

# Estimating uncertainties in incoherent scatter radar parameters from random variations in time series data

C. J. Davis and I. W. McCrea

Rutherford Appleton Laboratory, Chilton, OX11 0QX, UK

Received: 24 June 2003 – Revised: 28 July 2004 – Accepted: 20 August 2004 – Published: 3 November 2004

**Abstract.** The variation of parameters in time series data from the EISCAT UHF incoherent scatter radar has been used to estimate the uncertainties in measurements of electron concentration, electron and ion temperature and line-of-sight ion velocity. Good agreement was found with the results of previous analyses that estimated uncertainties in velocity and electron temperature from similar data using a complementary method. By contrast, uncertainties in ion temperature and electron concentration estimated from the same time series data were found to differ from the expected theoretical values by factors of 1.8 and 2.7, respectively. It was found that the relative sizes of these uncertainties more closely matched those predicted by a previous Monte Carlo simulation than values predicted by currently accepted theoretical formulae. Although the specific measured and theoretical uncertainties reported here relate to a very simple kind of ISR experiment, the method itself is general and can be applied to data from any incoherent scatter radar.

**Key words.** Ionosphere (auroral ionosphere; instruments and techniques). Radio science (instruments and techniques)

## 1 Introduction

The estimation of uncertainties in the parameters measured by incoherent scatter radar has been the subject of much discussion in the scientific literature (e.g. Evans, 1969; du Castel and Vasseur, 1972; Rishbeth and Williams, 1985; Jones et al., 1986; Breen et al., 1996 and Williams et al., 1996). The four most recent of the above studies focused on determining a theoretical formulation for such uncertainties as a function of the signal-to-noise ratio (SNR) in incoherent scatter radar (ISR) data. The data used in these studies came from some of the earliest EISCAT observing programmes, in which autocorrelation functions of the ionospheric backscatter from uncoded radar pulses were computed using particularly simple signal processing algorithms. In these data, each transmitter frequency was handled separately, and each autocorrelation function was derived from a separate set of range-

gated samples, so that a given signal or background gate was completely independent of its neighbours. This greatly simplified the derivation of a theoretical formula for calculating the random errors expected in such data (see the discussion in Williams et al., 1996).

Modern incoherent scatter programmes are now considerably more complex, using a variety of coding techniques and having much greater flexibility in the way that signal samples are processed. For example, backscatter from several frequencies may be pre-combined together within the radar correlator, in such a way that the original information can no longer be recovered. A given sample can also be used in the computation of multiple autocorrelation functions and the range gating strategy, rather than being fixed, can be specified by the data analyst during the post-processing phase. These techniques have many advantages, but one effect of this complexity is that the expected uncertainties in the plasma parameters derived from such data are much more difficult to predict from first principles. The functional form of their relationship to the SNR, however, generally remains the same as in the older experiments, except in a limited subset of cases (an example being phase-coded experiments, in which the self-noise of the code puts a lower limit on the achievable uncertainty at high SNR). The motivation for the current work was therefore to devise a general method of determining empirical formulae to characterise such uncertainties in any ISR data, largely irrespective of the exact nature of the observing programme. This would allow the uncertainties that are inherent in measurements made during these more complicated experiments to be quantified more easily than is currently the case.

In order to be certain that such empirical results were valid, it was first necessary to establish that the new method would also work on older data of the type used in the above-mentioned studies. The present study reports the results of deriving empirical relationships between uncertainty and SNR from such data by two different methods, based on time series and multiple frequency data, and the comparison of these with the earlier theoretical formulae.

The mainland EISCAT radar system (Folkestad et al., 1983; Rishbeth and Williams, 1985) consists of a monostatic VHF parabolic trough antenna and a fully steerable parabolic

UHF antenna both situated at Tromsø, Norway, together with two further UHF receiver sites at Kiruna, Sweden and Sodankylä, Finland, which complete the tristatic UHF system. This paper uses data from the Tromsø UHF radar, though many similar observing programmes were also run on the VHF system during the early years of EISCAT operation. As stated above, however, the method presented here is equally applicable to data from any ISR.

