A COMPARISON OF PARAMETRIC AND NONPARAMETRIC METHODS OF NORMALISING MICROARRAY DATA

M.R. Khondoker[†1,2], C.A. Glasbey[1], B.J. Worton[2]

[1] *Biomathematics & Statistics Scotland, Edinburgh, UK;* [2] *University of Edinburgh, Edinburgh, UK*

[†] E-mail: *mizanur@bioss.ac.uk*

DNA microarrays facilitate the simultaneous monitoring of expression levels of thousands of genes in cell samples. Normalisation is an important first step in the analysis of microarray data, to correct for effects arising from imperfections in the technology rather than real biological differences. Loess smoothing is commonly used for correcting for trends in log-ratio data. On the other hand, parametric models, such as the additive plus multiplicative variance model, have been preferred for scale normalisation; several transformation methods have emerged, including log and arcsinh. However, the variance structure of microarray data may be of a more complex nature than can be generally accommodated by a parametric model. We evaluate a new nonparametric approach that incorporates location and scale normalisation simultaneously using Generalised Additive Models for Location, Scale and Shape (GAMLSS, Rigby and Stasinopoulos, 2005, *Applied Statistics* **54**, 507-554). We compare its performance in inferring differential expression with Huber *et al.*'s (2002, *Bioinformatics* **18**, S96-S104) arcsinh transformation. Simulations show GAMLSS to be as powerful as Huber's method when the parametric model is correct, and more powerful when the model is wrong.