Emerging pattern of global change in the upper atmosphere and ionosphere

J. Laštovička¹, R. A. Akmæv², G. Beig³, J. Bremer⁴, J. T. Emmert⁵, C. Jacobi⁶, M. J. Jarvis⁷, G. Nedoluha⁵, Yu. I. Portnyagin⁸, and T. Ulich⁹

¹Institute of Atmospheric Physics, Bocni II, 14131 Prague, Czech Republic
²Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA
³Indian Institute of Tropical Meteorology, Pune-411 008, India
⁴Leibnitz-Institute of Atmospheric Physics, Schloss-Street 6, 18225, Kühlungsborn, Germany
⁵Naval Research Laboratory, Washington, D.C. 20375, USA
⁶Institute for Meteorology, University of Leipzig, Stephanstr. 3, 04103 Leipzig, Germany
⁷British Antarctic Survey, Cambridge, CB3 0ET, UK
⁸Institute for Experimental Meteorology, Lenin Str. 82, Obninsk 249038, Russia
⁹Sodankylä Geophysical Observatory, Tähteläntie 62, 99600 Sodankylä, Finland

Received: 11 May 2007 – Revised: 4 October 2007 – Accepted: 4 October 2007 – Published: 28 May 2008

Abstract. In the upper atmosphere, greenhouse gases produce a cooling effect, instead of a warming effect. Increases in greenhouse gas concentrations are expected to induce substantial changes in the mesosphere, thermosphere, and ionosphere, including a thermal contraction of these layers. In this article we construct for the first time a pattern of the observed long-term global change in the upper atmosphere, based on trend studies of various parameters. The picture we obtain is qualitative, and contains several gaps and a few discrepancies, but the overall pattern of observed long-term changes throughout the upper atmosphere is consistent with model predictions of the effect of greenhouse gas increases. Together with the large body of lower atmospheric trend research, our synthesis indicates that anthropogenic emissions of greenhouse gases are affecting the atmosphere at nearly all altitudes between ground and space.

Keywords. Atmospheric composition and structure (Thermosphere – composition and chemistry; Evolution of the atmosphere) – Ionosphere (Ionosphere-atmosphere interactions)

1 Introduction

Life on Earth is more directly affected by climate change near the surface than in the middle and upper atmosphere, but as the story of the Earth’s ozone layer illustrates, changes at higher levels of the atmosphere may be important for life on Earth, as well. Moreover, with society becoming more dependent on space-based technologies, it is important to understand the long-term evolution of the upper atmosphere.

The upper atmosphere consists of the mesosphere (~50–90 km), thermosphere (~90–800 km), and its ionised part, the ionosphere, which is embedded within these regions. The thermosphere is the operating environment of many communication and weather satellites, as well as the International Space Station, and knowledge of the atmospheric drag exerted on these objects is necessary for accurate prediction of their location. The structure of the ionosphere directly affects the performance of the Global Positioning System (GPS) and other satellite navigational systems, whose signals are phase-shifted by ionospheric plasma. The mesosphere critically affects lower atmosphere-thermosphere coupling via atmospheric waves and electromagnetic phenomena; this complex region is less understood than the regions above and below.

The 0.6°C increase in global surface air temperature during the twentieth century (e.g. IPCC, 2001) has been attributed predominantly to the increasing atmospheric concentration of greenhouse gases. In the upper atmosphere, the radiative effects of greenhouse gases, particularly CO₂, become more pronounced and produce a cooling, rather than a warming effect. At these altitudes, CO₂ is optically thin and is unable to contain outgoing infrared radiation; thermal energy is transferred by collisions with ambient gas to the excited states of CO₂ molecules and then lost to space via its infrared radiation. A much stronger greenhouse effect is observed on Venus where the 96% concentration of carbon...
dioxide in the atmosphere results in a troposphere that is more than twice as warm as Earth’s and a thermosphere that is 4–5 times cooler (e.g. Bougher and Roble, 1991).

The same relative increases in greenhouse gas concentrations observed near the surface are expected to occur in the upper atmosphere. Consequently, the chemical composition and thermal structure of the upper atmosphere may be altered substantially by human activities, which, through feedbacks, could potentially affect the lower atmosphere. As with global change near the Earth’s surface, the challenge facing upper atmospheric climate scientists is to detect long-term trends and attribute them to their primary causes, so that society can mitigate against harmful change scenarios.

Roble and Dickinson (1989) and Rishbeth and Roble (1992) presented the first theoretical simulations of the effects of hypothetical future increases in greenhouse gas concentrations on the upper atmosphere and ionosphere. These studies predicted substantial temperature reductions in the upper atmosphere with doubled CO$_2$, an effect that was termed “greenhouse cooling” (Cicerone, 1990). The global atmosphere is very nearly in a hydrostatic equilibrium, which means that the height of a given pressure surface is determined by the average atmospheric temperature below. The cooling is therefore expected to result in thermal contraction of the upper atmosphere and we may expect a significant decline in thermospheric density at fixed heights (not necessarily at fixed pressure levels), as well as a downward displacement of ionospheric layers (Rishbeth, 1990). Mesospheric behaviour might be affected also by the stratospheric ozone depletion (e.g. Bremer and Berger, 2002; Akmaev et al., 2006) observed over the last three decades (e.g. Staehelin et al., 2001).

Information about long-term trends in various upper atmospheric parameters has accumulated to a level such that a coherent pattern of upper atmospheric climate change is beginning to emerge. In this article, we present the first comprehensive assessment of observed long-term changes and trends in the mesosphere, thermosphere, and ionosphere. We identify gaps in our knowledge and understanding of long-term trends in various parameters, as well as possible inconsistencies among trends or their substantial changes with time, and we investigate whether the observed trends are consistent with the theoretically predicted effects of increasing greenhouse gas concentrations. International collaboration on upper atmospheric trend studies is realized via the IAGA/ICMA Working Group “Long-term Trends in the Mesosphere, Thermosphere and Ionosphere” and the SCOSTEP/CAWSES Working Group “Climatological Variations of the Ionosphere and Upper Atmosphere”, and leading scientists from both groups have contributed to this article.