During the commissioning stages of the EISCAT radar, a Monte Carlo simulation study of the relationship between SNR and the uncertainties expected for the proposed radar system was carried out (Murdin, 1980). The simulated data were based on a very simple transmission and signal processing scheme of the type used in the earliest EISCAT programmes, as described above. The results of this simulation differed significantly from earlier theoretical formulae (as detailed in Rishbeth and Williams, 1985) in two areas. Firstly, the theoretical calculations predicted that, for a given SNR, the relative uncertainties in ion and electron temperatures would be of the same order, while the Monte Carlo simulation predicted that the relative uncertainties in ion temperature would be 1.33 times those in electron temperature. Secondly, the theoretical calculations suggested that relative uncertainties in the electron concentration would be half those of the electron and ion temperatures, while the Monte Carlo simulation predicted that these would instead be of the same order.

## 2 Uncertainty estimates using CP-4-A data

Previous authors (Williams et al., 1996; Breen et al., 1996) used data from a specific type of EISCAT programme, known as CP-4-A or POLAR. The CP-4-A programme used a simple correlator algorithm that enabled signals from six frequency channels to be recorded separately. Uncoded pulses of length  $500\mu\text{s}$  were used, and each autocorrelation function was computed from an independent set of samples. All six channels made measurements of the same scattering volume at very similar times (within a few ms). This enabled an estimate of the uncertainty in each parameter to be made by studying the spread of recorded values in the six frequency channels. Such analysis was used to test the validity of the theoretical formulae that predict a relationship between uncertainty of the measurements and the signal-to-noise ratio (SNR). Results were published which showed favourable comparisons between theory and measurement for line-of-sight plasma velocity (Williams et al., 1996) and electron temperature (Breen et al., 1996). The currently accepted theoretical formula for the uncertainties in EISCAT velocity data,  $\delta V_p$ , from such observing programmes is (Williams et al., 1996):

$$\delta V_p = 1.17\sqrt{\lambda} \left[ 1 + \frac{2}{R} \right] \frac{T_i^{0.25}}{(\tau r t)^{0.5}} \left( 1 - 0.16 \frac{T_e}{T_i} \right) \left( 1 + \frac{7.10^{-3}\lambda}{\tau\sqrt{T_i}} \right) \sqrt{1 + \frac{1}{m(1+R)^2}}, \quad (1)$$

where  $\lambda$  is the wavelength,  $R$  is the SNR over the ion-line bandwidth,  $T_i$  is the ion temperature,  $T_e$  is the electron temperature,  $\tau$  is the pulse length,  $r$  is the pulse repetition frequency,  $t$  is the integration time and  $m$  is the number of independent background gates.  $R$  is related to the SNR over the filter bandwidth by the expression

$$R = R_f \frac{B_f}{B} = R_f B_f \frac{\lambda}{4.8\sqrt{8k_B T_i}}, \quad (2)$$

where  $R_f$  is the SNR over the filter bandwidth  $B_f$ ,  $m_i$  is the ion mass,  $k_B$  is Boltzmann's constant and  $T_i$  the ion temperature.

Likewise, the most recent theoretical estimation of the relative uncertainties in electron temperature for this type of EISCAT programme (Breen et al., 1996) can be expressed as:

$$\frac{\delta T_e}{T_e} = \frac{\delta T_i}{T_i} = 2.30 \sqrt{\frac{\left( 1 + \frac{2}{R} + \frac{m+1}{m} \frac{1}{R^2} \right)}{(Bnt\tau)}}. \quad (3)$$

This formulation came about after considering a CP-4-A type analysis (comparing the variation of the parameter measured simultaneously over 6 different frequencies) and by studying the difference between points measured in a time series of data. For the latter technique, these authors used data integrated over several minutes and minimised geophysical variation between data points by choosing times where geomagnetic activity was low. Similar relationships between SNR and uncertainties in the various plasma parameters were also published by Baron (1977), who considered data from the Chatanika radar. One might intuitively expect uncertainties in ion and electron temperatures to show a similar dependence on SNR, since variations in these two parameters affect the incoherent scatter to a similar extent (though in different ways).

