EFDA–JET–PR(12)12

M. Maslov, M.N.A.Beurskens, J. Flanagan, M. Kempenaars
and JET EFDA contributors

Statistical Error Bars Estimation for Thomson Scattering Diagnostics

Statistical Error Bars Estimation for Thomson Scattering Diagnostics

M. Maslov[1], M.N.A.Beurskens[1], J. Flanagan[1], M. Kempenaars[1]
and JET EFDA contributors*

JET-EFDA, Culham Science Centre, OX14 3DB, Abingdon, UK

[1]*EURATOM-CCFE Fusion Association, Culham Science Centre, OX14 3DB, Abingdon, OXON, UK*
** See annex of F. Romanelli et al, "Overview of JET Results",*
(23rd IAEA Fusion Energy Conference, Daejon, Republic of Korea (2010)).

Preprint of Paper to be submitted for publication in
Review of Scientific Instruments

ABSTRACT.
Thomson scattering is a routine diagnostic for measuring electron temperature and density on almost every magnetic confinement fusion plasma device. Due to very low cross section of the scattering, the measured signal to noise ratio is generally relatively low and statistical error bars of a single measurement often dominate over systematic uncertainties caused for example by imprecise calibration. In this work an analytical expression is introduced to determine statistical error bars for temperature, density and pressure measurements.

1. INTRODUCTION

Various Thomson Scattering diagnostics were built at many fusion devices around the world. While methods of data analysis, i.e. derivation of temperature and density from measured scattered light spectrum is not changing much from case to case, estimation of measurement uncertainties (so called error bars) can be quite different. The simplest estimate for density error bars is for example $\sigma(n) = (1/N_{PE})^{1/2}$, where N_{PE} is the total number of photoelectron events detected in all spectral channels. Another methods take into account χ^2 values of the least mean square, or min (χ^2) fit used for determination of temperature, i.e. using the information how well the measured spectrum fits into a theoretical Thomson Scattering spectrum. This method works reasonably well in cases where scattered light spectrum is fitted with multiple points, for example when the light detector is a CCD camera with hundreds of pixels [3]. The most complex and probably the most accurate method is Monte-Carlo simulation, i.e addition of a noise to measured signals and performing the fit multiple times to derive standard deviation of the output T_e and n_e values [3,4].

The purpose of this work is to derive an analytical expression for calculating statistical uncertainties of measured T_e and n_e, based only on signal-to-noise ratios in individual spectral channels and the spectrometer calibration data. This would allow to acquire accurate statistical error bars for arbitrary T_e, n_e, signal level and background noise without involving time-consuming Monte-Carlo calculations.

2. DERIVATION OF THE ERROR BARS

A generic Thomson Scattering diagnostic is measuring scattered light amplitude in several spectral channels. During a calibration process, the amount of light expected in each channel for every T_e value for a given density is determined, therefore the temperature and density measurement is basically a finding of the best match between measured and expected signals in all channels. Mathematically it means finding the minimum of the following value:

$$\chi^2 = \sum_{i=1..n} w_i [S_i - N \times F_i]^2 \qquad (1)$$

Here n is the number of spectral channels, S_i is the signal measured in each of the channels, F_i is the expected signal in each channel for unit density as a function of temperature, $F_i = F_i(T)$, N is the plasma electron density, w_i is a weighting factor for each channel measurement. Weighting is usually defined as $w_i = 1/\sigma^2(S_i)$, where $\sigma(S_i)$ is the uncertainty of signal measured in a particular channel, but for the purpose of this work we can consider it as an arbitrary non-negative number. Function $F_i(T)$ is a product of a spectral function of a particular channel and the scattered light spectrum (can be calculated using formulas of [1] for example), integrated over the whole wavelength range. Spectral functions are determined via calibration of the diagnostic and the functions $F_i(T)$ then calculated numerically for any given temperature value. Spectral calibration and the expected signals F_i for JET LIDAR Thomson scattering diagnostics [2] are shown on figures 1 and 2 as an example. χ^2 in expression (1) is a 2[nd] order polynomial function of N therefore for every given T value it has only one minimum at

