Ionospheric F1 layer long-term trends and the geomagnetic control concept

A. V. Mikhailov
Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation, Troitsk, Moscow Region 142190, Russia

Received: 8 April 2008 – Revised: 11 August 2008 – Accepted: 5 September 2008 – Published: 28 November 2008

Abstract. A previous approach to the ionospheric long-term trend analysis has been applied to the foF1 observations from Slough and Rome in order to investigate a possible relationship between the foF1 and the long-term variation of geomagnetic activity. A 40-year period, starting in 1962, has been used for the analysis. According to the results obtained earlier for F2 and E-region trends, geomagnetic control of the long-term variation has also been revealed for the foF1. Thus, it is now possible to speak about the geomagnetic control of the ionospheric trends in the whole ionosphere. This is not surprising as the Earth’s ionosphere is a single entity that is strongly controlled, either directly or indirectly, by the magnetic field. As with the F2-region, this geomagnetic control is provided via neutral composition and temperature changes. A very long-term (centennial) increase in geomagnetic activity in the 20th century is seen in the long-term foF1 variations as well. After its removal, the residual foF1 trends are very small and insignificant. In principal, this means that the observed foF1 long-term variations have a natural origin and can be attributed to solar and geomagnetic activity long-term variations. However, the situation in the thermosphere has been changing since 1997 and available foF2 observations at the two stations reveal information about the “break down” of the geomagnetic control in the F2-region. Possible reasons of these changes are discussed.

Keywords. Atmospheric composition and structure (Thermosphere – composition and chemistry) – Ionosphere (Ionosphere-atmosphere interactions; Mid-latitude ionosphere)

1 Introduction

During the last few decades, long-term trends of ionospheric parameters have been widely discussed due to general interest to the anthropogenic impact on the ecological system and on the Earth’s upper atmosphere as a part of it. The interest to the problem has been greatly stimulated by the model calculations by Roble and Dickinson (1989), Rishbeth (1990), Rishbeth and Roble (1992) who predicted the ionospheric effects of the atmosphere greenhouse gas concentrations increase. Although these trends are very small and have no practical importance, they may serve as an indicator of long-term changes in the Earth’s upper atmosphere, and their investigation may be interesting and important from this point of view.

The majority of analyses have been devoted to long-term trends in the ionospheric F2 and E-layer parameters since these observations are the most abundant and consistent, while trends in the F1-layer have been considered, as far as it is known, only in some papers (Bremer, 1998, 2001; Sharma et al., 1999; Xu et al., 2004). In these papers it is stressed: a) a positive trend in the foF1 long-term variation, b) a relationship with the global cooling of the thermosphere due to the greenhouse effect (Bremer, 1998, 2001; Sharma et al., 1999), c) the insignificance of the geomagnetic activity (Xu et al., 2004).

It should be noted that despite obvious and well-known contradictions with the ionospheric trend observations (Mikhailov, 2006a), the origin of the greenhouse hypothesis of the trends remains very popular. However, apart from the ionospheric trends, it should be stressed that the greenhouse hypothesis has received serious support from Keating et al. (2000), Emmert et al. (2004) and Marcos et al. (2005) who’s results revealed a steady decrease of the thermospheric density over the period of 2–3 solar cycles.

The leading role of geomagnetic activity in the ionospheric F2 and E layer parameter long-term variations has been

Correspondence to: A. V. Mikhailov
(avm71@orc.ru)
proposed and shown by Danilov and Mikhailov (1999, 2001, 2002); Mikhailov and Marin (2000, 2001); Mikhailov and de la Morena (2003). This is the so-called geomagnetic control concept which explains the main morphological features of the ionospheric trends in the F2 and E-regions by natural variations of solar and geomagnetic activity in the framework of contemporary ionospheric storm mechanisms. It should be noted thus far, that the the geomagnetic control concept is the only hypothesis which can explain the ionospheric trends morphology in a consistent way (Mikhailov, 2002, 2006a). The ionospheric F1-layer is tightly related, via the neutral composition, to the F2-layer, therefore one should expect the geomagnetic control in the foF1 long-term variations as well.

The aim of the paper is to check if such a control really takes place. Firstly, a priori, it is not obvious as F1-layer exhibits a small sensitivity to geomagnetic activity (Buresova and Laštovička, 2001; Buresova et al., 2002) resulting from its formation mechanism (Mikhailov and Schlegel, 2003). Secondly, problems with the foF1 identification from the ground-based ionosonde observations (Shchepkin and Vinitzky, 1981) and a poor quality of routine foF1 data may turn out to be an additional obstacle on this way.

2 Method description

A general approach to the trends analysis in the F2-layer was described by Mikhailov et al. (2002) and in the E-region – by Mikhailov and de la Morena (2003). This method, which turned out to be successful and allowed us to reveal the role of geomagnetic activity in ionospheric trends, was used with some modifications in the present study. The mid-latitude daytime ionosphere at F1-layer heights is controlled by the scheme of photo-chemical processes which is basically the same as in the F2-region. Therefore, as earlier in the case of foF2 and foE analyses, we proceed from an assumption that foF1 long-term variations are due to solar and long-term geomagnetic activity variations which may be presented by $R_{12}$ and 11-year running mean $A_p$ indices, where $R_{12}$ is 12-month running mean sunspot numbers. The method includes the following steps.

