Article

Constraining Coronal Heating: Employing Bayesian Analysis Techniques to Improve the Determination of Solar Atmospheric Plasma Parameters

Adamakis, Sotiris, Walsh, Robert W. and Morton-jones, Anthony J.

Adamakis, Sotiris, Walsh, Robert W. ORCID: 0000-0002-1025-9863 and Morton-jones, Anthony J. (2010) Constraining Coronal Heating: Employing Bayesian Analysis Techniques to Improve the Determination of Solar Atmospheric Plasma Parameters. Solar Physics, 262 (1). p. 117. ISSN 0038-0938

Solar Physics
DOI: 10.1007/•••••-•••-•••-••••-•

Constraining coronal heating: employing Bayesian analysis techniques to improve the determination of solar atmospheric plasma parameters

Sotiris Adamakis[1] · **Anthony J. Morton-Jones**[1] · **Robert W. Walsh**[1]

Abstract One way of revealing the nature of the coronal heating mechanism is by comparing simple theoretical one dimensional hydrostatic loop models with observations at the temperature and/or density structure along these features. The most well-known method for dealing with comparisons like that is the χ^2 approach. In this paper we consider the restrictions imposed by this approach and present an alternative way for making model comparisons using Bayesian statistics. In order to quantify our beliefs we use Bayes factors and information criteria such as AIC and BIC. Three simulated datasets are analyzed in order to validate the procedure and assess the effects of varying error bar size. Another two datasets (Ugarte-Urra *et al.*, 2005; Priest *et al.*, 2000) are re-analyzed using the method described above. In one of these two datasets (Ugarte-Urra *et al.*, 2005), due to the error estimates in the observed temperature values, it is not posible to distinguish between the different heating mechanisms. For this we suggest that both Classical and Bayesian statistics should be applied in order to make safe assumptions about the nature of the coronal heating mechanisms.

Keywords: Corona, Models; Corona, Structures; Heating, Coronal; Analysis, Statistical; Methods, Statistical

1. Introduction

Magnetically confined plasma loops are the fundamental building blocks of the solar atmosphere. Whole loop structures are observed over an extensive spectral range while extending over a large range of length-scales and dynamical time-scales. In particular, the Solar and Heliospheric Observatory (SOHO), the Transition Region and Coronal Explorer (TRACE) and now Hinode record loop-like features from small-scale brightenings lasting for tens of seconds to large-scale (of the order of the solar radius), apparently static loops that last for many hours.

Recent interest in loops has centred on the definitive determination via remote sensing of the basic parameter values within these features. One dimensional hydrostatic simulations of loop plasma (*e.g.* Peres, 2000) produce a temperature (T)

[1] University of Central Lancashire sadamakis@uclan.ac.uk, ajmorton-jones@uclan.ac.uk, rwwalsh@uclan.ac.uk

and density (ρ) structure along the loop that results from a balance between thermal conduction along the field-lines, optically thin radiative loss and a predefined coronal heating term. Generally, this profile extends from a cooler temperature at the loop base (or footpoint) area up to a hotter temperature at the loop apex. However, it has been demonstrated that the temperature gradient ($\frac{dT}{ds}$ where s is the distance along the loop) at each point along this profile is very dependent on where the energy deposition preferentially occurs (*e.g.* Priest *et al.*, , 2000). If the heat input is predominantly located at the base, the temperature in the "coronal" part of the loop will be relatively constant (*i.e.* $\frac{dT}{ds} \approx 0$). In comparison, if the energy is released near the loop apex, a significant temperature gradient ($\frac{dT}{ds} > 0$) travelling away from the apex can result. By observing the local temperature (and/or density) at successive spatial locations along a well-defined loop and then comparing the resulting profile with one generated from a one-dimensional hydrostatic model, then this could provide a means of constraining the possible preferred spatial location of the heating within that loop structure.