$$N = \frac{\sum_{i=1..n} w_i S_i F_i}{\sum_{i=1..n} w_i F_i^2} \tag{2}$$

So for every temperature value, there is only one density which satisfies the χ^2 = min requirement, therefore N can be considered not as an independent variable but as a function of temperature, N = N(T), so χ^2 becomes a function of only one variable T.
Finding the minimum of χ^2 is equivalent to solving the equation:

$$\frac{d\chi^2}{dT} = -2\sum_{i=1..n} w_i (S_i - N^* F_i)(\frac{dN}{dT}F_i + \frac{dF_i}{dT}N) = 0 \tag{3}$$

To simplify (3) and other expressions used here, we define the following values:

$$A = \sum w_i S_i F_i\,, \quad B = \sum w_i F_i^2\,, \quad C = \sum w_i S_i \frac{dF_i}{dT}, \quad D = \sum w_i F_i \frac{dF_i}{dT}, \quad \theta = \frac{\mathrm{d}\chi^2}{\mathrm{dT}} \tag{4}$$

Then $N = \frac{A}{B}$, $\frac{dN}{dT} = \frac{C}{B} - 2\frac{AD}{B^2}$ and the expression (3) transforms to

$$\theta = -2\sum_{i=1..n} w_i (S_i - \frac{A}{B}F_i)(\frac{C}{B}F_i - 2\frac{AD}{B^2}F_i + \frac{A}{B}\frac{dF_i}{dT}) =$$

$$= -2\sum_{i=1..n} w_i (\frac{C}{B}S_i F_i - 2\frac{AD}{B^2}S_i F_i + \frac{A}{B}S_i\frac{dF_i}{dT} - \frac{AC}{B^2}F_i^2 + 2\frac{A^2 D}{B^3}F_i^2 - \frac{A^2}{B^2}F_i\frac{dF_i}{dT}) = \tag{5}$$

$$= -2(\frac{AC}{B} - 2\frac{A^2 D}{B^2} + \frac{AC}{B} - \frac{AC}{B} + 2\frac{A^2 D}{B^2} - \frac{A^2 D}{B^2})\text{, or finally}$$

$$\theta = 2\frac{A}{B}(\frac{AD}{B} - C)$$

As it was already mentioned, whichever data processing routine is used to analyse a particular

Thomson Scattering data, it will end up finding the minimum of χ^2 or equivalently solving the $\theta = 0$ equation. The resulting error bars on temperature, density and pressure will be equal to:

$$\sigma(T) = \sqrt{\sum_{i=1..n} \left(\frac{dT}{dS_i} \sigma(S_i) \right)^2} \tag{6}$$

$$\sigma(N) = \sqrt{\sum_{i=1..n} \left(\frac{dN}{dS_i} \sigma(S_i) \right)^2} \tag{7}$$

$$\sigma(P) = \sqrt{\sum_{i=1..n} \left(\frac{dP}{dS_i} \sigma(S_i) \right)^2} \tag{8}$$

Here $\sigma(S_i)$ is the standard deviation of the measured signal in each spectral channel. It is usually a combination of detector intrinsic noise, quantum noise of plasma background and of the Thomson signal, but may be calculated differently for different system. *In this work we assume that $\sigma(S_i)$ is a known value or can be reasonably estimated by the diagnostic operator.*

dT/dS_i, dN/dS_i and dP/dS_i are the responses of the output temperature, density and pressure values produced by the data processing routine to deviation of the measured signals S_i. We will look at these values one by one.