1. Observed mid-latitude monthly median foF1 values for 10:00, 11:00, 12:00, 13:00, 14:00 LT are reduced to 12:00 LT moment to give an average noon foF1 value. The dependence $foF1 \propto (\cos \chi_o)^p$ where $\chi_o$ – solar zenith angle and $p=0.2$ (Rishbeth and Garriott, 1969) is used for this reduction. The use of an average foF1 increases the reliability of the observed noon foF1 values. The same procedure was applied to reduce these average foF1 obtained for each month to solar zenith angle of 15 June as the method works with annual mean values. In fact, foF1 is regularly registered only in summer, this takes place especially under solar maximum conditions. During medium and low solar activity foF1 are also observed in equinoxes. Therefore, annual mean values were calculated over 2-5 months (April–August). In the case of foF2 analysis, all 12 months are used to find annual mean values.

2. A regression of this noon foF1 with $R_{12}$

$$foF1_{reg} = a_0 + a_1 R_{12}^\alpha$$

is used to define monthly relative deviations

$$\delta foF1 = (foF1_{obs} - foF1_{reg})/foF1_{obs}$$

Initially it was proposed by Danilov and Mikhailov (1998, 1999), we analyze the relative rather than absolute $\Delta foF1$ deviations considered in the ionospheric trend analyses by most of the authors. Relative deviations $\delta foF1$ allow us to combine different months and get annual mean $\delta foF1$ which are used in the analysis, the final method being based on the 11-year running mean $\delta foF1$ values. Universally a simple arithmetic running mean smoothing with an 11-year gate is applied. The optimal 5 (for April–August) different values of $\alpha$ are specified to provide the least standard deviation (SD) after a regression (see later) of 11-year smoothed $\delta foF1$ values with $A_{p,132}$ (11-year running mean $A_p$ indices). The 11-year $\delta foF1$ smoothing requires all 5 values of $\alpha$ to be available simultaneously at each step of the SD minimization. This implies an application of special multi-regression methods (Press et al., 1992).

The expression (1) is of a general type and depending on $\alpha$ it can describe both the linear and non-linear relationship of foF1 with $R_{12}$. The regression coefficients $a_i$ are specified by the least squares method for each month and a given $\alpha$ value. It should be stressed that the expression (1) does not provide the best approximation of the observed foF1 versus $R_{12}$ dependence (other dependencies may give less of a sum of residuals), but it should be considered in terms of the following $\delta foF1_{132}$ regression with $A_{p,132}$ to find the minimal SD (see later). Therefore, the regression (1) is not a “model”, in the usual sense of the word, as it is accepted in other approaches. This regression is used to remove the solar activity part from the observed foF1 variations. A “pure” foF1 dependence on solar activity (presented by the $R_{12}$ index) a priori is not known for each month.

3. One-hour gaps in foF1 within the 10:00–14:00 LT interval are filled by interpolation, but the large gaps in observations are not filled, so some years are marked as “zero”. During the 11-year $\delta foF1$ smoothing the arithmetic mean is calculated over the non-zero years only. Due to this 11-year smoothing 10 years with observations (5 years in the beginning and 5 years in the end) are lost in the analysis, therefore only stations with long enough observational periods could be used. In the
The geomagnetic activity effect is removed from the 11-year running mean \( \delta f_{oF1} \) variation using a regression with \( A_{p132} \)

\[
\delta f_{oF1_{132}} = b_0 + b_1 A_{p132}(t + n) + b_2 A_{p132}^2(t + n) \quad (3)
\]

where \( \delta f_{oF1_{132}} \) and \( A_{p132} \) are 11-year running mean values, \( t \) denotes the years and \( n \) (varies from 0 to −5) is a time shift in years of \( A_{p132} \) with respect to \( \delta f_{oF1_{132}} \) variations which is selected to give the least SD for the residuals after Eq. (3). The regression coefficients \( b_i \) are found by the least squares method.

Our previous analysis (Mikhailov et al., 2002) has shown that the best results (the least SD) can be obtained if an additional smoothing is applied to \( \delta f_{oF1_{132}} \) and \( A_{p132} \) variations. Such a smoothing is made by a 5-order polynomial approximation of these parameter variations.

The residual linear trend with the slope \( K_r \) (in \( 10^{-4} \) per year) may be estimated over the residuals after the regression (3).

The test of significance for the linear trend parameter \( K_r \) (the slope) is made with Fisher’s \( F \) criterion (Pollard, 1977)

\[
F = r^2(N - 2)/(1 - r^2)
\]

where \( r \) is the correlation coefficient and \( N \) is the number of pairs considered. Keeping in mind that we work with smoothed variations, we put the number of degrees of freedom \((N - 2) = 4\) (the 5th order polynomial is defined by 6 coefficients).

