The data sample that is evaluated here may therefore be biased in several respects due to habits of international aviation. For instance, pilots of North-Atlantic flights try to utilise or avoid the jet stream in order to save fuel and time. Properties of the jet stream zone are therefore
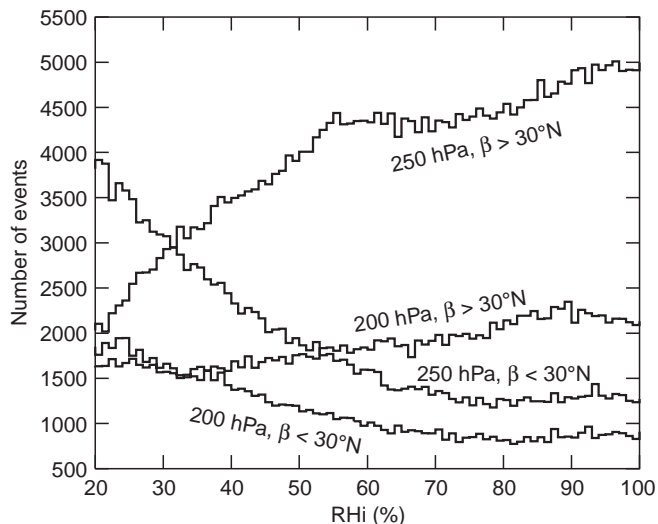
expected to have a relatively large weight in the data sample. Pilots also avoid flights through stormy regions and try to bypass them widely. This practice could lead to an underestimation of the true frequency of ice supersaturated regions in the upper troposphere since these are expected to appear often "over 100 km-wide regions downstream of cyclonic or convective storm systems" (Detwiler and Pratt, 1984).

The fractional area assumed by ice-supersaturated regions is certainly a function of altitude (or distance from the tropopause), latitude and longitude. A global picture of the spatial distribution of ice-supersaturated regions cannot be obtained from the MOZAIC data because of their incomplete spatial sampling. As a first guess of the global distribution of ice-supersaturated regions one might consider the distribution of "potential contrail cover" (regions where contrails would appear if there was air traffic; see Sausen *et al.*, 1998, their Fig. 2). The potential contrail cover given in that paper for air routes from Europe to North America is of the order 12–14% which is consistent with the fraction of data indicating ice supersaturation in the MOZAIC data. The potential contrail cover is much larger at higher altitudes in the tropics (100 and 150 hPa, i.e. altitudes that are never reached by the MOZAIC flights), generally exceeding values of 20%. This indicates, that ice supersaturation is more frequent or that ice-supersaturated regions are more extended in these tropical upper tropospheric regions than at northern middle latitudes.

5 Summary and conclusions

In the present work we have evaluated MOZAIC data from three years, 1995–1997, in order to determine statistical properties of the upper-tropospheric and lower-stratospheric relative humidity and temperature fields. The main results are:

1. The probability of measuring a certain amount of ice supersaturation in the troposphere decreases exponentially with the degree of ice supersaturation. Mean supersaturation in supersaturated regions amounts to about 15%, whereas ice nucleation would need on the average a supersaturation of about 30%. This explains the frequent observation of persistent contrails in otherwise clear air.

2. The probability of measuring a certain relative humidity in the stratosphere (both with respect to water and ice) decreases exponentially with the relative humidity. Mean values are $\langle RH \rangle \approx 12\%$ and $\langle RHi \rangle \approx 23\%$.

3. 13.5% of the analysed data imply ice supersaturation, in the stratosphere (presumably just above the tropopause) 2% of the data imply ice supersaturation. This value depends on the definition of the tropopause. Supersaturation with respect to water is negligible.

4. Ice supersaturated regions in the troposphere are 3–4 K colder than drier regions which indicates that uplifting of air-masses is one of the sources of ice supersaturation in the upper troposphere.

Fig. 4. Frequency distributions (events per 1% *RHi* bin) of relative humidity with respect to ice, *RHi*, in tropospheric subsaturated regions. The data are stratified according to latitude (tropical/sub-tropical, $\beta < 30°$N; mid-latitude/polar, $\beta > 30°$N) and pressure level (200 and 250 hPa)

5. In ice supersaturated regions in the troposphere the water vapour mixing ratio is more than 50% larger than outside which implies that moisture advection is another important source of supersaturation in the upper troposphere.

