Short-term prediction of ionospheric parameters based on auto-correlation analysis

Ivan Kutiev (1), **Plamen Muhtarov** (1), **Ljiljana R. Cander** (2) **and Mireille F. Levy** (2)
(1) *Geophysical Institute, Sofia, Bulgaria*
(2) *Rutherford Appleton Laboratory, Chilton, Didcot, Oxon , U.K.*

Abstract
A prediction method based on a simple auto-regressive model has been developed for short-term prediction of ionospheric characteristics. The method determines the auto-correlation function for the hourly values of the parameter of interest, using the time series from the previous 25 days. The resulting weighting coefficients can then be used to forecast future values of the parameter. The method has been applied to predict f_0F_2 up to 24 h ahead for stations Uppsala, Slough, Poitiers and Sofia. Error statistics are presented.

Key words: *ionospheric F region – prediction – auto-correlation method*

1. Introduction

Accurate information on the state of the ionosphere is necessary in order to optimize operation of High Frequency (HF) radio systems. Operators often need to plan frequencies a few hours in advance, which requires a short-term forecast of ionospheric characteristics like f_0F_2, the critical frequency of the F_2 layer. Ideally, this should be done using a global numerical model of the coupled thermosphere/ionosphere (Fuller-Rowell *et al.*, 1996), but in practice all necessary input data may not be available and the running times are prohibitive. A simpler solution consists in using empirical methods to predict the behaviour of ionospheric parameters at a given location. One possibility is the use of neural networks (Cander *et al.*, 1998), which have the advantage of taking non-linear phenomena into account.

Here we present a linear technique based on auto-regressive filtering. The Auto-Correlation Method (ACM) was developed for filling gaps due to equipment failures in the time series of measured ionospheric parameters (Muhtarov and Kutiev, 1999). It was then realized that the same approach could be applied to forecasting. In Muhtarov and Kutiev (1999), Muhtarov *et al.* (1998), Dick *et al.* (1998), the auto-correlation method was used to predict f_0F_2 up to 24 h ahead for a number of stations, with encouraging results. The ACM is based on a superposition of diurnal periodic components with an exponential term representing short-term decorrelation of the time-series. In this paper, we describe a simpler model obtained by first removing background diurnal variations using monthly medians.

Mailing address: Dr. Ljiljana R. Cander, Rutherford Appleton Laboratory, Radio Communications Research Unit, Chilton, Didcot, Oxon OX11 0QX, U.K.; e-mail: l.cander@rl.ac.uk

2. The modified auto-correlation model

In ionospheric studies, the average diurnal variations of a characteristic are traditionally represented by the monthly medians. In the present approach, instead of the ionospheric parameter itself, we use its relative deviations from the median values. In the case of f_0F_2, this quantity is given by: $\Phi = (f_0F_2 - f_0F_{2\,\mathrm{med}})/f_0F_{2\,\mathrm{med}}$. We consider Φ as a random process over time. The advantage of this quantity is that it does not contain any periodic components and can be regarded as a stationary process over an observation period of a few days. We assume an auto-regressive model of the form

$$\Phi(t_0) = \sum_{k=1}^{n} \beta_k \Phi(t_0 - \tau_k), \tag{2.1}$$

where n is the order of the regressive model and the β_k for $k = 1, \ldots n$, are weighting coefficients to be determined. We have at our disposal a sample of N measured values Φ_i at times t_i ($i = 1, \ldots N$), ordered in decreasing time. As the process is stationary, the weighting coefficients β_k are solutions of the system of equations (Korn and Korn, 1968)

$$\sum_{k=1}^{n} \beta_k r_{\Phi\Phi}(\tau_{ik}) = r_{\Phi\Phi}(\tau_i); \quad i = 1, 2, \ldots, n; \tag{2.2}$$

where $\tau_{ik} = \tau_i - \tau_k$ and $r_{\Phi\Phi}$ denotes the auto-correlation function of Φ. In practice we do not know the true $r_{\Phi\Phi}$, but only the normalized empirical auto-correlation function $\rho_{\Phi\Phi}$, defined by (Korn and Korn, 1968)