## 3 Using variation in time series data to estimate uncertainties in EISCAT parameters

The method of estimating uncertainties in EISCAT data presented in this paper is similar to the method used by Breen et al. (1996) with the following differences:

1. The data used were analysed at the minimum possible integration period of 10 s (this is the time resolution of data output from the correlator in the programmes used). This maximised the amount of variability due to random noise in the system while minimising the possibility of introducing significant variation due to real changes in the ionosphere over the time scale considered.
2. Since only an empirical relation was sought, data were binned and plotted against  $\text{SNR}^{-0.5}$ , as this allowed for terms in  $\text{SNR}^{-0.5}$  and  $\text{SNR}^{-1}$  to be considered when fitting a polynomial to the data. The nature of Eqs. (1) and (3) suggests that both such terms may be significant.

The basis of the method is to assume that any difference between data points is entirely due to random noise. The variance in each individual measurement can then be estimated from this difference using the relation:

$$\delta V_p^2 = \frac{\sum_{i=1}^n (V_{pi} - \langle V_p \rangle)^2}{(n-1)}. \quad (4)$$

where  $\langle V_p \rangle$  is the mean of the two velocity measurements. This variance should, of course, be larger for data with low SNR and smaller for data with high SNR. For  $n=2$ , Eq. (4) reduces to:

$$\delta V_p^2 = \frac{(V_{p2} - V_{p1})^2}{2}, \quad (5)$$

Although the plasma velocity is used in the above example, the same equations apply equally to the other plasma parameters. In order to provide a cross reference with previous analyses, data from the CP-4-A run from October 1989 were used. These data were analysed at a ten-second resolution, with data from each individual frequency analysed separately. At this resolution, it was only necessary to use a subset of the data to provide an adequate number of data points. Data from 2–3 October were used in this study, providing around ten thousand data points per frequency channel, which was considered more than adequate for the statistical analysis.

Having been analysed, estimates of the relative uncertainty in electron concentration,  $N_e$ , electron and ion temperatures,  $T_e$  and  $T_i$  and of the absolute uncertainty in line-of-sight velocity,  $V_p$ , were calculated using Eq. (5) for overlapping pairs of consecutive points in each time series (e.g. 12, 23, 34, etc.). While it would have been possible to use any interval between data points, or indeed Eq. (4) directly, to estimate the variance in plasma parameter measurements in the absence of any geophysical variation, the use of adjacent points was adopted to minimise the possibility of such variation, and in order to investigate the dependence of the variance on SNR. For each observation, the SNR is known and so the variance estimates were binned according to the concomitant values of  $\text{SNR}^{-0.5}$ . Given the large amount of data it was possible to be strict when considering data suitable for this comparison and no pair of data points was used whose associated SNR values lay in different bins. The standard deviation was then calculated for each bin. In order to ensure that the mean estimate was not influenced by extreme values, the outliers were first removed using Chauvenet's criterion (e.g. Gumbel, 1958). A polynomial in terms of  $\text{SNR}^{-0.5}$  was then fitted to the mean values, using the standard deviations on each mean as weights to the fit. This ensured that bins containing few data points did not have an unnecessarily large effect on the result (bins with fewer than 20 points were discarded). In this way, an empirical relationship between the uncertainty of a parameter and  $\text{SNR}^{-0.5}$  was estimated.