### 3 Geomagnetic control

Figure 1 shows the aspect of geomagnetic activity variation on \( \delta f_{oF1} \) for the two stations with the longest available period of observations – Slough/Chilton (1959–2007) and Rome (1957–2007).

The 11-year running mean \( \delta f_{oF1} \) at 12:00 LT are given in comparison with the \( A_{p132} \) index variation. Periods of decreasing geomagnetic activity (1956–1967 and 1987–2007) are seen to correspond to the positive trends in \( f_{oF1} \) with a 4–5 year time shift with respect to the \( A_{p132} \) variation. The inverse situation takes place for the period of increasing geomagnetic activity (1968–1986) with the negative \( f_{oF1} \) trends. This is the geomagnetic control of the \( f_{oF1} \) long-term variations which were revealed earlier for the F2- and E-region long-term trends (Mikhailov and Marin, 2000; Mikhailov and de la Morena, 2003). This result tells us that one should be careful with the selection of time periods for trend analysis and should not put together years belonging to different (rising/falling) periods of geomagnetic activity, however, this is not taken into account in other publications devoted to the ionospheric trend analyses.

Although the two stations qualitatively demonstrate similar reaction to the geomagnetic activity variation, the magnitude of this reaction is different (cf. the period after 1992) and this is another aspect of the geomagnetic control. Rome, as a low-latitude station, exhibits a more pronounced \( f_{oF1} \) increase after 1992. This results in a positive trend estimated over the whole period (dash line in Fig. 1), while this trend is negative for Slough (solid line), the time period being practically the same for the two stations. Again, this result tells us about the necessity to separate different phases (rising/falling) of geomagnetic activity during long-term trends analysis. It should be stressed that such an obvious geomagnetic control takes place when the geomagnetic activity effects are not removed from the \( \delta f_{oF1} \) variations using the expression (3). The final trends are considered later.
Figure 2 (top panels) shows the $\delta f_{oF1}$ dependence on $A_{p132}$. Two peculiarities should be noted in relation with these results. At first, the $\delta f_{oF1}$ increase is about two times larger at Rome compared to Slough over the last 10 years. And secondly, the type of dependence has strongly changed after 1992 – it became much steeper compared to the whole previous period. This is a new feature telling us about changes in the Earth’s thermosphere. As the F1-layer is controlled by photo-chemical processes (see later), this new type of dependence (1) implies that the same changes in geomagnetic activity measured in Fig. 2, top panels) exhibits stronger variations than for the previous period. The causes of this change are not known, but in the framework of our approach, it is possible to propose the following. The state of the thermosphere is mainly controlled by solar and geomagnetic activity – at least all modern empirical thermospheric models are based on this idea. Presumably, solar activity effects are removed from $f_{oF1}$ variations by expression (1), so we have only the geomagnetic activity effects left. The new type of $\delta f_{oF1}$ dependence on $A_{p132}$ after 1992 (Fig. 2, top panels) implies that the same changes in geomagnetic activity measured in $A_{p132}$ now correspond to stronger variations of neutral composition. It should be stressed that actual causes of these changes are not known yet and they may not be directly related to changes in the geomagnetic activity efficiency. It is only possible to state the change in this dependence after 1992.

4 Residual $f_{oF1}$ trends

One may try to remove the geomagnetic activity effect from the $\delta f_{oF1}$ trend and to obtain the residual trend as it was done in the F2-region (Mikhailov et al., 2002). At first, the obvious dependence on the phase of geomagnetic activity (Fig. 1) is removed using the Eq. (3). The results for the two stations are shown in Fig. 3 (top panels). Unlike the F2-region (Mikhailov et al., 2002, their Fig. 3), the quality of fitting, in this case, is not that good. Although the residual trends are clearly seen – a negative trend $K=-2.53 \times 10^{-4}$ per year on Slough and a positive trend $K=+2.09 \times 10^{-4}$ per year on Rome – they are insignificant as the confidence level is $\leq 90\%$ according to Fisher’s F criterion. These residual trends may be also removed by adding a complementary trend to $\delta f_{oF1}$. The results are shown in Fig. 3 (bottom panels). After this step the residual trends become very small – $K_{c}=-1.95 \times 10^{-6}$ per year on Slough and $K_{c}=-6.93 \times 10^{-7}$ per year on Rome and they are absolutely insignificant. Taking the average $f_{oF1}=5$ MHz, these two trends may be rewritten as $-9.7 \times 10^{-6}$ MHz per year on Slough and $-3.5 \times 10^{-6}$ MHz per year on Rome. It is interesting to note that the complementary trends $K_{c}$ have a different sign – positive for a mid-latitude station Slough and negative for a low-latitude station Rome. Similar
analysis for F2-layer trends gave \( K_c = +4.4 \times 10^{-4} \) per year and \( K_r = -2.32 \times 10^{-6} \) per year at Slough, at Rome we have \( K_c = -8.5 \times 10^{-4} \) per year and \( K_r = -1.55 \times 10^{-5} \) per year, the residual trends being insignificant. The meaning of these complementary trends is considered in the Discussion section.