Confirming the dominant energy deposition position has been more difficult than first imagined, including for example the dataset introduced by Priest *et al.*, (2000) (PDS from now on) which has been interpreted by separate authors as uniform (Priest *et al.*, 2000), base (Aschwanden, 2001) and apex (Reale, 2002; Mackay *et al.*, 2000) heating. Other datasets examine the variation in density rather than temperature (Ugarte-Urra *et al.*, , 2005; UUDS from now on). In many of the above investigations, the important step of model comparison between the hydrostatic simulation and the observed plasma parameter values is undertaken by employing a weighted chi-squared analysis. Possibly this is an adequate, "quick-look" approach to tackling the model comparison. However two statistical "obstacles" present themselves here:

1. The precision of the temperature observations may not be sufficient to discriminate one heating function from another. As we shall see, substantive changes in the nature of the heating function can result in only subtle changes in the temperature profile of the loop.
2. The current approach (as in PDS and UUDS) for comparing one heating function model with another is to minimize the well-known statistic:

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(T_i - \widehat{T}_i\right)^2}{\sigma_i^2} \tag{1}$$

where n is the number of grid points, T_i is the observed temperature at distance s_i along the strand, $\widehat{T}_i$ is the theoretical prediction of the model concerning the temperature and σ_i is the standard deviation of the observed temperature. The procedure, applied to the solar coronal loop heating problem, would be to isolate the correct heating functional form (*e.g.* base or apex heating with an exponential profile) based on its capability to furnish the minimum statistic value.

There are a number of deficiencies to this procedure. First of all, for any heating functional form, there is a continuous range of parameter values, with an infinite set of possible parameter combinations, each of which is a candidate model. Therefore, it is not possible to properly compare one family's performance in fitting the data with another (*e.g.* apex heating with basal heating using an exponential form),

without resorting to selecting specific values of the parameters. Such an approach was employed by UUDS, where a grid method of equally spaced parameter combinations was used. Of course, the concern is that somewhere we have missed certain parameter values between grid points which could reverse the conclusion reached in the model comparison assessment.

Furthermore, using a procedure such as the one described above, there is no straightforward way of telling if we have significant evidence of one heating function being superior to another. The issue here is that the minimum χ^2 statistic approach is not well suited to model comparison problems, since it is primarily a goodness-of-fit statistic.

Another point of interest is that the minimum χ^2 statistic approach is only strictly valid under the normal errors assumption. As with the UUDS observations, this is often clearly not the case — the error bars may be asymmetric. This means a minimum χ^2 based assessment may not be reliable.

Finally, merely taking the model with the minimum χ^2 statistic does not tell us anything about the quality of the model. One can always improve a model fit by adding increasingly more parameters until, at the point of nonidentifiability, the model fit equates to all observation values exactly ("joins the dots"), thus producing a χ^2 statistic value of zero, which would be then the model of choice based on the minimum χ^2 statistic criterion. However, it is clear that what we have done in that case is not come closer to the true picture: rather an artefactual model has been constructed which is unlikely to reflect the true picture. This whole problem is one of *overfitting*, whereby model fit is apparently improved by adding increasingly more parameters, and is not taken into account by simply using the χ^2 statistic.

The practical way round these issues is to resort to a simulation approach. In this paper we describe the use of a Bayesian Markov chain Monte Carlo (MCMC) analysis to solar coronal loop data, which embeds hydrodynamic modelling techniques (Walsh, Bell, and Hood, 1995; see Section 2.3) within a basic Metropolis-Hastings algorithm (Metropolis *et al.*, 1953; Hastings, 1970). Then in Section 3, we apply our method to three simulated datasets in order to test the efficiency of the method; in particular we examine the artificial inflation and compression of the error bars with subsequent inflation and compression of the, *e.g.*, parameter credible intervals. Section 4 investigates some real coronal loop datasets (PDS and UUDS) and subsequently presents quantitatively based conclusions concerning the nature of the heating of the loops examined. A discussion on how this work can be further progressed can be found in Section 5.

2. Statistical Methods

2.1. General Approach

In principle, the provision of observational temperature data with distance from the base of the loop will yield information concerning the nature of the heating function. Our approach employs a Bayesian analysis of the data using MCMC techniques, which are increasingly being employed upon astrophysical datasets (Adamakis, Morton-Jones, and Walsh, 2007). The Bayesian approach incorporates prior information we may have on model parameters we are interested in with the observational data

(likelihood) to form the updated or posterior information we have on the parameters. This uses Bayes' Theorem:

$$\pi(\mathbf{P}|\mathbf{T}) = \frac{L(\mathbf{T}|\mathbf{P})\pi(\mathbf{P})}{f(\mathbf{T})} \tag{2}$$

for observations $\mathbf{T}$ and parameter space $\mathbf{P}$, where $L(\cdot|\cdot)$ is the likelihood function, $\pi(\cdot)$ the prior distribution, $f(\cdot)$ the marginal likelihood and $\pi(\cdot|\cdot)$ the posterior distribution. The prior distribution may or may not reflect knowledge we may have on a parameter. If it doesn't we use a so-called noninformative prior.

