

GENETIC ALGORITHMS: PLASMA DIAGNOSTIC SIGNAL ANALYSIS

A.P. Millar, D.C. McDonald and D.A. Diver

Dept. of Physics & Astronomy, University of Glasgow, G12 8QQ, Scotland

Genetic Algorithms (GAs) are used in combination with Bayesian Statistics to create a powerful and robust spectral analysis tool.

1. Introduction

The Thomson Scattering of electrons is used to ascertain the electron temperature profile in the COMPASS-D experiment. A Nd:YAG laser fires at 20 Hz. The 504 ns interval centred on each laser pulse is a **segment**. For further details see [1].

For the three channels of each spectrometer, reference data is generated (see Fig. 1).

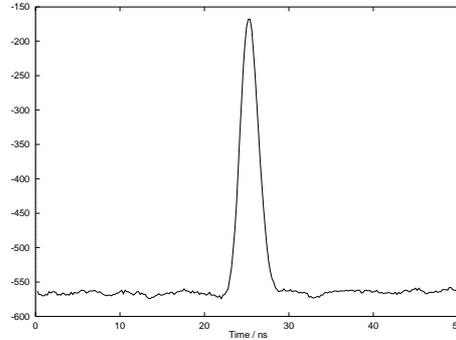


Figure 1. Reference data for channel 1 spectrometer 5

The time the pulse is observed can vary slightly. From observations, this variation is approximately Gaussian, with a variance of approximately one time bin.

The reference data has arbitrary amplitude and d.c. offset. This was normalised so that the resulting data g_i° has a maximum of one, and d.c. level of approximately zero. This was achieved by subtracting the median and rescaling.

The function $g_i(a, b, \mu)$ is the affine transformed reference curve g_i° as given by Eq. (1).

$$g_i(a, b, \mu) = a g_{i+\mu}^\circ + b \quad (1)$$

where a is the amplitude, b is the d.c. offset and μ is the temporal offset.

If the three channels are analysed together, then we can define the vector $\underline{a} = (a_1, a_2, a_3)$ where a_i is the amplitude of the i^{th} channel. Similar definitions hold for the vectors \underline{b} and $\underline{\mu}$. The reference functions are now \underline{g}_i° and $\underline{g}_i(\underline{a}, \underline{b}, \underline{\mu})$.

If the temporal offsets ($\underline{\mu}$) are suppressed, then the problem is applicable to the linear least-squares method. This method is fast, but error analysis can force the rejection of a sizeable percent of data points. This data rejection problem is especially noticeable when Electron Cyclotron Resonance Heating is used.

The problem of fitting reference to observed data can be analysed using Bayesian Statistics (see [2]). If the errors are Gaussian, the likelihood function is given by Eq. (2).

$$P(D|H_{\underline{a},\underline{b},\underline{\mu}}I) = \pi^{-n/2} \sigma_D^{-n} \exp \left[-\chi_D^2/2 \right] \quad (2)$$

where D , H are hypotheses of relevance and best fit parameterisation, and I is the background information, and

$$\chi_D^2 = \frac{1}{\sigma_D^2} \sum_{i=0}^{n-1} [f_i - g_i(\underline{a}, \underline{b}, \underline{\mu})]^2 \quad (3)$$

The prior information is that the three offsets μ_i form a Gaussian distribution with variance σ_μ^2 . If the average of the three offsets is $\bar{\mu}$ then this can be written:

$$P(H_{\underline{a},\underline{b},\underline{\mu}}|I) = \pi^{-3/2} \sigma_\mu^{-3} \exp \left[-\chi_\mu^2/2 \right] \quad (4)$$

where

$$\chi_\mu^2 = \frac{1}{\sigma_\mu^2} \sum_{i=1}^3 [\mu_i - \bar{\mu}]^2. \quad (5)$$

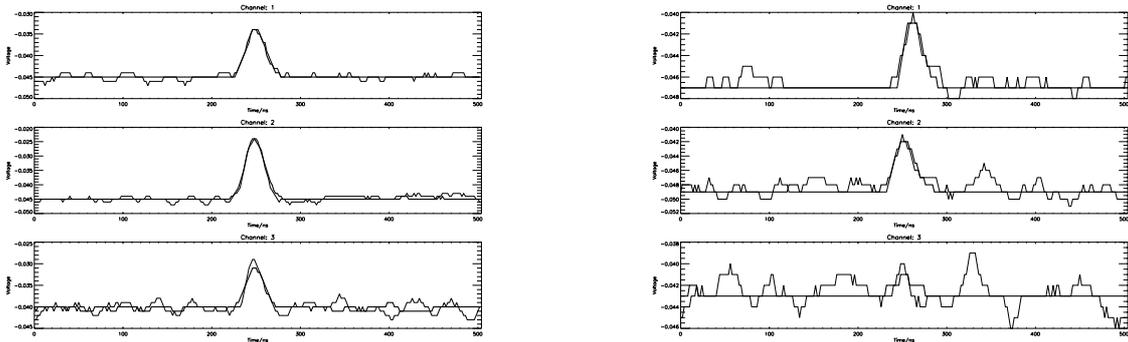
The ‘‘best-fit’’ values of \underline{a} , \underline{b} and $\underline{\mu}$ correspond to the values that maximise the plausibility of the model H , given the data is relevant and other information: $P(H_{\underline{a},\underline{b},\underline{\mu}}|DI)$ (see Eq. (6)).