5 Interpretation

Similar to F2-region, the revealed \( f_{o}F1 \) trends demonstrate the dependence on geomagnetic activity (Fig. 1). And this is not surprising as both ionospheric regions are strongly controlled (at least during daytime) by neutral composition which depends on geomagnetic activity. Therefore, it will be useful to consider the trends in the two ionospheric regions simultaneously. Figure 4 similar to Fig. 1 gives \( A_{p132} \) and \( \delta f_{o}F2_{-132} \) variations for Slough/Chilton and Rome for the whole available period of observations. The geomagnetic control in \( f_{o}F2 \) long-term variations with a 5-year shift with respect to \( A_{p132} \) is seen on both stations over the whole period, with the exception of the last 5 years after 1997. Contrary to the expected increase in \( \delta f_{o}F2 \) due to a continuous decrease in the geomagnetic activity we have a sudden change in the \( \delta f_{o}F2 \) dependence taking place at the two stations simultaneously. The \( \delta f_{o}F2 \) vs. \( A_{p132} \) dependences are shown in Fig. 2 (bottom panels). Similar to F1-region, there is a change in the dependence exactly for these last years, although it is less pronounced and started later compared to the F1 layer case. Obviously both effects in the F1 and F2-regions are related.

Physical interpretation of the F1 layer results may be given from an analysis of a simple scheme of photochemical processes controlling the daytime ionosphere at F1 layer heights (Mikhailov and Schlegel, 2003). The main processes are photo-ionization of neutral \([O], [O_2], [N_2] \) by solar EUV radiation, the conversion of primary ions to molecular ones via ion-molecule reactions followed by the dissociative recombination of molecular ions with electrons. This scheme of processes may be written as

\[
\begin{align*}
q(O^+) = & \gamma_1[N_2] + \gamma_2[O_2] \\
q(N_2^+) = & \gamma_3[O][N_2^+] \\
q(O_3^+) = & \gamma_4[O_2][O^+] = \alpha_2[O_2^+]N_e \\
\gamma_1[N_2][O^+] + & \gamma_2[O][N_2] = \alpha_1[NO^+]N_e \\
N_e = & [O^+] + [O_2^+] + [NO^+] 
\end{align*}
\]

where \( q_i \) – primary ion production rates, \( \gamma_i \) – ion-molecule reaction rate coefficients, and \( \alpha_i \) – dissociative recombination rate coefficients. Equilibrium concentration of \( N_2^+ \) ions is negligible compared to the main ions.

For the sake of a simple analysis, in accordance with Ivanov-Kholodny and Nikoljsky (1969), let us consider the ionosphere at F1-region heights consisting of atomic \( O^+ \) and molecular \( M^+=NO^+ + O_2^+ \) ions. From Eq. (4) we have for \( O^+ \)
ions \( q(O^+) = \beta [O^+] \) where \( \beta = \gamma_1 [N_2] + \gamma_2 [O_2] \) and for molecular ions \( M^+ = NO^+ + O_2^+ \) produced both by direct photo-ionization and via ion-molecule reactions we may write

\[
q(O^+) + q(M^+) = q = \alpha_{ave}[M^+]N_e
\]

where \( \alpha_{ave} = \alpha_1 + [NO^+] + \alpha_2 [O_2]^+ \) is the average-weighted dissociative recombination rate coefficient.

Keeping in mind that \( M^+ = N_e - [O^+] \) and \( [O^+] = q(O^+)/\beta \) we get from Eq. (5) the equation

\[
N_e^2 - N_e q(\beta) = 0
\]

which may be rewritten as

\[
N_e = \frac{q(O^+)}{\beta} + \frac{q}{\alpha_{ave} N_e}
\]

The first term in Eq. (7) presents the \( O^+ \) ion concentration and the second term – the concentration of \( M^+ \) ions. A comparison (Mikhailov and Schlegel, 2003) with the model calculations, when the complete set of pertinent processes is taken into account, has shown that a simplified approach (Eq. 7) may be used for physical interpretation.

Atomic \( O^+ \) and molecular \( M^+ \) ion concentrations change in opposite directions under varying geomagnetic activity. Usually \( [O^+] \) variations dominate over \( [M^+] \) ones and this determines the sign of \( \delta_0F1 \) changes (Mikhailov and Schlegel, 2003). Roughly \( [O^+] \) is proportional to \( [O]/[N_2] \) and well-established changes of this ratio (e.g. Prölls, 1995) under varying geomagnetic activity explain both the \( \delta_0F1 \) vs. \( \Delta_{p132} \) dependence (Fig. 2, top panels) and the geomagnetic control of the \( foF1 \) long-term variations (Fig. 1). In the framework of this mechanism a steeper \( \delta_0F1 \) vs. \( \Delta_{p132} \) dependence for the last decade (Fig. 2, top panels) simply tells us about a sharper increase of the \( [O]/[N_2] \) ratio compared to the whole previous period. This is a new and interesting result obtained indirectly via analysis of the ionospheric long-term trends.