The posterior distribution can be simulated using MCMC techniques. In our case, we use the Metropolis-Hastings algorithm to draw parameter values for each parameter in our model from the posterior distribution. This is not a straightforward application of the Metropolis-Hastings algorithm however, as we need to convert proposed heating function parameter values into model temperatures, $\widehat{T}_i$, at each distance, s_i. Using single-variable updates, the established hydrodynamic modelling code is used, with input being the current parameter values (including the proposed value of the parameter under consideration at a given iteration), to produce the $\widehat{T}_i$. The $\widehat{T}_i$ can then be used to construct the likelihood of the temperature observations, T_i. A further complication arises because we must have physically sensible temperature profiles with distance for the loop. If the $\widehat{T}_i$ is not monotonically decreasing from the apex to the base, then we must reject this generated set of parameter values. Thus, with k the number of parameters in the model, our approach can be summarised as:

Step 1. For the jth parameter, with current parameter value p_j, generate the proposed value, p_j^*, from a proposal distribution.

Step 2. Using the current set of parameter values, $(p_1, p_2, \ldots, p_j^*, \ldots, p_k)$, call the hydrodynamic code to generate the $\widehat{T}_i$.

Step 3. Reject the proposal p_j^* if the $\widehat{T}_i$ are not monotonically decreasing. If so retain the current value p_j and go to Step 4. Otherwise accept the new parameter value for the parameter p_j with probability

$$\alpha(p_j, p_j^*) = \frac{\pi((\mathbf{P}_{k-j}, p_j^*)|\mathbf{T})q((\mathbf{P}_{k-j}, p_j^*), (\mathbf{P}_{k-j}, p_j))}{\pi((\mathbf{P}_{k-j}, p_j)|\mathbf{T})q((\mathbf{P}_{k-j}, p_j), (\mathbf{P}_{k-j}, p_j^*))} \tag{3}$$

where $\mathbf{P}_{k-j} = (p_1, p_2, \ldots, p_{j-1}, p_{j+1}, \ldots, p_k)$ and $q(\cdot, \cdot)$ the proposal distribution and go to Step 4.

Step 4. Move to the next parameter p_{j+1} and repeat the process.

We follow the same procedure for multivariate Metropolis updates, with the only difference that instead of updating a parameter at each time, we update the whole set of parameters, *i.e* instead of p_j^* we have $\mathbf{P}^*$. Although single-variable Metropolis updates seems to behave slightly better than multivariate Metropolis updates (Neal, 2003), it can be very time consuming. For this reason we have used multivariate Metropolis updates for the examples presented in Section 3.

In this way, through thousands of iterations, the marginal posterior distribution for each parameter is built up. This distribution can then be used in many ways

to assess the parameter values, *e.g.* by drawing up 95% credible intervals. If this whole analysis is repeated for each heating function model, then model comparison techniques can be used to provide a quantitative assessment of the likelihood of one model over the other (see Section 2.4).

One further point to be made is that, with this approach we could in theory calculate the chi-squared statistic for each set of accepted parameter values under each heating function model, and take the model which provides the minimum value. This is likely to be more reliable than using the grid approach because we are naturally gravitating stochastically around the most likely parameter values which correspond also to the minimum chi-squared values (if the likelihood is normally distributed). Hence we would be less likely than with the regularised, non-stochastic grid approach to miss out on the values which would reverse our decision. However, for the reasons laid out in Section 1, we believe these alternative diagnostics are superior.

2.2. Data Distribution

The likelihood of the data should represent the way that our observations are distributed. This can change according to the way we gather the data; for example, this could be due to the instrument we use to gather the data or whether we have symmetric or assymetric error bars. This must be taken into account in the analysis.

2.2.1. Asymmetric and symmetric error bars

In case the error bars are not symmetric we have to deal with non-symmetric distributions. One popular right skewed distribution with positive support is the Gamma distribution. In this case, the observations $T_i,\ i = 1, \ldots, n$, will have model likelihood function:

$$\begin{aligned} L(\mathbf{T}|\mathbf{P}) &= \prod_{i=1}^{n} f(T_i|\mathbf{P}) \\ &= \prod_{i=1}^{n} T_i^{\gamma_i - 1} \frac{\exp(-T_i/\delta_i)}{\Gamma(\gamma_i)\delta_i^{\gamma_i}} I(T_i \in \mathcal{S}_1) \end{aligned}$$

where $\mathbf{T} = (T_1, \ldots, T_n)$, γ_i, δ_i are the parameters of the Gamma distribution, $\mathcal{S}_1$ is the domain of $f(\cdot|\cdot)$, $\mathbf{P}$ is the parameter vector and

$$I(T_i \in \mathcal{S}_1) = \begin{cases} 1, & \text{if } T_i \geq 0 \text{ and } T_{i+1} \leq T_i \text{ from the apex to the base.} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

is the indicator function of the temperature.