$$\rho_{\Phi\Phi}(\tau) = \frac{\sum \Phi_i \Phi_{k(i)}}{\sqrt{\left[\sum \Phi_i^2\right]\left[\sum \Phi_{k(i)}^2\right]}} \tag{2.3}$$

where the summation in the numerator is taken over the pairs of Φ values having the same time difference τ in the data sample. Note that we assume that the empirical mean $\bar{\Phi} = \frac{1}{N}\sum_i \Phi_i$ is zero, which is justified for ionospheric series over periods of several days. For convenience, the word «normalized» will be omitted hereafter.

We have to bear in mind that the empirical auto-correlation function is only an estimate of the true auto-correlation function. The accuracy of that estimate depends on the size N of the data sample, which should be large compared to the number n of coefficients to be computed. This is not viable for ionospheric time-series, since the required period of time would then be too long for the auto-regressive model to remain valid during the whole period.

The solution adopted here is to assume a correlation function of the form (Muhtarov and Kutiev, 1999)

$$r_{\Phi\Phi}(\tau) = \exp\left(-\frac{|\tau|}{T}\right), \tag{2.4}$$

where T is the time constant of the process $\Phi(t)$. Note that we do not need to include periodic components as for the ACM model form (Muhtarov and Kutiev, 1999) because the diurnal variations of f_0F_2 are already taken into account in the monthly medians. Numerical tests give optimal values for n and N of 24 (1 day) and 600 (25 days) respectively.

3. One-day prediction

To apply the model to 24-h prediction of f_0F_2, we calculate the next 24 hourly values using eq. (2.1). To do this, we take the time series for the past 25 days, calculate the hourly medians of f_0F_2 and then the corresponding values of Φ. We now calculate the empirical auto-correlation function $\rho_{\Phi\Phi}$ over that 25-day period for values of time lags τ_k up to 4 days, and estimate the time constant T by least squares fitting for the system

$$\rho_{\Phi\Phi}(\tau_k) = \exp\left(-\frac{|\tau_k|}{T}\right), \quad k = 1, \ldots, 24. \tag{3.1}$$

Typical values for T are of the order of 10 h. Finally, within the 4-day range of τ, using eq. (2.4) we select the 24 time lags with the highest auto-correlation values which are used to obtain the weighting coefficients β_k from system (2.2). Φ_0 is then obtained from eq. (2.1). The procedure is repeated for each hour i (i = 1… 24).

An example of one-day prediction of f_0F_2 is shown in fig. 1. The one-day prediction is performed for each of the days between 21st and 31st of July 1981 for f_0F_2 data at Slough. The dots represent the measured f_0F_2 (in tenths of MHz) and the full line shows the model predictions. On 25th July an intense geomagnetic storm started at 12:30 UT; the following night f_0F_2 dropped to 50% of its median level. It is seen that the prediction follows this drop fairly well, but with a few hours delay. This reaction time is natural, since values of the few previous hours usually have the highest weights and therefore the biggest contribution in the sum in (2.1). The ionosphere recovered during the next few days, although the daytime values of f_0F_2 on 28th, 29th and 30th are obviously corrupted. In general, the model also predicted a recovery until the 30th of July.

The prediction scheme calculates the 24 values of f_0F_2 at once; it does not use the previously predicted hourly values in the current hourly prediction. For the present analysis, we choose to start prediction from midnight, but it can be

