Bayesian Analysis (2010) **5**, Number 4, pp. 691–696

Comment on Article by Vernon et al.

David A. van Dyk*

This paper tackles the computationally challenging task of comparing the predictions of the sophisticated Galform computer model for Galaxy Formation to observed light curves—data on the number of galaxies observed per unit volume in a given bin of luminosity for a particular band of light. The authors are to be commended for their clearly careful and diligent model-checking of this complex computer model. Judging from Figures 12 and 13 they where able to find parameter values that agree much more closely with the observed luminosity functions then what was previously available. (Although when comparing with data for which the model was not tuned, as in Figure 14, the results are more ambiguous.) By exploring the distribution of the parameters that result in acceptable model fits, the authors are able to draw conclusions about the complex relationships among the parameters of scientific interest. This appears to be an important step forward in our understanding of the formation and evolution of galaxies and at the same time demonstrates the power of the authors' sequential strategy for searching an enormous space for increasingly likely parameter values.

It may be helpful to illustrate my understanding of the authors' strategy in terms of standard statistical methodology using a simple problem. Suppose $Y_i \sim \mathrm{N}(\mu_i(\theta), \sigma_i^2)$ are independent for $i = 1, \ldots, n$, with each σ_i^2 known. The loglikelihood function is $L(\theta|Y) = -\sum_{i=1}^n (Y_i - \mu(\theta))^2/2\sigma_i^2$. If $\mu(\theta)$ is not overly complex, we can maximize L and consider its curvature or contours to make inference and learn about θ. We can also evaluate $L(\theta|Y)$ at its maximizer, $\hat\theta$, or values of θ near $\hat\theta$ to check whether the proposed Gaussian model is adequate for the data. If $L(\hat\theta|Y)$ is significantly smaller than we would expect we conclude that the model is inadequate. The authors consider a problem in which $\mu(\theta)$ is very complex, the likelihood can only be evaluated with substantial numerical effort, and standard optimization, quadrature, and sampling techniques are apparently impossible or impractical. Instead they search the parameter space by simply evaluating the objective function, $L(\theta|Y)$ in my simple example, at numerous values of θ. The evaluation points are then culled by thresholding on $L(\theta|Y)$. A new set of values of θ are selected in the newly discovered highest-likelihood region of the parameter space, the likelihood is reevaluated at these points, and the evaluation points are culled again using a more stringent threshold. This is repeated until a set of parameter values is obtained that adequately predict the observed data or until all possible value of θ have been eliminated by the likelihood threshold. Although the authors do not refer to the Gaussian loglikelihood function, the actual objective functions that they employ bare a remarkable resemblance to it. Using my notation, the *implausibility function* defined in (13) replaces the sum over i in $L(\theta|Y)$ by a maximization over i and the function in (16) would reduce to $L(\theta|Y)$ in the independent case. In both cases $\mu(\theta)$ involves emulation and the implausibility functions differ from the loglikelihood by a factor of $-1/2$. Thus, the authors aim to reduce implausibility as I aim to increase the likelihood. Having identified the set of parameter values that adequately predicts the

*Department of Statistics, University of California, Irvine, CA, mailto:dvd@ics.uci.edu

 DOI:10.1214/10-BA524D