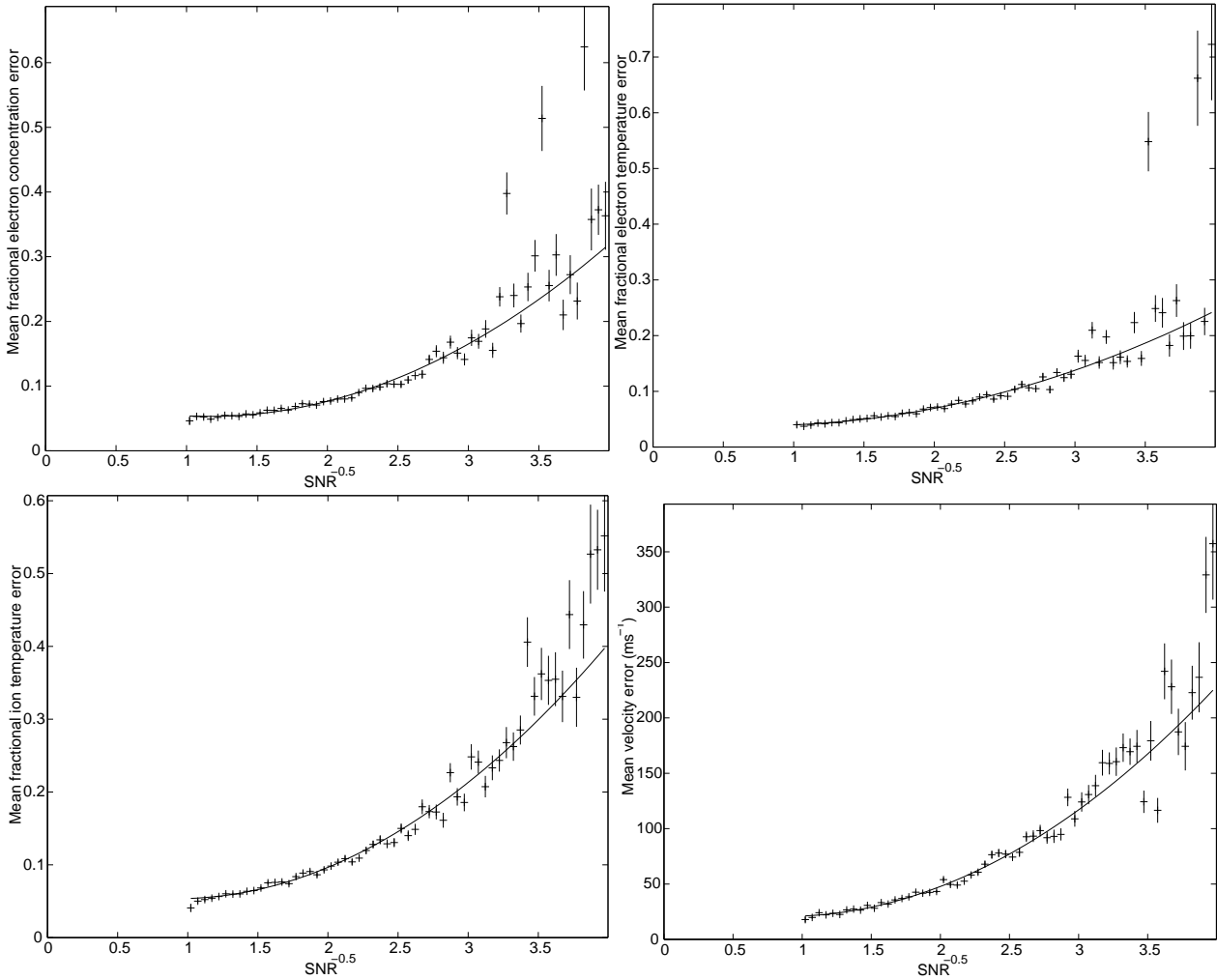
The rigorous way of determining the appropriate order of polynomial to fit to a particular data set involves fitting a

range of polynomial orders to the data and assessing the suitability of each by using the  $\chi^2$  test, with the polynomial order producing the lowest  $\chi^2$  value taken to be most representative of the data. In this case, however, the differences in  $\chi^2$  values produced by second and third order polynomial fits were so small that it was decided to use the lower order polynomial, as this could be justified in terms of the theoretical formulae. In the case of uncertainties in electron concentration and electron and ion temperatures, the theoretical formulae are expressed in terms of fractional errors and so these were used in estimating empirical relationships for these parameters. Using the second-order polynomial coefficients, empirical values were calculated for each bin and these were then compared with the mean theoretical uncertainty calculated for the data points in each bin.