$$P(H_{\underline{a},\underline{b},\underline{\mu}}|DI) \propto \exp \left[-\frac{1}{2}(\chi_D^2 + \chi_\mu^2) \right]. \quad (6)$$

Maximising $P(H_{\underline{a},\underline{b},\underline{\mu}}|DI)$ corresponds to minimising $\chi^2 = \chi_D^2 + \chi_\mu^2$. The function χ^2 is non-linear and a GA was used to find the global minimum. See [3], [4] and references therein for further details.

2. Results

The output for typical spectrometers is shown in Fig. 2 (thin lines). Also included is the GA fits (thick lines).



(a) Data from spectrometer 5 during segment 96

(b) Data from spectrometer 3 during segment 91

Figure 2. Data taken from COMPASS-D shot 26900

For each such set of three fits, the integrated pulse areas are then fitted to the relevant calibration curve (see fig. 3(a)) to obtain an electron temperature. This is repeated for all

spectrometers and the spatial distribution of electron temperature is obtained (see fig. 3(b)). More temperatures were usually recovered using our technique than was possible with a simple Least Squares approach.

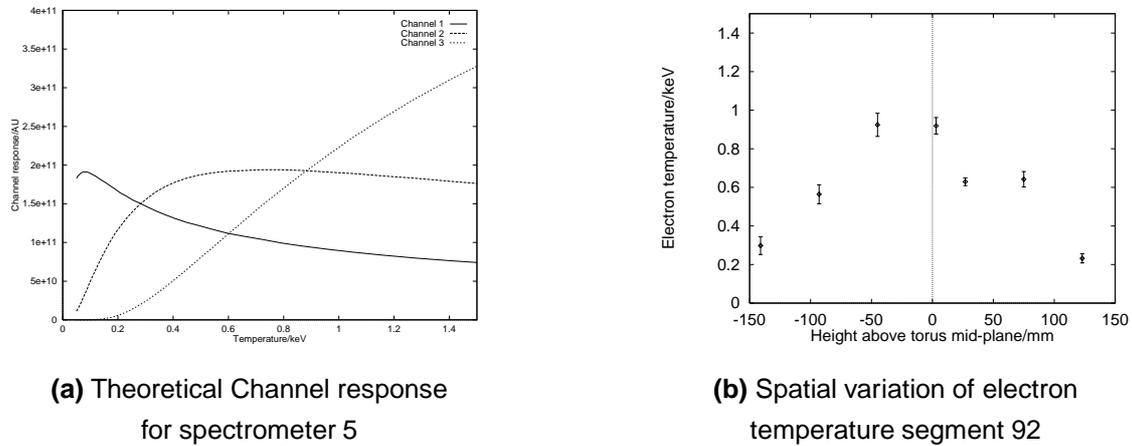


Figure 3. Graphs relating to temperature calculation for shot 26900

3. Conclusions

Genetic Algorithms are a powerful class of computational method for finding the global extremum of a function. Bayesian Statistics provides a near optimal statistic for determining the best model from a selection. By linking the GA “engine” to a Bayesian statistic more information has been gleaned from poor Signal to Noise ratio data than simple least squares method. The success of this novel technique suggests that a similar analysis is applicable to a more general class of problems.

Acknowledgements

This work is funded by the UK Engineering and Physical Research Council via a PhD Studentship [APM] and grant GR/L35232.

We would especially like to thank Keith Stammers and the COMPASS-D team of UKAEA Fusion, Culham Science & Engineering Centre for help, support and criticism.

References

- [1] I.H. Hutchinson: Principles of Plasma Diagnostics, chapter 7, pages 239–258. Cambridge University Press, 1992.
- [2] P.M. Lee: Bayesian Statistics: an introduction. Oxford University Press, 1989.
- [3] D.A. Diver: “Breeding spectral decomposition”, European J. Phys. **16**, 211, 1995.
- [4] D.A. Diver and D.G. Ireland: “Spectral decomposition by genetic algorithm”, Nucl. Instr. & Meth. in Phys. Res. A **399**(2–3), 414–420, 1997.