Different magnitudes of the \( foF1 \) increase in Rome and Slough observed after 1992 (Fig. 1) may be related to different latitudes of the stations – \( \Phi_{inv} = 37.5^\circ \) for Rome and \( \Phi_{inv} = 49.8^\circ \) for Slough. The \( F1 \) layer trend results (Fig. 1) were obtained mostly on summer months (April–August). There is always a summer/winter latitudinal gradient of \([O]\) with lower atomic oxygen concentrations towards the summer pole (e.g. Picone et al., 2002), that is, the \( [O]/[N_2] \) ratio is smaller and the compensating role of \( M^+ \) ions in Eq. (7) is larger.

Both \( F_2 \) - and \( F_1 \) regions exhibit the geomagnetic control – that is, the decreasing phase of geomagnetic activity is accompanied by positive ionospheric trends and vice versa (Figs. 1 and 4). The similarity in \( foF1 \) and \( foF2 \) long-term variations is provided by neutral composition (O, O\(_2\), N\(_2\)) variations which are closely related in the \( F1 \) and \( F2 \) regions due to a process of molecular diffusion which is very efficient at these heights. A very good correlation between \( \delta N_e \) at \( F1 \) layer heights and \( \delta NmF2 \) has been shown by Mikhailov and Schlegel (2003, their Table 3) for disturbed conditions using ISR observations.

But this coherence breaks down after 1997 (cf. Figs. 1 and 4). While the \( foF1 \) increase continues as a result of the decreasing geomagnetic activity, in the \( F2 \) region we have a decrease at least for the last available 5 years. At present, the theory of the mid-latitude daytime \( F2 \) region is well-developed. According to this theory, under increasing of the \( [O]/[N_2] \) ratio, the only way to decrease \( NmF2 \) is to enhance the poleward thermospheric wind \( Vnx \), which produces the downward plasma drift decreasing \( NmF2 \). It is interesting to note the difference in the reaction of the \( F1 \) - and \( F2 \) regions to this presumably enhanced \( Vnx \). While in the \( F1 \) region the new type of the \( \delta_0F1 \) vs. \( \Delta_{p132} \) dependence appeared approximately 10 years ago, in the \( F2 \) region the deviation from the normal type of this dependence has been observed only for the last 4–5 years (Fig. 2). The \( F1 \) region, being in photo-chemical equilibrium, is totally controlled by neutral composition variations, while in the daytime \( F2 \) region,
the poleward Vnx compensates to some extent the positive effect of the [O]/[N$_2$] ratio increase. The bend in the $\delta f_{oF2}$ vs. $A_p$ dependence (Fig. 2, bottom panel) obviously tells about the increasing Vnx in time which overpowers the [O]/[N$_2$] effect. It should be stressed that the conclusion on the Vnx increase is based on the knowledge of the daytime middle latitude F1 and F2 layer formation mechanisms only, no direct experimental support for such Vnx long-term variations at the thermospheric heights is available at present as far as it is known (see also Discussion section).

6 Rishbeth and Roble results analysis

A classical paper by Rishbeth and Roble (1992) is referred to by every publication devoted to ionospheric trends, and the present paper is not exclusion. While very modest deviations predicted in the F2 and E-regions look reasonable, although we are still very far from the doubled CO$_2$ scenario considered in their paper, the results on the F1-region with a 40–50% increase in electron concentration (their Fig. 9) look absolutely unreal. The mechanism of such Ne increase is not seen from the point of view of the present day theory of the F1 layer formation. Model calculations analogues to Mikhailov and Schlegel (2003) were made to estimate possible variations of NeF1 using geophysical conditions, temperature and neutral composition changes as they are given by Rishbeth and Roble (1992). The NeF1 variations were estimated at a constant pressure level and at a fixed height 175 km for the Millstone Hill ISR location. According to the isobaric concept by Rishbeth and Edwards (1989, 1990) which is a more appropriate F1 layer, where vertical plasma transfer is inefficient unlike F2 layer, one should not expect any changes of Ne at a constant pressure level. The calculations have exactly confirmed the prediction, while at the fixed height 175 km we have a 13% increase in Ne. The latter is explainable as under a 40 K decrease in Tn the O/N$_2$ ratio increases at a fixed height (see Discussion). Thus, during daytime (and F1-layer is a daytime formation) under the CO$_2$ cooling effects in the thermosphere predicted by Rishbeth and Roble (1992) one may expect a 13% increase in NeF1 at a fixed height, which is much less than in the cited paper. Therefore, what is the origin of the 40–50% increase in NeF1?