## 4 Results

Plots of the error terms  $\delta N_e/N_e$ ,  $\delta T_e/T_e$ ,  $\delta T_i/T_i$  and  $\delta V_p$  against  $\text{SNR}^{-0.5}$ , are shown in Fig. 1. These plots show values calculated using time series data from one of the six frequencies. It can be seen that there are obvious trends for all parameters, justifying the fitting of a polynomial to the data. The vertical error bars, used to provide some measure of weight to the fit, represent the error in the mean for each SNR bin. In general, bins corresponding to the lowest SNR have fewest points and so are given a lower weighting. The coefficients of the fitted polynomials were then used to generate formulae for the “observed” uncertainties in each of the parameters and the values of  $\delta N_e/N_e$ ,  $\delta T_e/T_e$ ,  $\delta T_i/T_i$  and  $\delta V_p$  were compared with those obtained using the empirical formulae of Breen et al. (1996) and Williams et al. (1996) respectively. The relative uncertainty in electron temperature was assumed to be the same as that in ion temperature, while the relative uncertainty in electron density was assumed to be half of this level, following the theoretical formulae given in Rishbeth and Williams (1985).

Figure 2 shows the comparison between observed and theoretical uncertainties for each parameter. The same trends were seen in each of the six frequencies used from the observing programme, but only sample plots from a single frequency are shown. The deviation from a linear trend towards high SNR (low  $\text{SNR}^{-0.5}$ ) in some of the plots is an indication that the polynomial does not represent the variation with SNR very well at these values. Second order polynomials were used, as they could be justified in terms of the theoretical uncertainty estimates. The fact that they do not represent the data for high SNR ratios may be an indication that there is some additional factor that is currently not accounted for in the theoretical calculations. Blind application of these second-order polynomial coefficients to reconstruct the uncertainties in EISCAT data from the SNR alone would misrepresent the accuracy of data with a high SNR. A linear regression was therefore carried out in order to estimate the gradient of each of these comparison plots. This gradient can then be used to scale the values predicted by the theoretical



**Fig. 1.** Second-order polynomial fits to the uncertainty levels for the four main EISCAT parameters binned according to  $\text{SNR}^{-0.5}$  for a sample frequency channel (1). The error bars represent the standard deviations of the distributions in each bin and are used to weight the fit. The four panels show the relative uncertainty in electron concentration, electron temperature, ion temperature, together with the uncertainty in line-of-sight velocity.