2.1. TEMPERATURE

Remember that fitting routine is solving the $\theta = 0$ (3) equation for the input S_i values. Therefore for any variation ∂S_i, the output temperature T will change to $T + \partial T$ in a way that the $\theta = 0$ condition will be preserved. Therefore

$$\frac{dT}{dS_i} = -\frac{\partial\theta / \partial S_i}{\partial\theta / \partial T} \tag{9}$$

From the expression (5) and from definition of the supplementary variables A,B,C,D (4) we can derive:

$$\frac{\partial\theta}{\partial S_i} = 2\frac{Aw_i}{B^2}(DF_i - B\frac{dF_i}{dT}) \tag{10}$$

The denominator of (9) can be calculated in a similar way from (5) taking into account that $dA/dT = C$. $dB/dT = 2D$, and that the $\theta=0$ condition must be fulfilled, i.e..

$$\frac{\partial\theta}{\partial T} = 2\frac{A}{B^2}(3CD + A\frac{dD}{dT} + B\frac{dC}{dT}) \tag{11}$$

Calculating $\partial\theta/\partial T$ as the expression (11) suggests can be done but probably excessive. In a real application it is suggested to calculate it numerically as $\partial\theta/\partial T = \theta\,(T + \Delta T)/\Delta T$ for a small temperature increment ΔT.

The final expression for $\sigma(T)$ is:

$$\boxed{\sigma(T) = \frac{1}{\partial\theta/\partial T}\frac{2A}{B^2}\sqrt{\sum_{i=1..n}\left(w_i\sigma(S_i)(DF_i - B\frac{dF_i}{dT})\right)^2}} \tag{12}$$

2.2. DENSITY

To calculate dN/dSi we will use definition N = A/B and remember that F_i is a function of temperature and therefore is a function of measured signals S_i as well, i.e. we need to calculate full derivatives accurately.

$$\frac{dA}{dS_i} = w_i F_i + \sum_{j=1..n} w_j S_j \frac{dF_j}{dT}\frac{dT}{dS_i} = w_i F_i + C\frac{dT}{dS_i}$$

$$\frac{dB}{dS_i} = 2\sum_{j=1..n} w_j F_j \frac{dF_j}{dT}\frac{dT}{dS_i} = 2D\frac{dT}{dS_k}\text{, so} \tag{13}$$

$$\frac{dN}{dS_i} = \frac{1}{B}\frac{dA}{dS_i} - \frac{A}{B^2}\frac{dB}{dS_i} = \frac{w_i F_i}{B} + \left(\frac{C}{B} - 2\frac{AD}{B^2}\right)\frac{dT}{dS_i} = \frac{w_i F_i}{B} + \frac{dN}{dT}\frac{dT}{dS_i}$$

and the final expression for density errors is:

$$\sigma(N) = \sqrt{\sum_{i=1..n}\left[\left(\frac{w_i F_i}{B} + \frac{dN}{dT}\frac{dT}{dS_i}\right)\sigma(S_i)\right]^2}$$

or using expressions (6) and (9):

$$\sigma(N) = \sqrt{\sum_{i=1..n}\left(\frac{w_i F_i \sigma(S_i)}{B}\right)^2 + \left(\frac{dN}{dT}\sigma(T)\right)^2 - \frac{2}{B}\frac{dN}{dT}\frac{1}{d\theta/dT}\sum_{i=1..n} w_i F_i \frac{d\theta}{dS_i}\sigma(S_i)} \tag{14}$$

2.3. PRESSURE

Since $P = N \times T$, its derivative is equal to $dP/dS_i = N \times dT/dS_i + T \times dN/dS_i$ therefore:

$$\sigma(P) = \sqrt{\sum_{i=}\left(N\frac{dT}{dS_i} + T\frac{dN}{dS_i}\right)\ \sigma^2(S_i)} =$$

$$= \sqrt{\left(N\sigma(T)\right)^2 + \left(T\sigma(N)\right)^2 + 2NT\sum_{i=1..n}\frac{dT}{dS_i}\frac{dN}{dS_i}\sigma^2(S_i)}$$

using (9) and (13) we derive the final expression:

$$\sigma(N) = \sqrt{\sum_{i=1..n}\left(\frac{w_i F_i \sigma(S_i)}{B}\right)^2 + \left(\frac{dN}{dT}\sigma(T)\right)^2 - \frac{2}{B}\frac{dN}{dT}\frac{1}{d\theta/dT}\sum_{i=1..n} w_i F_i \frac{d\theta}{dS_i}\sigma(S_i)} \tag{15}$$