The results by Rishbeth and Roble (1992) are given for 00:00 UT as the zonally averaged relative differences $\Delta$Ne/Ne (%) (their Fig. 9). This means that different local times were put together. Such an approach may be acceptable for thermospheric parameters whose diurnal variations are not that large, but not for the ionospheric F1-region with the Ne diurnal variations up to 100 times. Figure 5 gives NeF1 diurnal variations at 170–175 km observed by Millstone Hill ISR on magnetically quiet days 7 January 1997 and 10 July 1986 during solar minima with $F_{10.7} \approx 70$. The selected days are close to December solstice under solar minimum conditions used by Rishbeth and Roble (1992). The magnitude of the Ne observed diurnal variations are seen to be larger than two orders of magnitude. Hardly the ionospheric effects of CO$_2$ cooling are comparable with such large background NeF1 variations. Therefore, $\Delta$Ne/Ne may be very large during the nighttime hours when Ne is low. The ionosphere is not sunlit in December most of the time especially at high latitudes, where the authors obtained the largest $\Delta$Ne/Ne (their Fig. 9). On the contrary, in the summer hemisphere the ionosphere is sunlit most of the time at high latitudes and the contribution of nighttime hours, with relatively low Ne, is not large. Millstone Hill with $\varphi = 42.6$ N is not a high-latitude station, although this tendency is clearly seen at this latitude as well (Fig. 5). Summer conditions result in small (5–10%) $\Delta$Ne/Ne deviations (their Fig. 9), which may be considered as a real effect of CO$_2$ cooling in the F1-region.

7 Discussion

The undertaken analysis of the $f_{oF1}$ long-term variations has confirmed the expected result – these variations are also subjected to the geomagnetic control, which was earlier revealed in F2 and E-layer trends. The mechanism of this control is different in dissimilar ionospheric regions. In the mid-latitude daytime E-region, the geomagnetic control is provided via nitric oxide NO variations (Mikhailov and de la Morena, 2003; Mikhailov, 2006b). In the F1- and F2-regions, this control is mainly provided by neutral composition (O, O$_2$, N$_2$) variations, the meridional thermospheric wind being also important in the F2-region. Neutral composition variations are practically the same in the F1- and F2-regions due to a very efficient molecular diffusion process in the thermosphere. Therefore, the geomagnetic control concept initially developed for the F2 layer long-term trends
(Mikhailov, 2002, and references therein) is valid for the F1 layer trends as well. This concept is based on the contemporary theory of the F-layer storms (Rishbeth, 1991; Fuller-Rowell et al., 1994, 1996; Prölls, 1995; Rishbeth and Field, 1997; Rishbeth and Müller-Wodarg, 1999). The interaction of the background solar-driven and storm-induced thermospheric circulations is the basic process in this storm mechanism resulting in neutral composition variations.

The geomagnetic control concept of the ionospheric trends proceeds from the assumption that the storm mechanisms controlling the F-region on the short time scales (hours and days) are the same on the long time scales compared to a solar cycle length. The concept has emerged from the analysis of the F2 layer parameter long-term trends and turned out to be efficient in the interpretation of the ionospheric trend’s global morphology. Unlike a very popular greenhouse hypothesis of the ionospheric trends, which encounters many unsolved problems (Mikhailov, 2002, 2006a), the geomagnetic control concept explains, in a consistent way, the observed long-term variations in the E-, F1-, and F2-regions. A fairly sophisticated method (Mikhailov et al., 2002) has been developed to extract long-term trends from routine ionospheric observations, but namely this method has allowed us to explain the observed long-term variations of the ionospheric parameters by natural long-term variations of solar and geomagnetic activity. It is important to stress that we mainly use the same method for the trends analysis in the F2-, F1- and E-regions which is applied to the whole period of available observations. Therefore, any peculiarities revealed in the ionosphere long-term variations, for instance, unusual variations during the last decade (Figs. 2, 4) cannot be attributed to the method, but should have their origin in the Earth’s atmosphere itself.

As an attempt to interpret such revealed peculiarities we consider, for instance, a sharper reaction of the thermospheric neutral composition to the geomagnetic activity variations (Fig. 2), or an increase of the northward Vnx (Fig. 4). Such explanations should be considered only as plausible ones which follow from the formation mechanisms of the F1- and F2-regions as there are no direct observations of such long term trends in the thermospheric parameters thus far. But these proposed explanations are not the key point of the paper. The main result of our analysis is the geomagnetic control of the foF1 long-term variations similar to foF2 ones and this control practically takes place for the whole period of available observations (Figs. 1, 4).

Another aspect of the geomagnetic activity impact on the ionosphere long-term variations, is a complementary trend proposed by Mikhailov et al. (2002) for the interpretation of the F2 layer long-term trends. A similar approach has been applied to the F1 layer trends (Fig. 3). An addition of a linear complementary trend k to the observed δfoF1,32 practically completely removes the geomagnetic activity effect resulting in very small and insignificant residual trends. The complementary trend may be considered as a compensation of negative in Slough and positive in Rome trends initially presented in the observed δfoF1,32 values. These trends presumably have the same geomagnetic activity origin and are due to a very long-term increase in the geomagnetic activity in the 20th century. This increase in geomagnetic activity is widely discussed in recent publications (e.g. Clilverd et al., 1998, 2005; Mursula and Martini, 2006). But the effect of this very long-term increase in the geomagnetic activity cannot be removed even by using very smoothed indices, such as A p,132 and perhaps other long-term indices of geomagnetic activity are needed (Mursula and Martini, 2006).