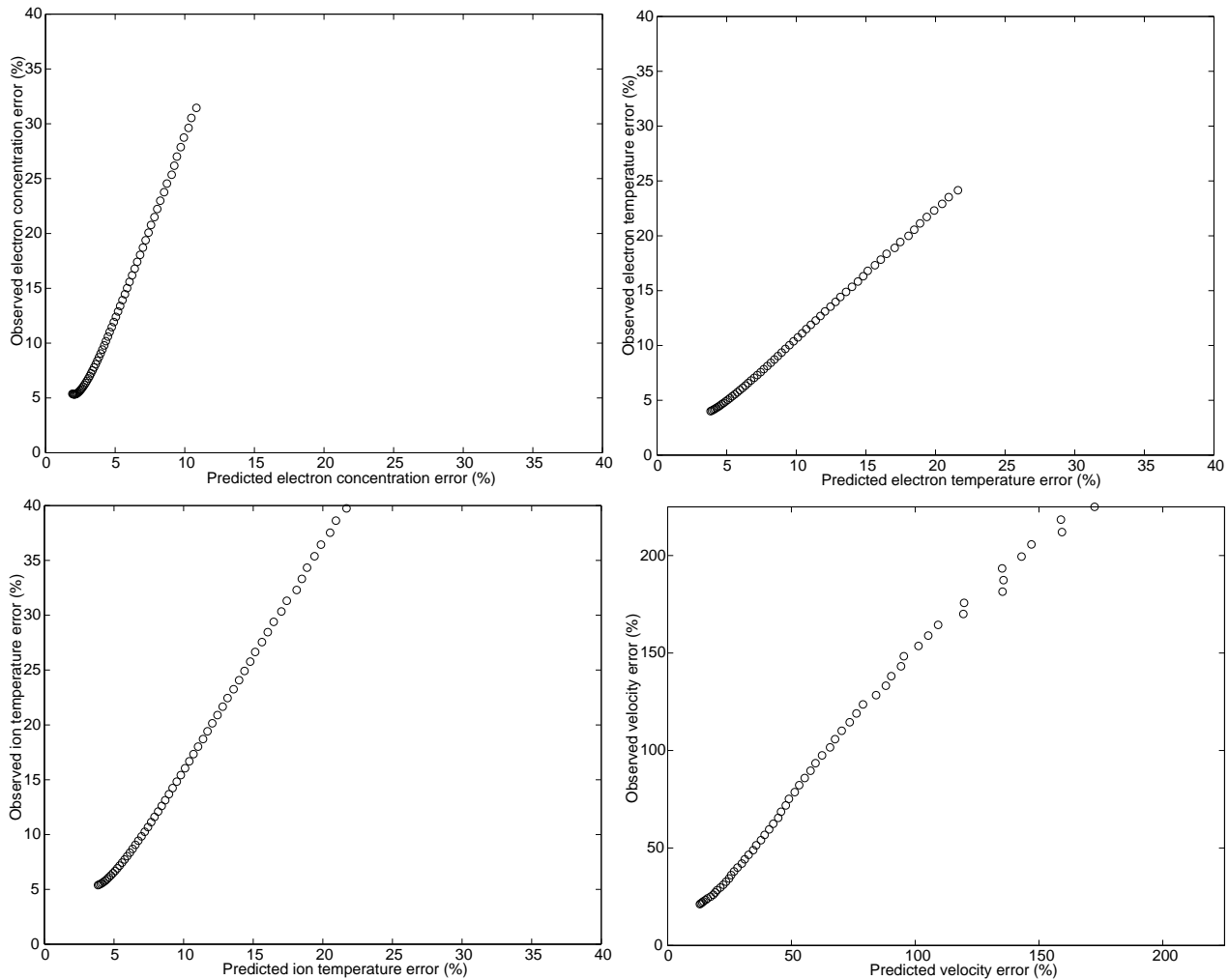
formulae. The six scaling factors obtained in this way, one for each frequency channel, are summarised in Table 1. If the observed uncertainty matched the theoretical uncertainty, this would produce a gradient of 1. As can be seen from the values shown in Table 1, the trends for each parameter are consistent but the uncertainty differs from the theoretical prediction by different amounts for each parameter.

The results in Table 1 show that the relationship between SNR and uncertainties in electron temperature is consistent with that calculated by Breen et al. (1996). In contrast, the results for ion temperature uncertainties differ from the predicted values by a factor of 1.8 while the empirical estimates of uncertainties in electron concentration differ from the expected theoretical values by a factor of 2.7.

Table 1 also demonstrates that, for any given signal-to-noise ratio, the uncertainty in line-of-sight velocity as estimated from the polynomial fit is around 1.5 times that pre-

dicted by Williams et al. (1996). Such a difference was predicted, as the current analysis used a matched filter technique to estimate the velocity. According to the table of values quoted by Williams et al. (1996), this should have resulted in a 35 to 50% difference between the two methods (for the value of the electron/ion temperature ratio measured in the data used).

As far as the authors are aware, there are no previously published studies of uncertainties in electron concentration as a function of SNR using actual measured data from incoherent scatter radar programmes. In order to provide a second estimate of this relationship, the data from all six frequencies of the CP-4-A programme were employed, using simultaneous results from each of the six channels rather than time series measurements on a single channel. This is essentially the same method as that adopted by Williams et al. (1996) to estimate uncertainties in the velocities.