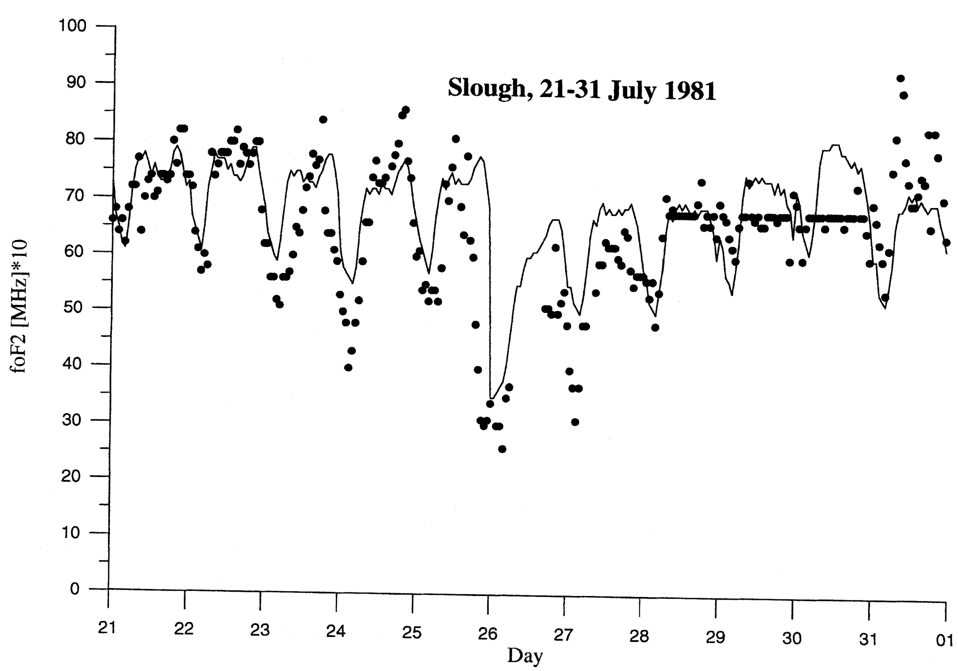

Fig. 1. The one-day prediction of f_0F_2 (full line) compared with data from Slough for the period 21-31 July, 1981. An intense geomagnetic storm starts at 12:30 UT on 25th July. The day numbers on the abscissa are placed at the beginning of the days.

started at any hour. The last measured value used for forecasting is then at 23[rd] h on the previous day. Clearly the prediction error must increase with the time lag from the last measured value. In fig. 2, the full line shows the standard deviation of the individual prediction errors (deviations) as a function of time lag, averaged over the one-day predictions performed for the two years 1981-1982 (upper panel) and 1985-1986 (lower panel) for Slough data. It should be noted that since Φ is defined as the relative deviation of f_0F_2, the above quantity is also the root mean square of the relative deviation of f_0F_2, or in other words the root-mean-square (r.m.s.) percentage deviation of f_0F_2 from the medians. For comparison, the r.m.s. relative difference between the data and its median

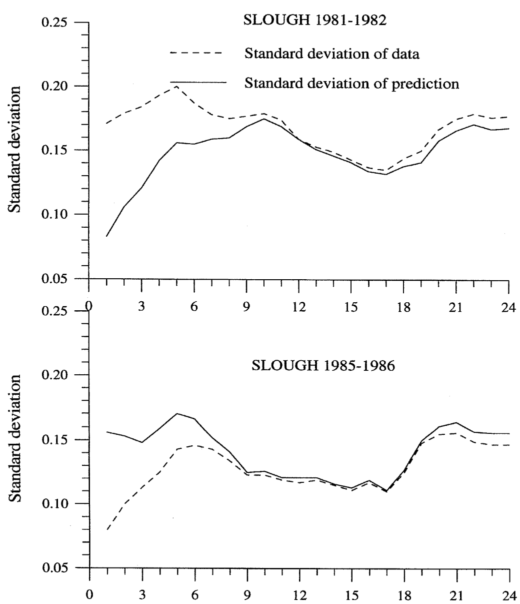

Fig. 2. Model prediction error (full line) and monthly median prediction error (dashed line) of the one-day predictions are plotted against the time lag in hours from the last measured value at 23 UT on previous day. Prediction errors (dimensionless relative standard deviations) are averaged over the data from the whole years 1981 and 1982 (upper panel) and 1985 and 1986 (lower panel).

values computed over the last 25 days for each hour is shown by the dashed line. The latter represents the prediction error of the conventional short-term forecasting method, which uses median values obtained within the last several days. The model prediction error is significantly lower than the monthly median error in the first hours of the prediction period, with the very low value of 8% for the first hour. After 10 h or so, corresponding to the time constant T, the model prediction error becomes similar to the monthly median error. The prediction error again becomes lower when the time lag is closer to 24 h, because the auto-correlation function increases.