It should be noted that the MOZAIC data are collected on commercial flights that mainly connect European with North American cities. Hence the data do not provide a complete global picture, and the values quoted (fractional areas, mean supersaturation, etc.) could be different in particular in supersaturated regions beneath the tropical tropopause. However, the exponential character of the humidity distributions is probably valid globally, since it was preserved when the data were separated according to tropical/sub-tropical and polar/midlatitude origin.

We have related the found exponential (or geometric) distributions to a stochastic model of processing of water vapour in the systems under consideration. A geometric humidity distribution results when a system is in steady state and when the rate at which water molecules are added to the vapour phase is proportional to, but smaller than the rate at which water molecules are removed from the vapour phase. The processes must consist of single independent random events, hence diffusion is an appropriate process whereas advection is not (although advection is important during the formation phase of ice-supersaturated regions). Unfortunately, it is not possible to derive the average residence times of water molecules in the considered systems from the parameters of the measured humidity distributions since there is no unique relationship between the residence times and the parameters that determine the observed distributions. Efforts should be made to determine these unknown residence times.

The determined humidity distributions are simple enough to use them for tests of parametrisations of the hydrological cycle in large-scale models. Model data could be used to construct synthetic humidity distributions of the lower-most stratosphere. These should display an exponential character and should have an exponent value in the range of those values that have been determined here. A synthetic humidity distribution that does not follow an exponential curve would be an indication of severe flaws in the parametrisation. Exponents that differ significantly from the values derived here (for equal bin width of 1%) point at a wrong ratio of process rates, i.e. there are processes that are likely too fast or too slow compared with other processes.

We conclude with a philosophical remark. Our experience from many research flights was that the relative humidity is a quantity that displays a tremendously intricate structure in space and time. Now we find the amazing result that over longer time and larger space scales the relative humidity can be described by a very simple law that depends on merely one parameter. We gave arguments that this simple law is the result of many single independent events, that all may themselves be rather complicated. Our results, therefore, support the view that, in spite of complications in details, nature by and large exhibits a tendency for simplicity.

Appendix: a stochastic mechanism that leads to a geometric humidity distribution

We found the supersaturation with respect to ice ($Si = RHi - 100\%$) in the troposphere and the relative humidity with respect to both water and ice in the stratosphere (at least down to low values of RH and RHi) to obey a modified geometric distribution law. This distribution law reads:

$$P_X(k) = q^k(1-q); \quad k \in \{0, 1, 2, \ldots\} \ . \tag{4}$$

Here, X is one of the three quantities considered, measured in certain units, e.g. percent, i.e. $X \in \{\frac{RH}{\%}, \frac{RHi}{\%}, \frac{Si}{\%}\}$. If the modified geometric distribution type would not hold for humidities below a certain threshold RH_0, we could just replace RH or RHi with $RH - RH_0$ and $RHi - RHi_0$ and retain the above distribution type. The relation between the parameter q and the exponents b listed in the Tables 1 and 2 is simply:

$$q = e^{-b} \quad \text{or} \quad b = -\ln q \ . \tag{5}$$

We note that we have considered the quantity X an integer variable, which will prove useful for the subsequent derivations. If X is instead considered a real variable, the appropriate distribution law is a continuous exponential distribution $P_X(u) = b \cdot \exp(-bu)$; $u \in \mathbf{R}, u \geq 0$. In the following we consider X an integer variable, although we will sometimes speak of an exponential distribution. The expectation value of P_X is

$$E[X] = q/(1-q) \ , \tag{6}$$

which might be interpreted depending on the meaning of X as average ice supersaturation in supersaturated regions in the troposphere, or average relative humidity with respect to water, $\langle RH \rangle$, or ice, $\langle RHi \rangle$, in the stratosphere.