2 Difficulties – methodology, data uncertainties, natural variability

Important points to be discussed in relation to long-term trend studies include various uncertainties in data, the long-term consistency of measurements, and effects of applications of different methods of trend determination. These issues are particularly important for detecting weak trends in noisy time series. Different methods have been developed for extracting long-term trends from particular data sets. These methods often include averaging and interpolation procedures, and thus it is very difficult to compare trend estimates obtained by different authors. Ulich et al. (2003) discussed some practical problems of determining long-term change.

Long time series of measurements need to be consistent. Instrumental changes and malfunctions need to be well documented and corrected for, and subsequent instruments must be well inter-calibrated with previous ones. Before attempting to compute any trend, one has to verify whether or not the time series in question includes instrumental effects. Furthermore, historic time series usually contain data gaps, which have to be treated with care. Some analysis methods require continuous data sets with constant sampling frequency, but interpolation can introduce incorrect data points, which might skew the results of the analysis.

Studies of the F2-layer peak height, $hmF2$, suffer from an additional complication. The height cannot be directly scaled from ionograms; it is computed from the critical frequencies of the E and F2 layer as well as the propagation parameter M(3000)F2 using empirical formulae. These formulae were derived in the 1970s based on various regional data sets and to be simple and straightforward enough to be used by the computers of the time, but these formulae are by no means globally valid, and their use results in different trends. One has to check carefully which one is applicable to the ionosonde in question.

Atmospheric time series usually vary on diurnal and seasonal time scales. In order to see beyond this variability, cyclic influences on the measurements need to be properly removed and/or considered in interpretation of results. This is usually done by fitting multi-parameter models to the data. Care must be taken that these cyclic parameters are adequately removed and/or considered, especially if the diurnal and seasonal sampling is not uniform.

Solar and geomagnetic activity can have a crucial impact on trend determination when data series are relatively short or when trends in the ionosphere or thermosphere are being estimated. However, the strong 11-year solar cycle can often be suppressed or minimized with proper selection of the analysed period, with application of corrections to solar activity, or by studying the residuals from empirical models based upon solar and geomagnetic inputs, season, latitude and time. Solar activity changes are usually represented by the F10.7 and E10.7 indices, or by sunspot numbers. Estimated trends can be sensitive to the choice of solar activity...
proxy; for example, the use of sunspot numbers instead of F10.7 or E10.7 produces less reliable trends for F2-layer parameters (Laštovička et al., 2006).

The temporal extent necessary to obtain reliable trend estimates from a data set is not easy to determine (Weatherhead et al., 2002). When studying a time series beginning at, say, solar maximum and ending at solar minimum, even if it covers several solar cycles, residual solar cycle effects affect the estimated trend. In this case, if the trend estimate is plotted as a function of the length of the time series it is derived from, a damped oscillation results, the decay rate of which can be used to estimate the number of years of data required to produce a reliable trend (Jarvis et al., 2002; Clilverd et al., 2003).

One relatively little known factor, which affects the magnitude of determined trends, is data smoothing, e.g., in the form of running averages, which may artificially improve the statistical significance of results if the reduction in the number of degrees of freedom is not accounted for properly; it also tends to reduce the magnitude of trends, as illustrated by the behaviour of trends in foF2 (Sect. 5).

Comparison of the results of different methods of trend determination applied to the same data set of foF2 values showed good agreement among some methods, but substantially different results were obtained by other methods (Laštovička et al., 2006; Sect. 5). This is the case for trends that are much weaker than overlapping natural variations. On the other hand, strong trends appear to be essentially method-independent, as documented by Križan and Laštovička (2005) for laminae in ozone profiles.

Long-term changes of solar activity and/or an increase of geomagnetic activity throughout the twentieth century (e.g. Stamper et al., 1999) might play an important role in the observed long-term trends. The role of solar and geomagnetic activity in the observed long-term trends increases with altitude from troposphere to ionosphere and is greater in the ionised component than in the neutral atmosphere. However, it should be noted that the increase of solar influence with altitude is not monotonic; for instance the solar effect in the mesopause region is relatively small (Matthes et al., 2004), making it easier to study long-term trends in this region. During the 20th century, the continuous increase in greenhouse gas concentration, and possibly other anthropogenic influences, combined with a weakening upward trend in solar and geomagnetic activity, has resulted in an increasing proportion of overall upper atmospheric trends being due to human activity relative to that due to changes of solar origin (Laštovička, 2005). For example, a change from dominant geomagnetic to dominant non-geomagnetic forcing of the long-term trend in the E-region ionosphere (~110 km altitude) was found around 1970 (Mikhailov and de la Morena, 2003). At present the geomagnetic activity may play an important or even dominant role only in foF2 and hmF2 trends.

### 3 What does the term “trend” mean?

Not all authors understand the term “trend” in the same way. Most authors understand the term “trend” as a “long-term trend”, i.e. as a long-term tendency of essentially monotonic linear/quasi-linear change, either an increase or decrease of the values of the studied variable with time. If the long-term behaviour is substantially non-monotonic or oscillatory, the term “long-term change” is used. Strictly speaking, trends are often not quite linear. However, in many cases the linear trend approximation is sufficient, and in other cases we use the linear approximation for easier comparison with trends in other parameters.