Note that if N and T were independently measured values, we would have the following expression for the pressure errors:

$$\sigma^*(P) = \sqrt{(T\sigma(N))^2 + (N\sigma(T))^2} \tag{16}$$

The last two terms under square root in (15) represent the covariance of N and T, i.e. the fact that they are derived from the same Thomson Scattering measurements and are not independent values.

3. MONTE-CARLO SIMULATIONS

To validate the formulas derived here, a set of Monte-Carlo simulations was performed. Spectral calibration and $F_i(T)$ functions were taken from JET LIDAR Thomson Scattering as a model. Standard deviation of measured signal was estimated as $\sigma(S_i) = \sqrt{\sigma(el)^2 + Bck + S_i}$, where *σ(el)* is a typical detector noise, *Bck* is plasma background light signal in photoelectrons, S_i is measured signal in photoelectrons. Weighting coefficients are set to wi = 1/s2 (S_i).

Simulation was done for 100 temperature values in the range 0.2-11.0keV. Plasma density was fixed at $10^{19}m^{-3}$ which provided 400-550 photoelectron events over all 6 spectral channels, depending on T_e. Plasma background together with detector noise was fixed at 20 photoelectron events per channel.

For each initial Te value, expected signal S_i was reconstructed and then a random normally distributed noise with standard deviation of $\sigma(S_i)$ was added. The result signals were then sent to the LIDAR data processing code which performed the fit and derived temperature and density values in the same way as it does for the genuine measurement signals. This was performed 5000 times for every initial Te value and standard deviation of the fit results was compared with error bars estimate done with formulas (12), (14) and (15). Results can be found on figures 3, 4, 5. We can see that the formulas derived here are in very good agreement with the simulation results.

On figure 4 a simple estimation of density error bar, $dN/N = 1/\sqrt{N_{phot}}$ where N_{phot} is the total number of photoelectrons in all channels, is also plotted. This simple estimate is below the accurately calculated error bars, and the difference is bigger for lower temperatures.

On figure 5 a simple estimation for pressure errors (16) is also plotted. Remarkably, if one assumes that N_e and T_e are measured independently, thus use (16) instead of (15), it would lead in this case to underestimated accuracy of the measurements for the most of Te range 0-8keV

CONCLUSION

Analytical expressions for statistical error bars estimation of Te, Ne and Pe measured by Thomson Scattering diagnostic was derived in this work. Validity of these expressions were tested using Monte-Carlo simulations for the JET Core LIDAR system parameters and it was shown that as long as noise level of individual measurement channels is evaluated correctly, the formulas are accurate. Output error bars are calculated as the standard deviation of the measured values, i.e. they outline interval with 66% probability to find the real value within it.

This method of error bars estimation is used for public JET Core LIDAR data since 2011. Formulas are applicable to any Thomson Scattering diagnostic or any other measurement instrument which is using similar to (1) data processing routine.

As a final note, we should pay attention to the fact that in calculations we assumed that the weighting coefficients w_i are constants, while in the fitting process they are adjusted according to signal to noise ratio of a particular channel. Thus they are effectively functions of measured signals S_i and strictly speaking it must be taken into account when calculating d/dS_i derivatives. In a general case when $\sigma(S_i) << S_i$ this can be ignored (and as we've seen, results are in agreement with Monte Carlo simulation), but when dealing with exceptionally low signals as few photoelectrons per channel, more accurate treatment may be required.