In the following we suggest a stochastic mechanism that would result in the observed modified geometric distribution. Let us assume a physical system (e.g. an ice supersaturated region in the upper troposphere) that contains water vapour, droplets, ice crystals, aerosols, etc. We are interested in the number of H_2O molecules in the vapour phase and term this (integer) number the "state of the system". Let us assume that there are processes that change the state from k to $k+1$ with a rate $\lambda(k)$ (s^{-1}). Such processes can include diffusion of water vapour, chemical production of H_2O molecules, or evaporation of drops and ice crystals. Let us further suppose that other processes change the state of the system from $k > 0$ to $k - 1$ with a rate $\mu(k)$. The latter processes could be diffusion, condensation, deliquescence of aerosols, or chemical reactions that consume water vapour. We assume that both the processes that enhance and diminish the number of H_2O molecules in the vapour phase have a random nature (therefore advection is not an appropriate process) and are of Poisson type, which means that the number of steps $k \to k \pm 1$ in an arbitrary time interval Δt is randomly distributed according to a Poisson law with means $\lambda \Delta t$ for upward steps and $\mu \Delta t$ for downward steps, respectively. The rates λ and μ must depend on the state k in a similar way: $\mu(k+1) = c\lambda(k)$, where c is a constant. The simplest case is given when the rates are both independent of the state k. Although this is not a realistic assumption for real physical systems we will use it here for the sake of simplicity; we note that the results obtained are still valid in the general case. Of course, for an empty system, i.e. $k = 0$, the rate $\mu(0)$ vanishes. For the chosen stochastic model it is further required that the ratio of the rates, $\varrho := \lambda/\mu$, must be strictly less than one, otherwise the system cannot reach a steady state, and there would be no expectation value for the state

of the system, in contradiction to our results. With these assumptions the system behaves like a M/M/1-queue in queueing theory (see Goodman, 1988), and it is allowed to apply the results of this theory here.

The proportion of time the system is in the state k (i.e. $k \leq X < k + 1$), is given by:

$$p_k = \varrho^k (1 - \varrho) \ . \tag{7}$$

Obviously, this fraction of time is identical to the relative frequency (probability) of measurements that give $N = k$ where N is the number of H_2O molecules in the vapour phase contained in the system:

$$p_k \equiv P_N(k) \ . \tag{8}$$

This probability law has already the form of a modified geometric distribution.

Now we change the point of view: let a certain number of n H_2O molecules be equivalent to a relative humidity increment of 1%. (Certainly, n is a *very* large integer.) Then we can ask for the probability, $P_X(m)$, that the relative humidity in the system is m%. A relative humidity of m% is obviously equivalent to a collection of states of the system in the range $mn \leq k < (m+1)n$. Thus we have:

$$P_X(m) = P_N(mn + 0) \ + \ P_N(mn + 1) + \cdots + P_N(mn + n - 1)$$

$$= \sum_{i=0}^{n-1} P_N(mn + i)$$

$$= (1 - \varrho) \varrho^{mn} \sum_{i=0}^{n-1} \varrho^i$$

$$= (1 - \varrho) \varrho^{mn} \frac{1 - \varrho^n}{1 - \varrho}$$

$$= \varrho^{mn} (1 - \varrho^n)$$

$$= q^m (1 - q) \ , \tag{9}$$

where we have introduced the abbreviation "q" for ϱ^n. This shows that the stochastic model given leads to the observed type of a modified geometric distribution for the relative humidity. A similar construction can be used to derive the geometric distribution for the supersaturation in ice supersaturated regions.

References

Bethan, S., G. Vaughan, and S. J. Reid, A comparison of ozone and thermal tropopause heights and the impact of tropopause definition on quantifying the ozone content of the troposphere, *Q. J. R. Meteorol. Soc.*, **122,** 929–944, 1996.

Bosch, K., *Großes Lehrbuch der Statistik*, R. Oldenbourg, München, Germany, 586 pp, 1996.

Chiou, E. W., M. P. McCormick, and W. P. Chu, Global water vapor distributions in the stratosphere and upper troposphere derived from 5.5 years of SAGE II observations (1986–1991), *J. Geophys. Res.*, **102,** 19 105–19 118, 1997.

Detwiler, A., and R. Pratt, Clear air seeding: opportunities and strategies, *J. Weather Modification*, **16,** 46–60, 1984.