A linear trend cannot continue indefinitely and in reality a quasi-stable trend should begin and end within a certain period of time. This is most evident with the Antarctic ozone hole, which first appeared in the late 1970s, is apparently peaking at present and is expected to disappear in the second half of this century. This behaviour was caused by the emissions of man-made ozone-depleting substances and their subsequent ban by the Montreal Protocol and its amendments, modulated by an increasing greenhouse effect. Such behaviour may be characterized by a sophisticated mathematical curve, but may often be approximated by the so-called piecewise contiguous linear trend model (Weatherhead et al., 2000; Reinsel et al., 2005), wherein essentially different linear trends are fit to the data within different time intervals, the latter also being model parameters. The upper atmospheric and ionospheric linear trends should probably be understood in a similar way. A different understanding of the term “trend” may lead to some of the contradictions referred to in Introduction.

A few authors use a slightly different and more mathematical approach. They use the fact that a true linear trend is located at infinite period and is not satisfactorily taken into account in a spectral decomposition. Assuming the trend is not a small temporal part of a major very long period cycle or due to some temporally-coincident stochastic change, then it is possible to temporally and spectrally decompose the time series to separate the oscillatory and true trend elements using, for instance, the wavelet transform technique. However, practical application of such an approach has to be made with care, as the trend necessarily has a non-infinite period limited by the length of the dataset or the length of the interval of the stationary linear trend existence.

### 4 Consistent pattern of long-term trends

The temperature in the mesosphere and lower thermosphere (MLT) is directly affected by changing concentration of greenhouse gases. Mesospheric temperature trends may be inferred from temperature measurements made by ground-based, in-situ (rocketsonde), and satellite instruments. Trends can also be estimated using indirect methods.
Beig et al. (2003) provided a first comprehensive review of observational and model results for temperature trends in the mesosphere (50–80 km) and mesopause region (80–100 km). The comparison of the results obtained by different observations distributed over several decades is complicated. Nevertheless, there are a number of occasions where the temperature trend results indicate consistency and some of the differences are explainable. Beig (2006) recently strengthened the conclusions of Beig et al. (2003) with the addition of newly reported results (e.g. Remsberg and Deaver, 2005).

The present status of mesospheric temperature trends based on the available measurements can be broadly described as follows: (1) 2–3 K/decade decreases in lower and middle mesosphere (50–70 km) temperatures, with the strongest trends occurring at tropical latitudes. (2) In the upper mesosphere (70–80 km), uncertain trend estimates ranging from 0 to −10 K/decade, with the large trends being essentially based on rocketsonde data. Rocket sensors may involve significant uncertainties if not properly corrected, therefore large rocket-based trends must be considered with great care. On the other hand, a few zero-trend results are associated with local/regional effects or seasonal sampling at high latitudes.

There are a growing number of experimental results consistent with a zero temperature trend in the mesopause region (80–100 km). The most reliable data sets show no significant trend. Seasonal analyses are not numerous below the mesopause. The seasonal variation around the mesopause reveals different trends in winter and summer. It may explain some of the diverse results reported so far. Offermann et al. (2004) show a systematic variation in the long-term temperature trends from month to month in the mesopause region. French and Burns (2004) have recently found a dramatic seasonality in the trends in the mesopause region temperature at higher latitudes of the Southern Hemisphere.

Trends in mesospheric temperatures have been subject of model investigations, as well. Roble and Dickinson (1989) predicted a global average cooling by about 10–15 K for the doubled-CO$_2$ scenario. Thereafter, several model results were reported as reviewed in Beig et al. (2003). These studies predicted a general cooling trend in the MLT region and attributed it to the greenhouse gas increase in the atmosphere. These studies suggest that global change effects are significant throughout the mesosphere. The troposphere should expand slightly as it is warmed (e.g. Seidel and Randel, 2006), but the upper atmosphere should contract much more due to the greater cooling. Although many model studies have considered only the doubled-CO$_2$ scenario, they have shown reasonably good qualitative agreement with observations, including both vertical and latitudinal structures of temperature trends. Three-dimensional models (Akmaev, 2002; Schmidt et al., 2006; Fomichev et al., 2007) indicate a decrease in cooling toward the summer pole in the upper mesosphere. The simulated cooling trend in the middle mesosphere produced only by the CO$_2$ increase is usually below the observed level. However, including other greenhouse gases and taking into account the “thermal shrinking” of the upper atmosphere results in a cooling close to the lower limit of observations. The inclusion of sea surface temperature (SST) changes into simulations does not substantially affect the middle atmospheric temperature with the exception of the summer mesopause where additional reduction in cooling or even heating is found (Schmidt et al., 2006; Fomichev et al., 2007). In the mesopause region, model simulations produce trends usually below 1 K/decade that appear to be consistent with most observations in this region. Recent model results of Gruzdev and Brasseur (2005), which include possible effects of changing gravity wave activity, do not predict any significant temperature change in the 92–95 km layer.

Stratospheric ozone depletion (e.g. Staehelin et al., 2001), resulting in weaker ozone heating, may have a noticeable impact on upper-atmospheric temperatures and densities as well (Akmaev et al., 2006). Bremer and Berger (2002) have shown that ozone depletion has contributed noticeably to the overall cooling and contraction of the mesosphere. Long-term changes of the mesospheric ozone itself are largely unknown. Recent simulations by Akmaev et al. (2006) have found that the effect of the stratospheric ozone depletion on total mass density, which is affected by cooling-induced contraction of the upper atmosphere, peaks near 110 km and also extends well into the thermosphere, adding about −0.5%/decade to the overall negative trend estimate near 200 km (Fig. 1).