The results of the one-day prediction performed during the whole years 1981 and 1982, separately for the stations Uppsala, Slough, Poitiers and Sofia are shown in fig. 3. The abscissas show the monthly median prediction error, the ordinates the model prediction errors. The dots show daily averaged values of the errors. The thin line $y = x$ correspond to the case where the model performance is the same as that of the monthly median prediction. The thick line shows the least squares linear fit of model prediction errors against monthly median prediction errors. Since the slope a of the regression lines is less than 1, the modified auto-correlation model indeed improves on the monthly median predictions. We define the «Prediction Efficiency» (PE) as 1-a, expressed in percent. It is a measure of the gain of accuracy that the model achieves in relation to the monthly median prediction. Figure 4 shows the same quantities for the years of solar minimum 1985 and 1986. Clearly PE is station dependent. Perhaps the smaller PE values at Sofia and Uppsala data are due to non-linear effects from irregularities in the equatorial and subauroral regions.

4. Conclusions

The modified auto-correlation model predicts deviations of f_0F_2 from its monthly medians using a regressive formulation based on a simple exponential expression for the auto-correlation function. Predictions up to 24 h ahead are obtained by calculating T from a best fit to the

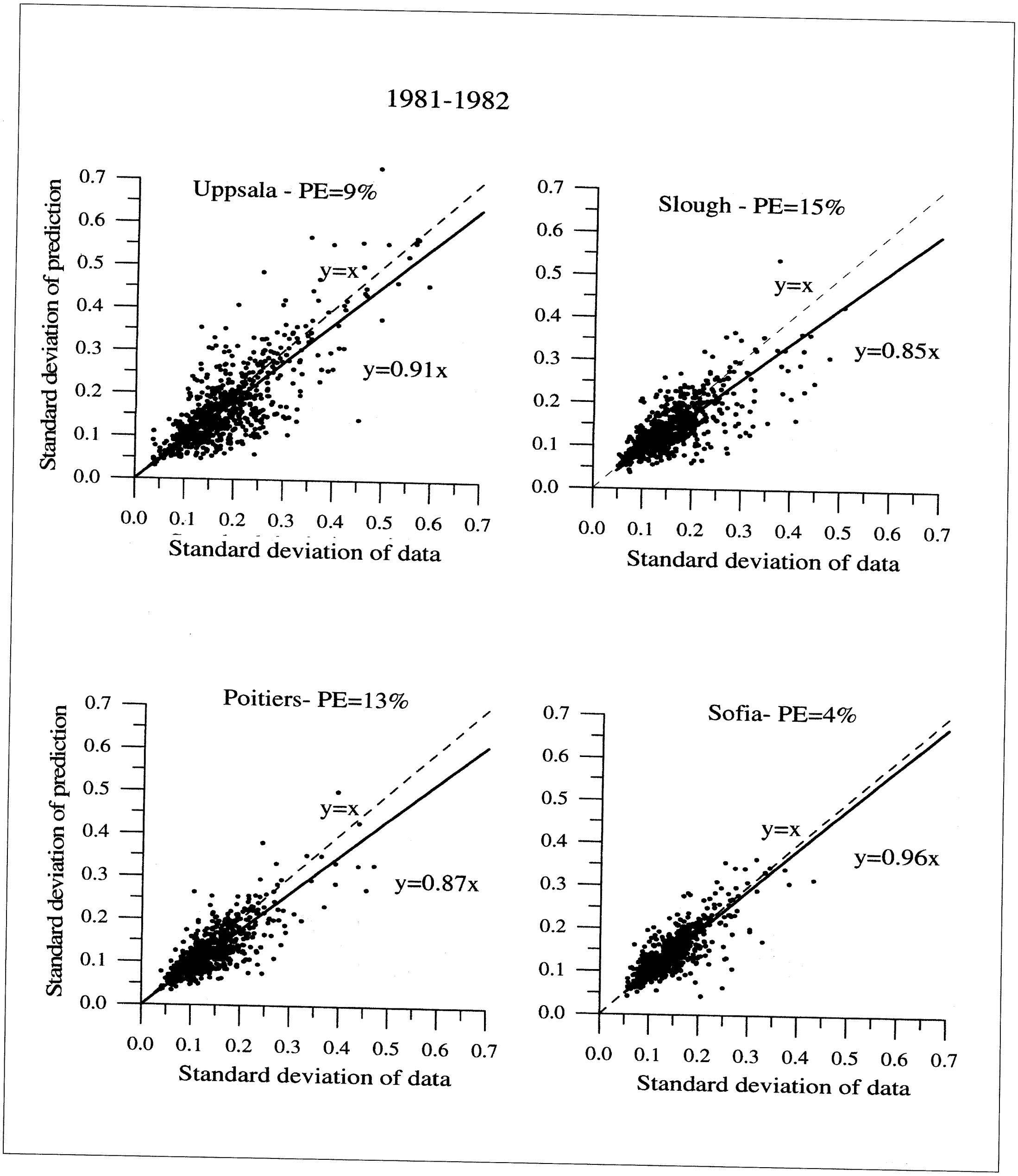

Fig. 3. The daily average model prediction error of the one-day prediction collected from the whole years of data 1981 and 1982 (dimensionless ordinate axis) are plotted against the respective daily average monthly median prediction error (dimensionless abscissa axis) for Uppsala, Slough, Poitiers and Sofia. The dashed lines mark the $y = x$, the full lines show the regression. The deviation of slopes between the two lines in percents is defined as «Prediction Efficiency» (PE).