On the other hand, if the data we collected give symmetric error bars, we should use a symmetric distribution. A Gaussian distribution will usually be most appropriate. Thus, the mode likelihood function of T_i will be:

$$L(\mathbf{T}|\mathbf{P}) = \prod_{i=1}^{n} f(T_i|\mathbf{P})$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(T_i-\mu_i)^2}{2\sigma_i^2}\right) I(T_i \in \mathcal{S}_1)$$

where $\mathbf{T} = (T_1, \ldots, T_n)$, μ_i, σ_i are the parameters of the Normal distribution and the indicator function is the same as in Equation (4).

2.2.2. Interpretation of error bars

It is important to define clearly the standard deviation of the data from the error bars using probabilistic arguments. For instance, if we gather temperature values and we believe with some probability pr_i that these values lie in the range $(T_{L,i}, T_{U,i})$ then we have:

$$P(T_{L,i} \leq T_i \leq T_{U,i}) = pr_i \tag{5}$$

where $(T_{L,i}, T_{U,i})$ are the lower and upper points of the error bar for the ith grid point respectively. The variance of the data distribution is calculated by solving Equation (5) with the acceptance that the observed T_i are the mode of that distribution. In the case of the Normal distribution we have:

$$\sigma_i = \frac{T_{U,i} - T_{L,i}}{2\Phi^{-1}\left[(1+pr_i)/2\right]}$$

where $\Phi(\cdot)$ is the Cumulative Distribution Function (CDF) of the Normal distribution with mean 0 and standard deviation 1. In the special case that $pr_i = 0.9973, i = 1, \ldots, n$ then $2\Phi^{-1}\left[(1+pr_i)/2\right] = 6$ and we get the "6σ" belief. In the case of the Gamma distribution, Equation (5) is difficult to solve analytically, so we turn to numerical methods, *e.g.* Newton-Raphson (Gelman *et al.*, 2003). The mode of the data distribution is calculated by the temperature values that are proposed from the model. We can then calculate the parameters from knowledge of the mode and variance.

2.3. Heating Function: Models and Parameters

The one-dimensional plasma equations employed in this model are (Walsh, Bell, and Hood, 1995):

$$\frac{D\rho}{Dt} + \rho\frac{\partial \upsilon}{\partial s} = 0 \tag{6}$$

$$\rho\frac{D\upsilon}{Dt} = -\frac{\partial p}{\partial s} + \rho g + \rho\nu\frac{\partial^2 \upsilon}{\partial s^2} \tag{7}$$

$$\frac{\rho^\gamma}{\gamma - 1}\frac{D}{Dt}\left(\frac{p}{\rho^\gamma}\right) = \kappa_0\frac{\partial}{\partial s}\left(T^{5/2}\frac{\partial T}{\partial s}\right) - \rho^2 Q(T) + H(s) \tag{8}$$

$$p = \frac{R}{\tilde{\mu}}\rho T \tag{9}$$

where $\frac{D}{Dt} = \frac{\partial}{\partial t} + \upsilon.\nabla$ is the total derivative, ρ is the density, t is the time, υ is the velocity of the plasma, p is the pressure, g is the gravitational acceleration, ν is the coefficient of kinematic viscosity, γ is the ratio of specific heats, $\kappa_0 = 10^{-11}$ for the

corona, Q is the radiative loss function, H is the heat input, R is the molecular gas constant and $\tilde{\mu} = 0.6$ in the ionised corona. For the radiation function we can assume that it is of the form $Q(T) = \chi T^{\theta}$. If we assume gravity and viscosity negligible, normalize $s, t, \upsilon, T, \rho, p$ by setting $s = l\overline{s}, t = t_c\overline{t}, \upsilon = \upsilon_c\overline{\upsilon}, T = T_c\overline{T}, \rho = \rho_c\overline{\rho}, p = p_c\overline{p}$ and assume constant pressure along the loop (*e.g.* the conductive velocity is much smaller than the sound speed), then we can rewrite Equations (6) to (9)