No direct observational information on neutral temperature trends is available for altitudes above 100 km due to a shortage of data, and information deduced from other measurements is scarce. The first study of trends in ion temperatures at heights near 350 km has recently been presented for Millstone Hill (42.6° N, 288.5° E) by Zhang et al. (2005). They observed a trend of about −17 K per decade over the period 1975–2005.
Bencze (2005) used the difference $h m F_2 - h' F$ (F layer half-thickness) as a parameter for studying the trend in the average F-layer temperature. He found a negative trend in this parameter, which suggests thermal contraction of the F layer, as expected as a consequence of greenhouse cooling. Bencze (2005) interpreted the trend for solar minimum years as a consequence of changing solar activity in solar cycle minima with a partial contribution from the greenhouse gas effect and an insignificant contribution from changes of geomagnetic activity.

As already mentioned in relation to Fig. 1, changes in atmospheric temperature have resulted in changes in atmospheric density. At thermospheric altitudes between ~200 and 800 km, atmospheric drag causes measurable decay of the orbits of satellites and space debris, and routine tracking observations can be used to derive long-term changes in thermospheric density. The results of published trend analyses of these data are summarized in Fig. 2; after accounting for known natural variability, all three studies (Keating et al., 2000; Emmert et al., 2004; Marcos et al., 2005) indicate that, at a given height, thermospheric density has declined during the past several decades, at a rate of 1.7–7.2% per decade. The results of the three studies are quite consistent; some of the small differences among the results can be attributed to differences in the heights and time periods sampled in each study. The density trends increase with height and are stronger for solar minimum conditions (Emmert et al., 2004); such behaviour is qualitatively consistent with model predictions of thermal contraction associated with increased greenhouse gas concentrations (e.g. Roble and Dickinson, 1989; Rishbeth and Roble, 1992; see also Sect. 6). Figure 2 also reveals good agreement between observational trends and model estimates; observational trends are located within the range of model estimates (Roble and Dickinson, 1989; Rishbeth and Roble, 1992; Akmaev et al., 2006; Qian et al., 2006).

In the orbital drag studies, empirical thermospheric models were used, prior to computing trends, to filter out the effects of solar extreme ultra-violet, or EUV, irradiance (the dominant source of long-term variability), as well as spatial, seasonal, and geomagnetic activity effects. Overall, the empirical models do a good job of filtering solar EUV effects, generally reducing the long-term EUV-dependent variation to within 5% (Emmert et al., 2004; Marcos et al., 2005). However, during solar minimum, recent orbit-derived density measurements show a strong EUV-dependent bias relative to the models; this bias is characterized by a sharp drop-off of the measured values at very low solar activity levels (Emmert et al., 2004; Marcos et al., 2005), and is possibly a limitation of the 10.7 cm solar radio flux used as a proxy for EUV irradiance. Despite this uncertainty in the specification of natural EUV-induced variability, the trend estimates appear to be quite robust, particularly since similar trends are obtained across all phases of the solar cycle. Other potential sources of error in the trend analyses have also been examined, but none can account for the existence of the negative trends (Emmert et al., 2004).

Long-term changes in the neutral atmosphere have to affect the ionosphere. Laštovička and Bremer (2004) reviewed long-term trends in the lower ionosphere (below 120 km) using various datasets obtained by several different methods of measurements and modelling. Below about 85–90 km, all data provide a positive trend in electron densities at fixed heights, in qualitative agreement with the cooling and thermal contraction of the mesosphere. Above 90 km, rocket data indicate a negative trend in electron densities at fixed heights, whereas various ground-based measurements show a contradictory positive trend. Rocket data, unlike ground-based
data, are very scarce and irregularly distributed in space and time; we therefore believe that the positive ‘ground-based’ trend is more reliable. For ions, enough observational data to study trends are not available. For the $2\times CO_2$ scenario, models (e.g. Beig, 2000) predict much more dramatic changes in the ion composition than in the total ion density, which is equal to the electron density above $\sim 80\, km$; for example, a dramatic reduction of NO$^+$ concentration is expected.

Analyses for the ionospheric E-layer parameters $foE$ and $h'E$ have been carried out for European stations (Bremer, 1998) as well as globally distributed ionosonde stations (Bremer, 2001; Bremer et al., 2004). The mean global trend for $foE$ derived from observations at 72 stations is $+0.0014\, MHz/year$ and is significantly different from zero, as shown in the lower part of Fig. 3. Here the older results (Bremer, 2001; Bremer et al., 2004) have been updated to include new data up to 2005. The mean global trend for $h'E$ is negative, with a significance level greater than 95% based on the observations at 33 stations (middle part of Fig. 3). The trend results (positive $foE$ and negative $h'E$ trends) are in qualitative agreement with model predictions of the increased greenhouse effect (Rishbeth and Roble, 1992), but the experimental trends are markedly stronger than the model predictions. The negative trend in $h'E$ is a consequence of thermal contraction of the mesosphere and lower thermosphere. The positive $foE$ trend is consistent with rocket mass spectrometer measurements of the ion density ratio [NO$^+$/O$_2^+$] in the E region (Danilov, 1997). The observed negative trends of [NO$^+$/O$_2^+$] cause increasing electron densities and thus increasing $foE$, because the dissociative recombination coefficient of NO$^+$ is greater than that of O$_2^+$. Acceleration of trends in $foE$ and ionospheric radio wave absorption after $\sim 1980$ supports their anthropogenic origin (Bremer, 2005). However, Mikhailov and de la Morena (2003) found $foE$ trends to be geomagnetically controlled before about 1970. This is consistent with Laštovička’s (2005) conclusion about the decreasing role of solar and geomagnetic activity from the beginning to the end of the 20th century. Bremer (2006) observed for a couple of stations that after eliminating the solar cycle effect, $\Delta foE$ and $\Delta O_3$ display a long-term negative correlation $r=-0.57$, which is consistent with model predictions of the role of ozone depletion in long-term trends in the upper atmosphere and ionosphere. According to Fig. 1 the ozone effect in the neutral density maximizes near the height of E region maximum. Omission of ozone effects contributes to the difference between model results of Rishbeth and Roble (1992) and observational results.