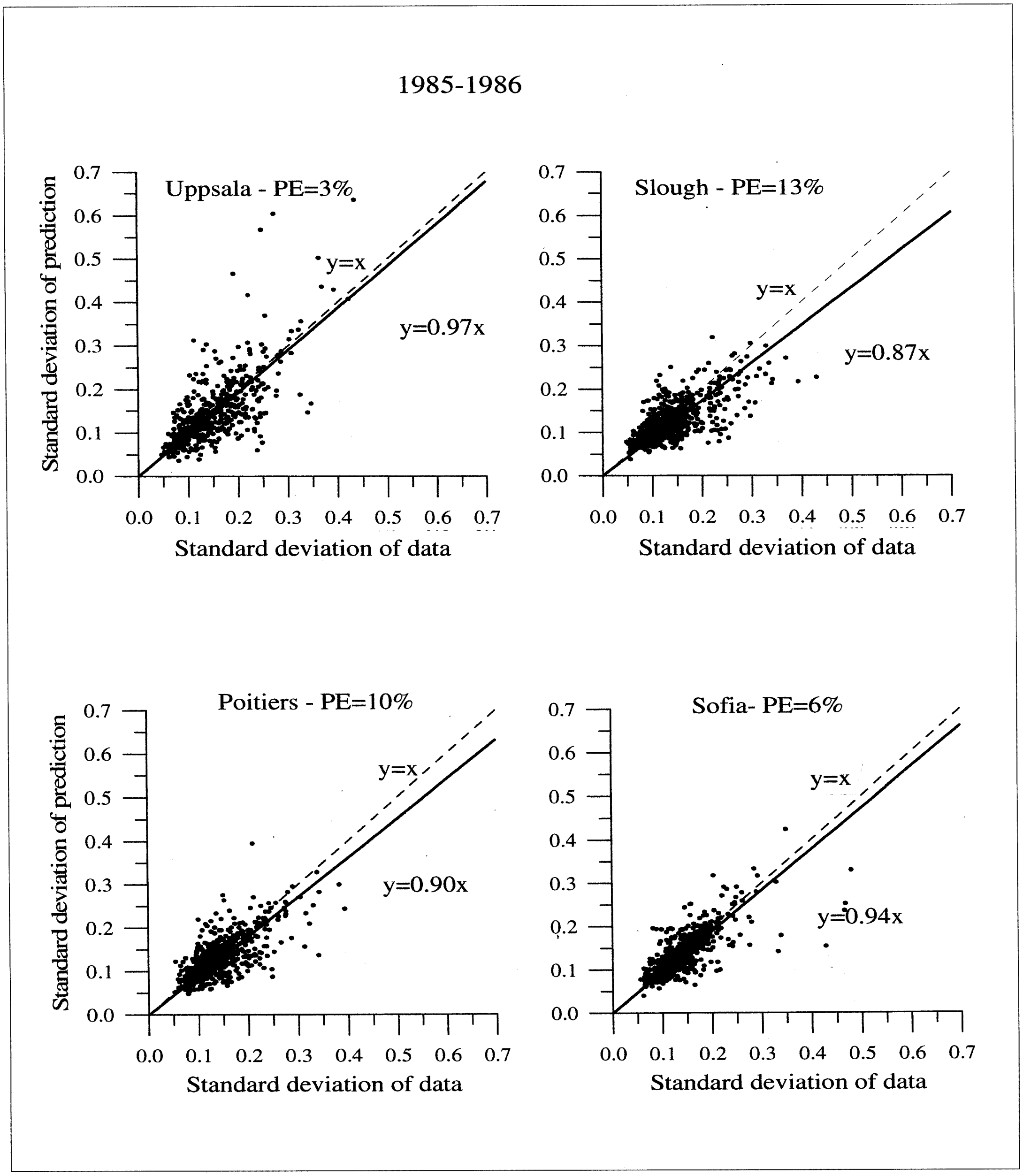

Fig. 4. The same as in fig. 3, for the years 1985 and 1986.

empirical auto-correlation function computed over the last 25 days. The model has been tested on data from Uppsala, Slough, Poitiers and Sofia, showing that this method gives better results than the monthly median prediction.

REFERENCES

CANDER, LJ.R., M.M. MILOSAVLJEVIC, S.S. STANKOVIC and S. TOMASEVIC (1998): Ionospheric forecasting technique by artificial neural network, *Electron. Lett.*, **34**, 1573-1574.

DICK, M.I., LJ.R. CANDER, M.F. LEVY, I.D. FINCH, I. KUTIEV and P. MUHTAROV (1998): Short-term ionospheric forecasting over Europe, in *Proceedings of the 2nd COST 251 Workshop «Algorithms and Models for COST 251 Final Product»,* edited by A. VERNON, COST251TD(98)005, 30-31 March 1998, Side, Turkey, 252-256.

FULLER-ROWELL, T.J., M.V. CODRESCU, H. RISHBETH, R.J. MOFFETT and S. QUEGAN (1996): On the seasonal response of the thermosphere and ionosphere to geomagnetic storms, *J. Geophys. Res.,* **101**, 2343-2353.

KORN, G.A. and T.M. KORN (1968): *Mathematical Handbook* (McGraw-Hill Book Co.), second edition, paragraph 11.3.

MUHTAROV, P. and I. KUTIEV (1999): Autocorrelation method for temporal interpolation and short-term prediction of ionospheric data, *Radio Sci.*, **34** (2), 459-564.

MUHTAROV, P., LJ.R. CANDER, M.F. LEVY and I. KUTIEV (1998): Application of the geomagnetically correlated statistical model to short-term forecasting of f_0F_2, in *Proceedings of the 2nd COST251 Workshop «Algorithms and Models for COST251 Final Product», 30-31 March 1998, Side, Turkey*, 241-245.

(received October 20, 1998;
accepted January 11, 1999)