Duhnke, K., H. G. J. Smit, D. Kley, A. Marenco, P. Speth, and J. Wefers, Untersuchung der in MOZAIC gemessenen Ozon- und Wasserdampfverteilung im polaren Strahlstrom über dem Nordatlantik, *Jül-Ber.* **3593,** Jülich, Germany, 1998.

Gierens, K., Numerical simulations of persistent contrails, *J. Atmos. Sci.*, **53,** 3333–3348, 1996.

Gierens, K. M., U. Schumann, H. G. J. Smit, M. Helten, and G. Zängl, Determination of humidity and temperature fluctuations based on MOZAIC data and parametrisation of persistent contrail coverage for general circulation models, *Ann. Geophysicae*, **15,** 1057–1066, 1997.

Goodman, R., *Introduction to stochastic models*, Benjamin/Cummings, Menlo Park, California, 368 pp, 1988.

Helten, M., H. G. J. Smit, W. Sträter, D. Kley, P. Nedelec, M. Zöger, and R. Busen, Calibration and performance of automatic compact instrumentation for the measurement of relative humidity from passenger aircraft, *J. Geophys. Res.*, **103,** 25 643–25 652, 1998.

Heymsfield, A. J., L. M. Miloshevich, C. Twohy, G. Sachse, and S. Oltmans, Upper-tropospheric relative humidity observations and implications for cirrus ice nucleation, *Geophys. Res. Lett.*, **25,** 1343–1346, 1998.

Hoinka, K. P., Statistics of the global tropopause pressure, *Mon. Weather Rev.*, **126,** 3303–3325, 1998.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister, Stratosphere-troposphere exchange, *Rev. Geophys.*, **33,** 403–439, 1995.

IPCC, *Climate change: the IPCC scientific assessment*, Cambridge University Press, Cambridge, UK, 365 pp, 1990.

Lindzen, R. S., Some coolness concerning global warming, *Bull. Am. Meteorol. Soc.*, **71,** 288–299, 1990.

Manabe, S., and R. T. Wetherald, The thermal equilibrium of the atmosphere with a given distribution of relative humidity, *J. Atmos. Sci.*, **24,** 241–259, 1967.

Marenco, A., V. Thouret, P. Nedelec, H. G. Smit, M. Helten, D. Kley, F. Karcher, P. Simon, K. Law, J. Pyle, G. Poschmann, R. Von Wrede, C. Hume, and T. Cook, Measurement of ozone and water vapour by Airbus in-service aircraft: the MOZAIC airborne program, an overview, *J. Geophys. Res.*, **103,** 25 631–25 242, 1998.

Murphy, D. M., K. K. Kelly, A. F. Tuck, M. H. Proffitt, and S. Kinne, Ice saturation at the tropopause observed from the ER-2 aircraft, *Geophys. Res. Lett.*, **17,** 353–356, (March Supplement), 1990.

Pierrehumbert, R. T., Thermostats, radiator fins and local runaway greenhouse, *J. Atmos. Sci.*, **52,** 1784–1806, 1994.

Rind, D., E. W. Chiou, W. Chu, J. Larsen, S. Oltmans, J. Lerner, M. P. McCormick, and L. McMaster, Positive water vapour feedback in climate models confirmed by satellite data, *Nature*, **349,** 500–503, 1991.

Sausen, R., K. Gierens, M. Ponater, and U. Schumann, A diagnostic study of the global distribution of contrails, Part I: present-day climate. *Theor. Appl. Climatol.*, **61,** 127–141, (1998).

Schumann, U., On conditions for contrail formation from aircraft exhausts, *Meteorol. Z.*, N.F., **5,** 4–23, 1996.

Sonntag, D., Advancements in the field of hygrometry, *Meteorol. Z.*, N.F., **3,** 51–66, 1994.

Spencer, R. W., and W. D. Braswell, How dry is the tropical free troposphere? Implications for global warming theory, *Bull. Am. Meteorol. Soc.*, **78,** 1097–1106, 1997.

van de Berg, L., A. Pyomjamsri, and J. Schmetz, Monthly mean upper tropospheric humidities in cloud-free areas from METEOSAT observations, *Int. J. Climatol.*, **11,** 819–826, 1991.