In the F1 layer, most $foF1$ data series ($h'\sim 200\, km$) from ionosonde observations over Europe have positive trends after elimination of the solar and geomagnetically induced parts (Bremer, 1998). This finding has been confirmed by analyses of long-term measurements at 51 worldwide-distributed ionosonde stations (Bremer, 2001; Bremer et al., 2004). Using an updated data set the derived mean global trend of $+0.0019\, MHz/year$ (upper part of Fig. 3) is significantly different from zero and in agreement with the model results of Rishbeth and Roble (1992) derived for an increasing greenhouse effect (Fig. 1 shows that ozone effect near 200 km is rather weak).

Summarizing, we may say that trends in the mesospheric temperature, electron concentration in the lower ionosphere, electron concentration and height of its maximum in the E region, electron concentration in the F1-region maximum, the thermospheric neutral density and ion temperature – all these trends are mutually consistent and qualitatively agree with model simulations of consequences of the enhanced greenhouse effect resulting in global cooling and contracting of the upper atmosphere. They form a consistent pattern of global change in the upper atmosphere.

Most of the observed trends are stronger than expected from initial models of Roble and Dickinson (1989) and Rishbeth and Roble (1992). Part of the difference can be attributed to omission of the effect of stratospheric ozone depletion in these early models; its inclusion significantly improves the agreement between models and measurements (e.g. Bremer and Berger, 2002). Uncertainties in model results arise also from the inadequate simulation of some
processes, or from uncertainty in the specification of model drivers, e.g. height profiles of CO\textsubscript{2} and other greenhouse gases. Observational information on trends in important minor constituents such as ozone and nitric oxide in the MLT region is essentially missing, and only scarce model predictions are available. On the other hand, the degree of quantitative uncertainty in some observational trends compared with model results is also relatively high. As shown in Fig. 2 for thermospheric neutral density, models of Roble and Dickinson (1989), Rishbeth and Roble (1992), and Qian et al. (2006) underestimate the observed trends, while the most recent model by Akmaev et al. (2006) apparently overestimates the observed trends. Trends, especially in the ionospheric parameters, have mostly been compared with the model predictions by Rishbeth and Roble (1992) for the 2×CO\textsubscript{2} scenario and, therefore, some quantitative disagreement is to be expected.

5 Non-stationary trends, discrepancies and missing information

There are three groups of parameters whose trends cannot be incorporated into the above global pattern due to discrepancies between results of various authors, possible changes in the direction of trends over time, or missing information. The key words of these three groups are dynamics, water vapour, and the F2 region of the ionosphere. Let us start with the first group, dynamics.

Changes in temperature and minor constituents have to affect winds. Routine wind measurements at the mesosphere/lower thermosphere (MLT) heights (80–100 km) have been carried out for several decades. The most prominent components of the mid-latitude MLT winds are, apart from planetary waves, the prevailing (mean) winds and the semi-diurnal tidal winds. The diurnal tide peaks at lower latitudes, and reaches substantial amplitudes at midlatitudes only during spring. Analysis of wind datasets has shown that the winds in the MLT region change on the interdecadal time scale, thus indicating possible climatic changes in the MLT wind regime. Some recent results on long-term changes in MLT winds have been published by Jacobi et al. (2001, 2005), Bremer et al. (1997), Merzlyakov and Portnyagin (1999) and Portnyagin et al. (2006). They show that different MLT wind parameters have significantly changed since the beginning of the regular wind measurements in the early 1960s.

Some of the longest MLT wind datasets have been obtained at northern midlatitudes. Figure 4 shows time series of the annual mean zonal and meridional prevailing winds over Obninsk (55° N, 37° E) and Collm (52° N, 15° E). In spite of differences in the measurement techniques (Jacobi et al., 2005) and strong year-to-year wind variations with characteristic periods of several years, the general tendencies in the climatic MLT wind variations at both stations are similar. The sign of the long-term linear trends changes in the early 1990s. The long-term variations are strongest in summer (not shown here). These results have been confirmed by MLT meteor radar wind observations over Kazan (56° N, 49° E) and MF (partial reflection) winds over Saskatoon (52° N, 107° W) during shorter periods (e.g. Jacobi et al., 2001). Based on these results it may be suggested that the non-linear long-term changes in the MLT prevailing winds are characteristic for the whole Northern Hemisphere mid-latitude belt. The total ozone and ozone laminae trend patterns indicate a corresponding change in the sign of trends in the mid-1990s, its origin being probably changes in trends in dynamics (e.g. Križan and Laštovička, 2005). This indicates a possibility of change of trends in dynamics in the whole middle atmosphere.

An important role in dynamics and energy deposition in the MLT region is played by atmospheric waves, sources and atmospheric filtering of which might be affected by anthropogenic changes in the atmosphere. MLT dynamics are forced through wave coupling, so trends in the mean circulation may be indicative of possible long-term trends in wave activity. The MLT wave activity is determined by the wave sources generally in the troposphere and stratosphere, the wave propagation conditions in the stratosphere/mesosphere, and the interaction processes with the mean circulation in the
MLT. Usually, wave activity is categorized as short-period gravity waves, tides with periods of 24 h plus harmonics, and planetary waves of periods of about 2–30 days.

Global analysis of MLT planetary waves and tides is only possible since the UARS satellite mission, so the time interval available is still too short for trend estimation. Therefore, our knowledge primarily relies on ground-based measurements, in particular radars. A few homogeneous time series exist since the late 1970s.