ACKNOWLEDGEMENT

Most of this work was performed on JET while the main author was affiliated to "Centre de Recherches en Physique des Plasmas, Association EURATOM - Confédération Suisse, EPFL, 1015 Lausanne, Switzerland" This work was supported by EURATOM and carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

[1]. Naito, O., Yoshida, H., Matoba T., Physics of Fluids **B 5**(11), 1993

[2]. Gowers, C. W., Brown, B.W., Fajemirokun, H., Nielsen, P., Nizienko, Y., Schunke, B., Review of Scientific Instruments **66** (1995), doi: 10.1063/1.1146321

[3]. Beurskens, M.N.A., Barth, C.J., Chu C.C., Lopes Cardozo, N.J., Review of Scientific Instruments **70** (1999), doi: 10.1063/1.1149702

[4]. Beurskens, M.N.A., Giudicotti, L., Kempennars, M., Scannel, R., Walsh, M.J., Review of Scientific Instruments **79** (2008), doi: 10.1063/1.2969097.

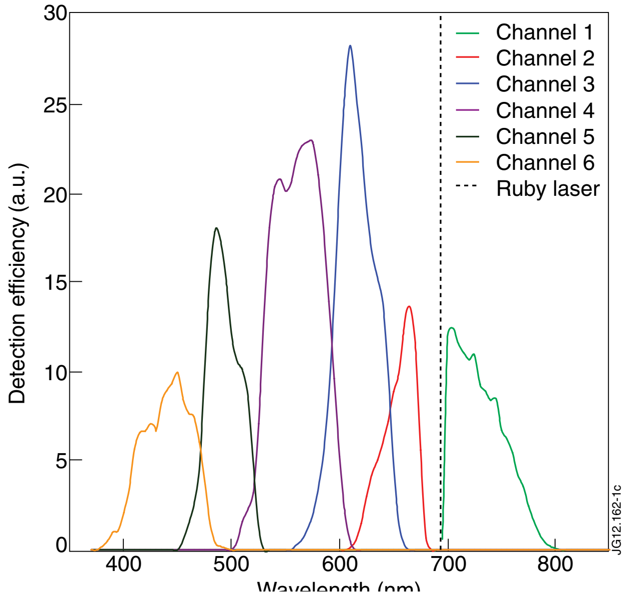

Figure 1: Spectral calibration of JET LIDAR Thomson scattering diagnostic.

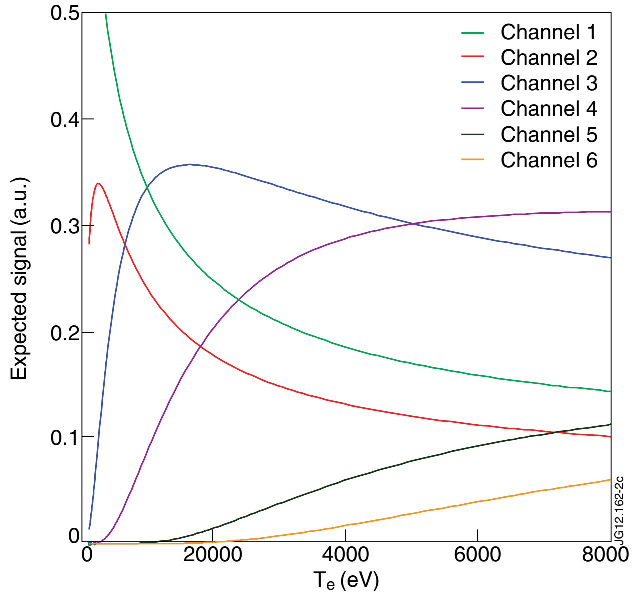

Figure 2: $F_i(T)$ or expected signal in different spectral channels as a function of T_e, as calculated for JET LIDAR diagnostic.