The required complementary trends are positive at Slough, both in F1- and F2-regions, and negative in Rome (see earlier). This difference in the sign of the complementary trends is in line with general F-layer storm morphology and physics. Normally, under increased geomagnetic activity, the disturbances are negative at high and middle latitudes and positive at lower latitudes, so the compensating trends should have opposite signs as we obtained in our calculations.

In relation to the geomagnetic control of the ionospheric trends and the dependence of trends on the phase (rising/falling) of the geomagnetic activity, it was stressed repeatedly the importance of the periods selection. When the effects of geomagnetic activity are not removed, Fig. 1 shows that one may get positive or negative trends with different magnitudes when different phases of geomagnetic activity are put together. The statement of Xu et al. (2004) that “the effect of geomagnetic activity was not significant” is just a result of a rude method application, as the geomagnetic activity effects are strong and cannot be missed. Further, rough methods applied for trend analyses may result in different sign of foF1 and foF2 trends (Bremer, 1998; Xu et al., 2004).

As it was mentioned earlier, basically F1 and F2 ionospheric regions are closely related via the scheme of photo-chemical processes. Although the meridional neutral wind strongly affects the mid-latitude F2-region, its contribution to NmF2 variations is relatively small during daytime hours as daytime Vnx and corresponding vertical drifts are relatively small (Buonsanto and Witasse, 1999). This idea is also used in the isobaric F2-layer concept by Rishbeth and Edwards (1989, 1990) where the effects of the vertical plasma drift are ignored. Therefore, during the daytime hours in question variations of neutral composition play the leading role and normally F1 and F2 ionospheric regions should exhibit the same sign of trends – the period after 1997 (Figs. 1 and 4) is considered later.

An interesting result is a change of the δfoF1,32 vs. A p,132 dependence after 1992, in the F2-layer this happened later (Fig. 2). The formation mechanism of the F1-region is relatively simple – photo-chemical equilibrium. As far as it is known, there are no indications on any systematic increase in solar EUV radiation for the last 15 years. However, some notes should be made in relation to this. A decrease of neutral temperature should result via a decrease of neutral scale height in an increase of the ion production rate. This seems...
to be in line with the expected CO₂ cooling of the thermosphere. However, under a 15% increase in [CO₂] over a 30-year period (Keeling et al., 2002) one may expect the thermospheric temperature decrease by 10–20 K at best (Rishbeth and Roble, 1992; Akmaev, 2003) and this is only around 2% of the neutral temperature at F1-region heights. Therefore, the analyzed effect can hardly be related to such small decrease in the neutral temperature. Moreover, similar foF1 increase took place during the decreasing phase of the geomagnetic activity before 1965 (Fig. 1) when no CO₂ cooling effects are expected. Therefore, the revealed foF1 increase after 1992 (Fig. 1) should be attributed to the neutral composition variations. According to the scheme of photochemical processes (Eqs. 4–7, also Mikhailov and Schlegel, 2003) such an increase in foF1 may be related only to the O/N₂ ratio increase. It should be stressed that real causes for such changes of neutral composition, which follow from our trend analysis, are not known. But as the method of the analysis, applied for the whole period in question, is the same and the type of the δfoF1 vs. A_p132 dependence is different after 1992, formally one can speak about a new reaction of the thermosphere on geomagnetic activity (Fig. 2, top panel). The other possibility to explain the steeping of the δfoF1 vs. A_p132 dependence, may be based on a correction of the conventional magnetic indices (Mursula and Martini, 2006). Anyway, this result is a good illustration of the use of ionospheric trends analyses for monitoring the long-term changes in the Earth’s upper atmosphere.

A large 4–5 year lag between long-term geomagnetic activity and long-term F2-layer parameter variations have been revealed earlier during our analysis of the foF2 and hmF2 long-term variations (Mikhailov et al., 2002; Mikhailov, 2002). The same time shift takes place in the foF1 variations (Fig. 1). No explanation to this time delay can be proposed at present. As a working hypothesis, one may think that the whole Earth’s atmosphere is involved with the processes provoked by long-term variations of geomagnetic activity. Long-term changes in the global atmospheric circulation and related variations in the thermospheric neutral composition and temperature is the most probable mechanism. Special investigations with 3-D first principle models are needed to answer this question.

An unexpected effect in the foF2 long-term variations has appeared after 1997 on both stations (Fig. 4). Despite the steady decrease in the geomagnetic activity since 1987 and an expected positive trend in foF2, which really took place for some years, we have a negative trend foF2 for the last 5 years. This contradicts the geomagnetic control concept of the ionospheric trends, which is seen to be valid for the whole previous period. Moreover, in the F1-region we have a pronounced positive trend for the same period exactly in accordance with the concept. Positive foF2 trend also implies an O/N₂ increase under decreasing geomagnetic activity (Mikhailov and Marin, 2001). But unlike the F1-region totally controlled by photo-chemical processes, thermospheric winds essentially contribute to the NmF2 variations. Under increasing of the O/N₂ ratio an enhanced poleward thermospheric wind Vnx is the only way to decrease NmF2 in the daytime mid-latITUDE F2-region.