Most ground-based instruments cover the upper mesosphere and lower thermosphere up to 95–100 km. Results using radar measurements are still inconclusive, generally showing a decrease in tidal amplitudes since the 1960s (Bremer et al., 1997) at mid and high latitudes, which is in accordance with modelling (Ross and Walterscheid, 1991), but the studies have relied on partly inhomogeneous datasets. Additionally, closer inspection of the semidiurnal tide series shows that the negative trend seems to cease after the mid 1980s (Jacobi et al., 1997; Portnyagin et al., 2006), or may even reverse. Observations at Scott Base, Antarctica reveal a positive trend since the late 1980s (Baumgaertner et al., 2005). This is qualitatively consistent with the change of trends after the 1980s in the prevailing winds shown in Fig. 4. Jarvis (2005) used Northern Hemisphere magnetometer measurements to analyse the changes since the early 20th century. He found a decrease of tidal amplitudes at midlatitudes, no clear trend at low latitudes, and some strong but insignificant trends at high latitudes. Also the magnetometer time series show a (slight) tendency for smaller/reversed changes at the end of the century.

Results of gravity wave analyses from ground-based systems are strongly dependent on system characteristics, for example, the wave periods seen by MF radars (>10 min) and other systems (meteor radars, >1 h; LF drift measurements, 0.7–3 h) are very different and results not comparable. The observed long-term decrease in the summer meridional wind (e.g. Bremer at al., 1997) should be connected with changes in the gravity wave flux divergence (Jacobi et al., 2003) owing to changes in middle atmosphere gravity wave wind filtering, but experimental evidence of gravity wave activity does not show a significant trend (Gavrilov et al., 2002).

Planetary wave analyses often describe either the total variability of winds or temperature in the period interval of about 2–30 days (Bittner et al., 2000), or they focus on specific waves such as the quasi 2-day wave (Chshyolkova et al., 2005). Few analyses of waves on a global or hemispheric scale are available, and none of these has considered interannual variability. There is a tendency for a slight increase of the overall day-to-day variability in the 1980s and 1990s (Bittner et al., 2000), but this result is based on only a few datasets, and is neither generally confirmed by a global analysis, nor explained theoretically. A more indirect analysis of planetary waves was performed using radio wave absorption variations (e.g. Laštovička et al., 1994). These analyses showed a possible positive trend in the late 1970s and 1980s, but not in the 1960s and 1990s, thus the trend is intermittent (Laštovička, 2002).

To summarise, the observed trends in MLT wave activity and winds either levelled off or changed the sign in the very late 1980s or early 1990s. These non-monotonic trends are observed for quantities that are strongly affected by waves coming from below, from the troposphere and stratosphere, i.e. the resulting trends are a mixture of tropospheric, stratospheric and MLT region trends in various parameters and cannot be directly compared with other trends in the MLT region; their different behaviour is natural.

Another area of problems concerns water vapour and related phenomena. According to regular balloon-borne measurements, the water vapour content in the lower stratosphere has increased by ∼1%/year over last 40–45 years of the 20th century (Oltmans et al., 2000; Rosenlof et al., 2001). Nedoluha et al. (1998) reported a ∼2%/year increase in water vapour near the stratosphere from 1991–1997 based on HALOE and ground-based microwave (WVMS) measurements, but during recent years there has been no indication of any significant changes (Nedoluha et al., 2003). Instead, HALOE, POAM, and balloon measurements have shown a considerable decrease in lower stratospheric water vapour by ∼0.4 ppmv following a sudden drop in 2001, which appears to coincide with anomalous cooling of the tropical tropopause (Randel et al., 2006).

From 1984–1992 global, annually averaged methane increased consistently by ∼0.01 ppmv/yr (Dlugokencky et al., 2003). Such an increase would result in a water vapour increase of up to ∼0.02 ppmv/yr in the lower mesosphere, where all of the methane is oxidized. In 1993, however, this increase slowed down considerably, and from 1999–2002 the methane concentration has remained constant.

Any increase in water vapour in the stratosphere, either from an increase in water vapour entering the stratosphere or an increase resulting from the oxidation of methane, will affect the water vapour mixing ratio in the mesosphere. Measurements of water vapour in the mesosphere have been available since the early 1990s from HALOE and from the WVMS instruments, but studies of water vapour trends in this region are complicated by solar-cycle variations in the solar Lyman-α radiation. Nevertheless, HALOE measurements in the upper mesosphere clearly do show that after almost a full solar cycle there has been an increase in mesospheric water (Nedoluha et al., 2003); however, this is not the case for very recent years since about 1997.

It is necessary to mention again that similar to trends in dynamics, trends in water vapour are largely determined by trends in tropospheric sources and, therefore, can hardly be directly compared with other trends in the MLT region.

Noctilucent clouds (NLC), consisting of fine ice particles, appear at heights of ∼82–85 km during summer at high latitudes. Their appearance is controlled by temperature and water vapour. Long-term trends in the occurrence frequency and brightness of NLCs observed from ground at latitudes
below about 65°N in Europe were the subject of discussion for a number of years. Finally it was found that there was no detectable trend in the NLCs over the period 1961–2000 (Kirkwood and Stebel, 2003). This agrees with the absence of a temperature trend above Scandinavia near 82–83 km in summer (e.g. Lübken, 2000). When observed from space, NLCs are generally referred to as polar mesospheric clouds (PMCs). Satellite observations cover a larger geographical area and the highest occurrence frequency of PMCs is around 80°N. The satellite observations indicate an increase in PMC occurrence frequency and brightness over the period 1978–2003 (DeLand et al., 2006). The NLC and PMC data sets describe different time periods and different geographical areas with different temperature-water vapour conditions, which may contribute to the observed differences in trends. In connection with this it should be mentioned that Stevens et al. (2005) have shown that space shuttle exhaust can contribute significantly to the number of PMC/NLCs observed in a season. Kirkwood and Dalin (2006) suggested that the shifting local time of satellite observations may introduce a false PMC trend. The absence of a trend in temperatures at NLC/PMC heights during the summer NLC/PMC season, combined with the absence of a clear water vapour trend covering the entire period of NLC/PMC observations seems to be consistent with absence of an unambiguous trend in NLC/PMCs.