**Fig. 2.** A comparison of observed uncertainties (calculated from the polynomial fits) against predicted uncertainties (derived for the same bins of data using the theoretical formulae quoted in this paper) for the same sample channel as used in Fig. 1. As for Fig. 1, the panels show the fractional uncertainties in electron concentration, electron temperature and ion temperature and the uncertainty in line-of-sight velocity.

An initial analysis produced very poor results, with no linear trend visible when comparing observed and predicted uncertainties. This was caused by each of the six channels being offset from each other by a small amount due to calibration differences. In order to minimise this effect, the electron concentrations measured by all six channels were calibrated to the first by a least-squares fitting and the procedure was repeated. This resulted in a much improved linear relationship between the observed and predicted uncertainties, with a slope of 1.86. Despite the calibration problems with this observing programme (which cannot be completely removed), this is strong evidence that, for any given value of SNR, the relative uncertainty in electron concentration is close to twice the level quoted by Rishbeth and Williams (1985) (i.e. it is comparable with the uncertainty in electron temperature). While the time series and multi-frequency methods give different results (gradients of 2.7 and 1.9 with respect to theory), both methods suggest that the relative uncertainty in electron

concentration should be larger than the current theoretical formulation. More confidence should probably be given to the method of estimating uncertainties from the difference between adjacent points by the use of time series data. This technique is not subject to calibration errors, since it uses data from a single frequency channel. As stated earlier, many modern ISR experiments (e.g. Turunen, 1986) pre-combine data from multiple channels, which is a perfectly valid procedure if the signal levels in each channel are comparable. In this case, the time series method would be the only one available.

## 5 Discussion

It is clear that, for some of the derived plasma parameters, the relative sizes of the uncertainties calculated using the method outlined in this paper disagree with values derived using the currently accepted theoretical formulae. In particular, the

**Table 1.** Ratios between observed and predicted uncertainty levels for the four main EISCAT derived parameters. These values are given by the gradients of least-squares fits to the data from six separate frequency channels in the CP-4-A data taken by the Tromsø UHF radar between 2–3 October 1989.

	Ratio of observed to predicted uncertainty levels						Mean Ratio
	Channel 1	Channel 2	Channel 3	Channel 4	Channel 5	Channel 6	
$N_e$	3.04	2.89	2.73	2.59	2.52	2.56	2.72±0.20
$T_e$	1.16	1.15	1.05	1.02	1.09	0.99	1.08±0.07
$T_i$	1.99	1.84	1.77	1.66	1.76	1.61	1.77±0.13
$V_p$	1.39	1.40	1.40	1.51	1.53	1.52	1.46±0.07

ratio of the relative uncertainties in electron concentration and ion and electron temperatures differs between the two methods. The uncertainties in those parameters that share related theoretical error formulae, namely  $N_e$ ,  $T_e$  and  $T_i$ , can be compared directly with the predictions of the Monte Carlo simulation performed by Murdin (1980). The mean gradients from the polynomial fits from the current analysis (as detailed in Table 1) produce values for  $N_e$ ,  $T_e$  and  $T_i$  with ratios of 1.24, 1.0 and 1.63 with respect to  $T_e$ . In comparison, the Monte Carlo simulation predicted ratios in the uncertainties of these same parameters of 1.22, 1.0 and 1.33. According to Murdin's predictions, which apply to precisely the kind of ISR experiments whose data have been used here, the relative uncertainties in electron concentration should not be a factor two less than those in electron temperature, and the relative uncertainty in electron temperature measurements should be lower than for the ion temperature. Of these, the most striking result is the relative size of the uncertainty in electron concentration. These predictions seem to have been validated by the present study.