It is necessary to remind us that during the foF2 long-term trend analysis we work with all 12 months to obtain annual mean δfoF2 (Mikhailov et al., 2002). In this case, daytime annual mean Vnx is poleward at middle latitudes (Hedin et al., 1991; Buonsanto and Witasse, 1999). A negative foF2 trend after 1997 implies Vnx enhancement which overpowers the O/N₂ increase. At present, there are no direct observations of such Vnx long-term increase. It is possible to mention the results by Danilov (2008) who has analyzed the correlation coefficient r between daytime and nighttime foF2 values (which normally is negative) and found an increase of abs(r) after 1980, which he has related to an increase in the poleward Vnx. Such an increase of Vnx does not look unreasonable as under decreasing geomagnetic activity (corresponding to a decrease of the auroral heating) the solar driven circulation (poleward during daytime) should enhance (e.g. Fejer et al., 2000; Balan et al., 2004). In principle, such Vnx increase should have been seen in the hmF2 long-term variations. Indeed, such analysis for Slough has shown (Mikhailov, 2006a; Fig. 1) strongly damped hmF2 variations compared to the previous (1950–1975) period. But hmF2 is strongly controlled by neutral temperature and the effect of the greenhouse cooling of the thermosphere can hardly be separated from the discussed Vnx effect as they work in one direction.

In the end, some words about possible effect of the CO₂ cooling. Just a decrease of neutral temperature should result in [O⁺] increase at fixed heights in the F1-region. This can be shown from simple arguments (Mikhailov and Schlegel, 2003). For the sake of simplicity, we may suppose that the thermosphere is isothermal and neutral species [O] and [M]≈[N₂] are distributed in accordance with the barometric law: [O]=O₀exp(–z/H) and [M]=M₀exp(–1.75z/H), where H=kTn/mg is the atomic oxygen scale height. The O⁺ production rate may be written as q(O⁺)=j_o[O]exp(–a Ch₂χ) where j_o is the ionization efficiency depending on the incident solar EUV flux and ionization cross-sections, Ch₂χ is the Chapman function for solar zenith angle χ, and a includes the column density of neutrals multiplied by absorption cross-sections. The linear loss coefficient β may be written as β=γ[M]. In this case [O⁺]=q(O⁺)/βoexp(0.75z/H). Therefore, [O⁺] will increase at a fixed height providing Tn decreases. Normally Δ[O⁺] controls the sign of ΔNe in Eq. (7), therefore one should expect an increase of Ne in the F1-region under decreasing Tn. But it should be kept in mind that all these arguments are valid only if concentrations [O]₀ and [M]₀ are unchanged at the basic level x₀. In reality, this hardly takes place under varying geomagnetic activity and changing global circulation. Therefore, it is a problem to separate and demonstrate the pure effects of the greenhouse cooling.

8 Conclusions

The main results of our analysis may be summarized as follows.

1. Using an earlier developed approach to the ionospheric long-term trend analysis, it was shown for the first time the relationship between \( f_0F1 \) trends and the geomagnetic activity long-term variations. By a complete analogy, with the earlier obtained results for the \( f_0F2 \) and \( f_0E \) trends, the periods of increasing geomagnetic activity correspond to negative \( f_0F1 \) trends while these trends are positive for the decreasing phase in geomagnetic activity. Therefore, it is possible to speak about the geomagnetic control of the \( f_0F1 \) long-term trends as well and this is not surprising as the Earth’s ionosphere is a single whole formation which is strongly controlled directly or indirectly by the magnetic field.

2. The mechanism of this geomagnetic control is provided via neutral composition and temperature variations in the framework of the contemporary theory of the F-layer formation under varying geomagnetic activity. Basically daytime F1 and F2-layer trends should exhibit synchronous variations as they are closely related via common neutral composition and the scheme of photo-chemical processes. However, thermospheric winds which strongly affect the F2 layer (unlike F1-region) can contribute to the F2 layer trends, for instance after 1997.

3. Similarly to the \( f_0F2 \) and \( f_0E \) trends, there exists a background \( f_0F1 \) trend which may be considered as a manifestation of a very long-term (centennial) geomagnetic activity increase which took place during the 20th century (Clilverd et al., 1998, 2005; Mursula and Martini, 2006). After the removal of this background effect, the residual \( f_0F1 \) trends are very small and insignificant. This means that observed \( f_0F1 \) long-term variations (basically) have a natural origin and may be attributed to solar and geomagnetic activity long-term variations.

Acknowledgements. Topical Editor M. Pinnock thanks two anonymous referees for their help in evaluating this paper.

References


A. V. Mikhailov: Ionospheric F1 layer long-term trends and the geomagnetic control concept 3803