There is no observational information about long-term trends in two other important minor constituents in the mesosphere and mesopause region, ozone and nitric oxide (NO), except for the analysis of ozone data from UARS/HALOE satellite observations, which provides no detectable trend for sunrise ozone but a strong negative trend for sunset ozone in a narrow region near 80 km, in anticorrelation with the water vapour trends (Marsh et al., 2003). However, due to the length of the analysed interval and uncertainty in the present-day water vapour trend, the mesospheric ozone trend remains uncertain. Model simulations for ozone suggest a general increase of ozone concentration at fixed pressure levels in response to increasing concentrations of greenhouse gases (the pattern at fixed heights is more complex) (Schmidt and Brasseur, 2004). As for nitric oxide, model simulations predict a substantial reduction of its concentration for the 2×CO₂ scenario (e.g. Beig, 2000).

The third area of problems with trends is the F2 region ionosphere. Two characteristics of the ionospheric F2 region, the critical frequency foF2 (which is directly related to the maximum electron density NmF2) and the height of the maximum, hmF2, are available for several decades from the global network of ionosondes. Trends in foF2 have been extensively studied (e.g. Danilov, 2002; Laštovička et al., 2006, and references therein). However, the results remain contradictory, even though a recent comparison of various methods of trend determination on a test dataset removed some of the discrepancies (Laštovička et al., 2006); computed foF2 trends were predominately negative but weak, of the order of −0.1 MHz/decade, but results of two out of six teams, which analyzed the same data set, still differed substantially. Long-term data homogeneity and problems with adequate removal of solar cycle effects belong to further causes of the discrepancies among various published results.

The peak of the F2 layer tends to remain at a fixed pressure level within the thermosphere (except during geomagnetic storms) and the altitude of that pressure level responds to changes in temperature. There have been conflicting results with most regions indicating upward (warming) hmF2 trends and some indicating downward (cooling) hmF2 trends (e.g. Ulich and Turunen, 1997; Bremer, 1998; Jarvis et al., 1998; Xu et al., 2004). Other studies show considerable variability among different sites (e.g. Upadhyay and Mahajan, 1998). Some of this variability is a consequence of residual solar cycle influence remaining in the trend estimates; this may be the result of complications which arise from the non-linearity of the response of the ionosphere to solar input (Xu et al., 2004). Accounting for these residuals and using some of the longest and most continuous data series available results in consistently negative hmF2 trends that converge to about −1.5 km/decade (Clilverd et al., 2003). This is perhaps the most reliable result on trends in hmF2 and it is consistent with the density changes calculated from satellite drag and model expectations. However, the cause of regional differences, for example between Western and Eastern Europe, still remains to be resolved. Another problem of hmF2 investigations is the necessity to calculate hmF2 from the observed M(3000)F2, which might be the source of some discrepancies between the results of various authors.

In addition to the problems with the consistency of hmF2 trends, there is an ongoing debate about their origin. Some authors believe in the predominantly greenhouse origin, some in the predominantly geomagnetic activity origin (the long-term changes of geomagnetic activity in the 20th century had qualitatively the same impact on trends in foF2 and hmF2 as the greenhouse effect), and some believe both are important, which is more probable for present-day trends. Another source has also been suggested: a long-term decrease of thermospheric atomic oxygen concentration (Danilov, 2005).

6 Modelling of long-term trends

After the pioneering studies by Roble and Dickinson (1989) and Rishbeth and Roble (1992) modelling of long-term changes in the neutral upper atmosphere has primarily focused on more realistic scenarios of greenhouse forcing. Recent studies have elucidated the role of different physical mechanisms, provided uncertainty estimates, and explained some of the observed features. Various models predict weak or no trends around the mesopause and in the summer upper mesosphere (e.g. Akmaev and Fomichev 2000; Bremer and Berger, 2002; Gruzdev and Brasseur, 2005; Akmaev et al.,
Fig. 5. Trends in Earth’s atmosphere. The atmospheric layers are defined by the temperature profile. The ionospheric layers are defined by the electron density profile (midnight at equator). Arrows indicate the direction of change. Red – warming; blue – cooling; green – no change of temperature; black – changes in electron density (horizontal) and heights of ionospheric layers (vertical).

2006; Schmidt et al., 2006; Fomichev et al., 2007). This result, explained by the simple fact that these regions of the upper atmosphere are already sufficiently cold, is in good agreement with the observations (Beig et al., 2003; Beig, 2006). The solar-cycle dependence of density trends in the thermosphere, first predicted theoretically for the 2×CO2 scenario (Roble and Dickinson, 1989; Rishbeth and Roble, 1992), has been found in observational analyses (Fig. 2) and has recently been confirmed in simulations with the observed greenhouse forcing (Qian et al., 2006). Stronger trends at solar minimum vs. solar maximum result from a combination of two factors. First, the thermosphere is substantially colder and so has a smaller scale height at solar minimum; any downward displacement of constant-pressure levels then manifests in a stronger density decrease at a fixed height. Secondly, increased amounts of nitric oxide during solar maximum produce stronger infrared cooling dampening the effect of greenhouse cooling under these conditions (Qian et al., 2006).