The initial aim of the current work was to evaluate the accuracy of radar measurements using differences between consecutive values in time series data. Theoretical formulae have been published, which predict error levels in electron temperature and ion velocity (Breen et al., 1996, Williams et al., 1996). These methods mainly relied on comparison of simultaneous data from several different frequencies. The CP-4-A data provided a useful set for comparison, since it could be used to estimate the uncertainties from both time series and multi-frequency methods. In addition, empirical relationships between the SNR and relative random uncertainties for electron concentration and ion temperature have been derived which were not in the previously published literature. The relative magnitudes of these uncertainties appear to agree closely with the simulation results of Murdin (1980). Murdin's calculations were very specific and applied only to ISR observations made with uncoded pulses and using simple correlator programs, such as the EISCAT CP-4-A data used in the present analysis. A good test to confirm that Murdin's simulations have wider validity would be to repeat them for a radar programme using a modern correlator algorithm. The results of such a simulation could then be tested for more recent data sets, using the method of estimating uncertainties from the differences between adjacent points in time series data.

## 6 Conclusions

Based on an analysis of data from the CP-4-A programme, it has been demonstrated that time series data can be used to generate empirical formulae for the relationship between signal-to-noise ratio and uncertainties in plasma parameters measured by incoherent scatter radar. These relationships for line-of-sight ion velocity and electron temperature broadly agree with the results of Williams et al. (1996) and Breen et al. (1996), who used similar data. The present study has extended their work by also investigating relative uncertainties in ion temperature and electron concentration. Our data provide evidence to support the simulation carried out by Murdin (1980), whose results are significantly different from those obtained using currently accepted theoretical relationships. In particular, the relative error in electron concentration was found to be almost double that predicted by Rishbeth and Williams (1985). In the future, we plan to extend this work to more modern radar programmes with modern coding schemes and more complex correlator algorithms. The results of such a study would be of great importance to the interpretation of all recent incoherent scatter data.

*Acknowledgements.* This work was funded by the UK Particle Physics and Astronomy Research Council. EISCAT is supported by Finland, France, Germany, Japan, Norway, Sweden and the UK. The authors would like to thank G. J. Daniel for useful discussions.

Topical Editor M. Lester thanks B. Isham for his help in evaluating this paper.

## References

- Breen, A. R., Williams, P. J. S., Etemadi, A., and Davda, V. N.: Uncertainties in measurements of electron temperature and the estimation of  $F$ -region electron heat conduction from EISCAT data. *J. Atmos. Terr. Phys.*, 58, 145–159, 1996.
- Baron M. J: The Chatanika radar system, in "Radar Probing of the auroral plasma", Ed. A. Brekke, Universit tsforlaget, Troms , 1977.
- du Castel, F. and Vasseur, G.: Evaluation des Performances d'un sondeur ionospherique   diffusion incoherente, *Ann. Telecommun.*, 27, 239–247, 1972.
- Evans, J. V.: Theory and practice of ionosphere study by Thomson scatter radar, *Proc. IEEE.*, 57, 496–530, 1969.
- Folkestad, K., Hagfors, T., and Westerlund, S.: EISCAT—an updated description of technical characteristics and operational capabilities, *Radio Sci.*, 18, 867–879, 1983.

- Gumbel, E. J.: *Statistics of Extremes*, Columbia University Press, 1958.
- Jones, G. O. L., Winser, K. J., and Williams, P. J. S.: Measurements of plasma velocity at different heights along a magnetic field line, *J. Atmos. Terr. Phys.*, 48, 887–892, 1986.
- Murdin, J.: *Incoherent Scatter Radar Measurements in the Auroral Zone*, KGI report No. 80:2, 1980.
- Rishbeth, H. and Williams, P. J. S.: The EISCAT ionospheric radar: The system and its early results, *Q. J. R. Astron. Soc.*, 26, 478–512, 1985.
- Turunen, T.: GEN-SYSTEM - a new experimental philosophy for EISCAT radars, *J. Atmos. Terr. Phys.*, 48, 777–786, 1986.
- Williams, P. J. S., Etemadi, A., McCrea, I. W., and Todd, H.: Errors due to random noise in velocity measurement using incoherent-scatter radar, *Ann. Geophys.* 14, 1480–1486, 1996.