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho \upsilon)}{\partial s} = 0 \tag{10}$$

$$\frac{\partial \upsilon}{\partial s} = \frac{\partial}{\partial s}\left(T^{5/2}\frac{\partial T}{\partial s}\right) - b\left[Q(T) - H(s)\right] \tag{11}$$

$$\rho = \frac{1}{T} \tag{12}$$

where all bars have been removed for convenience and $b = \frac{\rho_c^2 \chi_c T_c^{\theta_c} l^2}{\kappa_0 T_c^{7/2}}$. Equations (10) to (12) are solved with the following boundary conditions:

$$\frac{\partial T}{\partial s} = 0 \quad \text{at} \quad s = 0 \tag{13}$$

$$T = T_{foot} \quad \text{at} \quad s = 0.5 \tag{14}$$

where $s = 0$ is the apex of the loop. We have employed the form of the radiative loss function as outlined in (Hildner, 1974).

Whereas the optically thin radiation loss function can be estimated from observations, the form of the heating function still remains a mystery. In the analysis that follows, we assume that the heating function $H(s)$ has the following general form:

$$H(s) = \lambda \exp(\beta s) \tag{15}$$

Of course this is only one case. We can test different functions in order to see which one best describes our data (see implications of this in Section 5). Applying Equations (15) to (11) we get:

$$\frac{\partial \upsilon}{\partial s} = \frac{\partial}{\partial s}\left(T^{5/2}\frac{\partial T}{\partial s}\right) - bQ(T) + \alpha \exp(\beta s) \tag{16}$$

where $\alpha = b\lambda$. We have replaced $b\lambda$ with α, because b and λ will be extremely high posterior correlated.

Thus, we have a range of parameters to investigate. Firstly, there is α and β from the heating function. β is exceptionally important because altering its value can have a profound effect on the nature of the heating profile. For example, if it is positive then more heat is deposited in the lower part of the loop (footpoint or basal heating). On the other hand, if β is negtive then more heat is deposited in the upper part of the loop (apex heating). Note that when $\beta = 0$ we have the "unique"case of uniform heating (although see discussion on this in Section 5).

Secondly, we introduce T_{foot} as an extra parameter. Ugarte-Urra *et al.*, (2005) highlight the sensitivity of choosing this boundary condition. Finally, in our simplified HD equations, we assume an isobaric scenario. Thus, since pressure p remains unchanged along the loop for a given set of parameter values, we have decided to leave its

value floating. The pressure will always be equal to p_c. Given $p_c = \frac{R}{\mu}\rho_c T_c$, then if we assume we fix T_c (10^6K), then ρ_c will be a changing value to explore. Subsequently, varying ρ_c changes b which hence becomes our forth and final parameter. Please note that we assume that the length l of the loop to be well known and thus simply defined.

To sum up, let the parameter space be $\mathbf{P} = (b, \alpha, \beta, T_{foot})^T$ with observed temperature values $\mathbf{T} \in \mathcal{S}_1$ (the data). The values of $\mathbf{P}$ lie in the region $\mathcal{S}_2 \in (0,\infty) \times (0,\infty) \times \mathcal{R} \times [0,\infty)$. The restriction $\alpha > 0$ is because extra heat should be added to the system, not subtracted from it. Methods for choosing priors are discussed in Section 2.5.

2.4. Model Comparison

In Bayesian statistics, Bayes factor is considered to be the traditional way of testing two or more hypotheses in Bayesian statistics. Suppose H_1, H_2 are the two hypotheses we want to test. The odds form of Bayes's theorem is:

$$\frac{\pi(H_1|\mathbf{T})}{\pi(H_2|\mathbf{T})} = \frac{f(\mathbf{T}|H_1)}{f(\mathbf{T}|H_2)} \frac{\pi(H_1)}{\pi(H_2)} \quad . \tag{17}$$

Note that $\pi(H_k)$ is the belief we have about the truth of the hypothesis H_k before we observe the data, $\pi(H_k|\mathbf{T})$ is what we get after we observe the data and $f(\mathbf{T}|H_k)$ is the marginal density, *i.e.* the belief of the data after we sum over the parameter space. The first term of the right hand side of Equation (17) is the *Bayes factor*. The second term of the right hand side of Equation (17) is the prior odds of the two hypotheses. In the absence of any prior information we may assume this to be 1 (*i.e.* $\pi(H_1) = \pi(H_2) = 0.5$). In that particular case Bayes factor is equal to the posterior odds. However, if we have some prior information about the hypotheses there is always the option to include it in the analysis.