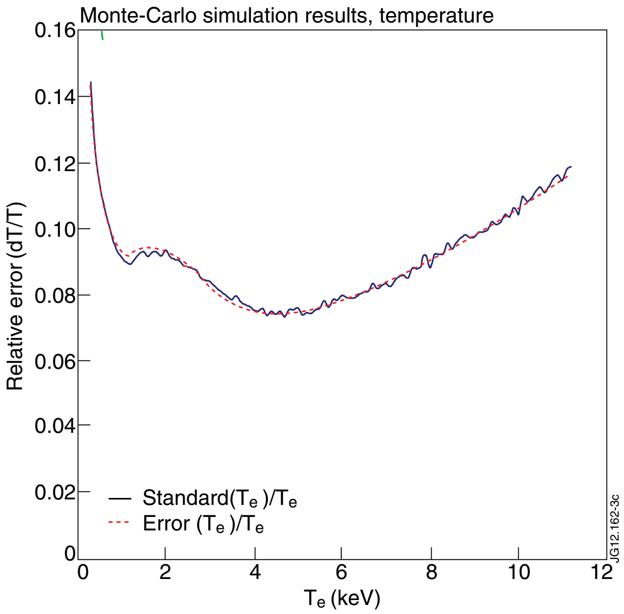

Figure 3: Comparison of relative T_e error bars, calculated using formula (12), with the standard deviation of 5000 min(χ^2) fits.

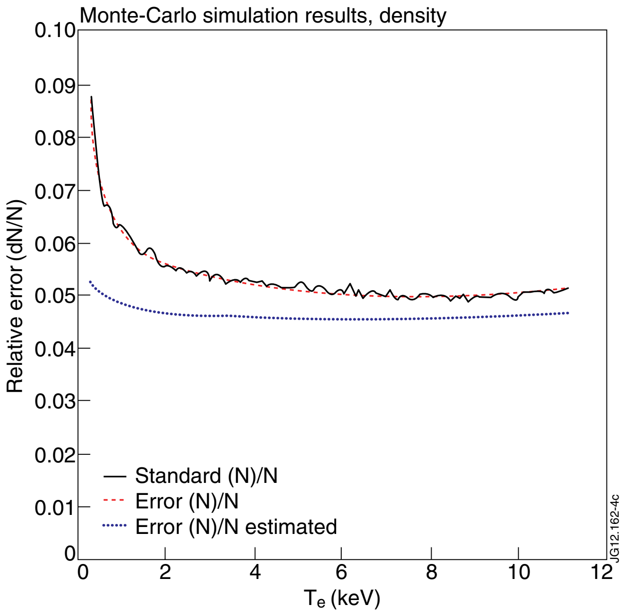

Figure 4: Comparison of relative N_e error bars, calculated using formula (14), with the standard deviation of 5000 min(χ^2) fits and with simplified estimate.

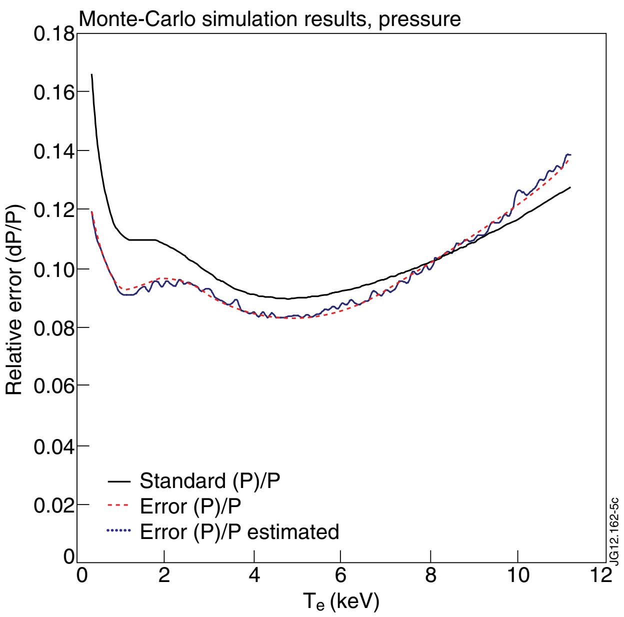

Figure 5: Comparison of relative P_e error bars, calculated using formula (15), with the standard deviation of 5000 min(χ^2) fits and with simple estimate done with (16).