Atmospheric infrared radiative transfer is fundamentally non-local and represents one of the processes coupling the upper and the lower atmosphere. Possible effects of documented long-term changes at lower levels, especially in the stratosphere, may be influential all the way to the top of the atmosphere, as Fig. 1 indicates (Bremer and Berger, 2002; Akmaev et al., 2006). Self-consistent modelling of global change in the coupled atmosphere-ionosphere system awaits the development of future “whole atmosphere” models, as current comprehensive climate-chemistry general circulation models (e.g. Gruzdev and Brasseur, 2005; Schmidt et al. 2006; Fomichev et al., 2007) do not yet encompass the F region. Available one-dimensional studies (e.g. Beig, 2000) predict quite dramatic changes in the thermospheric minor species and ion composition, including a reduction in NO number density of up to 95% in response to the 2×CO2 greenhouse forcing.

More specific outstanding science issues have yet to be addressed by suitable numerical models. Even though the treatment of infrared atmospheric CO2 radiation has improved substantially in recent years, several key input parameters remain uncertain, including collisional excitation rates (Castle et al., 2006) and the vertical distribution of carbon dioxide itself. Many modelers have used a global mean profile of the CO2 volume mixing ratio (VMR) compiled by Fomichev et al. (1998) from early rocket and satellite measurements. Recent satellite observations with a careful treatment of the infrared radiation inversion problem (e.g. Kaufmann et al., 2002; Grossmann et al., 2004) suggest that the VMR may be substantially lower (up to a factor of 2) in the upper mesosphere and lower thermosphere than previously thought. Since the thermosphere is nearly transparent to the outgoing CO2 radiation, the strength of the corresponding greenhouse forcing is roughly proportional to its VMR. The new data will undoubtedly affect the magnitude of the forcing and may to some extent explain the overestimation of the density response (Fig. 2) in simulations by Akmaev et al. (2006).

Just as increasing cloud coverage and sunlight reflection may dampen the greenhouse warming in the troposphere, several physical processes have recently been identified as substantial negative feedbacks acting to suppress the thermospheric response to any external forcing including the greenhouse cooling. In addition to the infrared radiative cooling by NO mentioned above, one of the feedbacks is simply related to the fact that the collisional excitation rate of CO2 molecules by atomic oxygen (one of the main cooling mechanisms in the thermosphere) is strongly dependent on temperature. The dependence is such that at colder temperatures the excitation rate rapidly decreases and so dampens the effect of greenhouse cooling, and vice versa at higher temperatures. It has also been found that, surprisingly, even such processes as molecular heat conduction may substantially contribute to the negative feedback and damping (Akmaev, 2003). Careful theoretical analyses and model diagnostics are needed to untangle the intricate chain of external forcing and internal feedbacks acting in the upper atmosphere and ionosphere and reconcile the results of various models and observations (e.g. Fig. 2).
The trends described in previous Sections allow us to construct a scenario of global change in the upper atmosphere at heights above 50 km, i.e. in the mesosphere, thermosphere and ionosphere. The known potential drivers of upper atmospheric long-term global change include increasing concentrations of greenhouse gases (CO₂, CH₄ and others), anthropogenic changes of the ozone layer and to some extent also of water vapour, and natural long-term variations of geomagnetic activity, which increased throughout the 20th century; the role of long-term change of general solar activity (described by sunspot numbers) is unimportant in the trends observed in the second half of the 20th century.

The observed temperature decline in the mesosphere and lower thermosphere, as well as the electron density trends in the lower ionosphere are at least qualitatively consistent with thermal contraction of the mesosphere as a consequence of the in-situ mesospheric greenhouse cooling, with a noticeable contribution from net cooling resulting from the stratospheric ozone depletion. The lack of strong temperature trends near the mesopause is consistent with model simulations of greenhouse gas increases. Higher up, trends in the ionospheric E and F1 region parameters, in ion temperature near 350 km (only one station), and in thermospheric densities between 200 and 800 km, are also qualitatively consistent with the expected effect of greenhouse cooling and thermal contraction of the atmosphere.

The trends described above form a consistent pattern of global change in the upper atmosphere at heights above 50 km as illustrated by the arrows in Fig. 5 that qualitatively agrees with predicted effects of increasing concentration of greenhouse gases. The upper atmosphere is generally cooling and contracting, and related changes in chemical composition (minor components) and effective recombination rate are affecting the ionosphere. For most upper atmospheric and ionospheric quantities, the dominant driver of long-term trends in the last 3–4 decades is, directly or indirectly, increasing greenhouse forcing. Together with the large body of lower atmospheric trend research, our survey indicates that anthropogenic emissions of greenhouse gases are affecting the atmosphere at nearly all altitudes between ground and space.

Some quantities cannot be included yet into the presented consistent pattern of upper atmospheric global change due to unresolved substantial discrepancies between results of various authors (foF2, hmF2, NLC/PMC). Other quantities exhibit changes in trends during the analyzed period (winds, waves, water vapour). Their trends cannot be straightforwardly compared with trends in other quantities, because they are to a large extent driven by trends in the troposphere and stratosphere. Observational information on trends in some other quantities (e.g. concentrations of ozone and NO in the MLT region) is still missing. Resolving these open problems, filling in gaps in our knowledge of long-term trends, and identifying the driving mechanisms will be the goal of future trend research.

Acknowledgements. J. Laštovička thanks for support the Academy of Sciences of the Czech Republic, project AVOZ 40720517, and the Ministry of Education, Youth and Sports of the Czech Republic, project OC.091. R. A. Akmaev has been supported by grants from the U.S. NSF CEDAR (Coupling, Energetics, and Dynamics of Atmospheric Regions) and NASA LWS (Living With a Star) programs. J. T. Emmert has been supported by NASA LWS and the Office of Naval Research.

Topical Editor M. Pinnock thanks A. Danilov, H. Rishbeth and V. Fomichev for their help in evaluating this paper.

References


Term Changes and Trends in the Atmosphere”, Sodankylä (Finland), 2006.

www.ann-geophys.net/26/1255/2008/