However, it is worth mentioning here that Bayes factor tends to be more sensitive to the choice of prior than the posterior probability of an interval (Kass, 1993; Kass and Greenhouse, 1989) and so choice of priors becomes even more critical. For example, if we use an improper prior (say Uniform with an infinite range) for a parameter of interest, this will result to ill-defined Bayes factors and posterior probabilities that prefer the simplest model, with probability one, regardless of the information from the data. This is widely known as the Bartlett's paradox (Bartlett, 1957; Lindley, 1957). Apart from improper priors, this might be a consequence of using priors with a very large spread, in an effort to make our distribution non-informative, which in turn can lead to false conclusions (Kass and Greenhouse, 1989). Thus, in the case where there is no available prior information, the spread of the prior should not be very large in order to have effective results with Bayes factors. Raftery (1996) propose a way to overcome this problem using the Laplace method for integrals. In this paper we present results from three procedures to calculate the marginal densities, $f(\mathbf{T}|H_k)$. Interested readers can follow up the descriptions of these methods in Kass and Raftery (1995) and Raftery (1996), which describe these approaches in detail:

1. Laplace method with posterior covariance matrix.
2. Laplace estimator with robust posterior covariance matrix.
3. Monte Carlo estimation.

In Section 3 we have also tried the harmonic mean estimator (Gelfand and Dey, 1994), but due to the fact that it suffers from infinite variance, it is far from convergence.

An alternative way of making model comparison is by using the Akaike Information Criterion (AIC) proposed by Akaike (1974) and the Bayesian Information Criterion (BIC) proposed by Schwarz (1978). The former proposes to choose the model that minimizes

$$\text{AIC} = -2(\text{log maximised likelihood}) + 2(\text{number of parameters})$$

whereas the latter chooses the model that minimizes

$$\text{BIC} = -2(\text{log maximised likelihood}) + (\log n)(\text{number of parameters})$$

where n in this case will be the number of the observed data points. AIC tends to be biased in favor of more complicated models, as the log-likelihood tends to increase faster than the number of parameters. BIC tends to favor simpler models than those chosen by AIC. In the special case that we are dealing with a symmetric likelihood function (*e.g.* Gaussian distribution) maximising the likelihood function would be like minimizng the χ^2 statistic. Thus, in the following we will also employ these tests.

2.5. Choosing Priors

A natural way to address the prior distributions for our parameters is by considering their supports (in this case their domain). Thus, for b, α and T_{foot} we would like a distribution which supports non-negative numbers and for β a distribution which support all the real numbers. Because of the fact that we do not have any prior information for the parameters b, α we will use improper priors for these two paramters ($\pi(b, \alpha) \propto 1$). Gamma and Normal distributions seem to be ideal for parameters T_{foot} and β respectively, *i.e.* $\beta \sim N(\beta_1, \beta_2^2)$ and $T_{foot} \sim Gamma(t_1, t_2)$, with the parameters $\beta_1, \beta_2, t_1, t_2$ yet to be defined.

To avoid indeterminate Bayes factors, we decided to avoid using improper priors for the parameter of interest β. When we start the analysis it could be that we do not have any prior information at all. For example, the crucial parameter here is β. If there was not any prior information, we could use for estimation purposes a Normal prior distribution centered at 0, *i.e.* $\beta_1 = 0$, with a big variation, *e.g.* $\beta_2^2 = 10^4$. This will make it "non-informative", but it will lead to Bartlett's paradox (see Section 2.4) when we want to include model comparison into our analysis. Thus, it would be preferable to include any prior information available. For example, we set 99% confidence on the parameter β being between -20 and 20. This means that we can assume $\beta_1 = 0, \beta_2 = 7.76$. Using similar notions we end up with $t_1 = 2, t_2 = 1$ (which will give 95.96% for T_{foot} to fall between 0 and 5 with mode at 1).

An alternative choice of prior distribution, based to the previous one, is by using "Dual" priors, which lead to well-defined Bayes factors, and can be described as following: choosing a Normal distribution with $\beta_1 = 0, \beta_2 = 7.76$ will give more weight into the values of β that are closer to β_1. This means that $f(0)/f(20) = 27.70$. So, instead of using a prior that will give more weight to certain values of the parameter, we may want to introduce a prior which will allow to jump between two values with equal prior probability. In that case the $\pi(\mathbf{P}) \propto 1$ described above

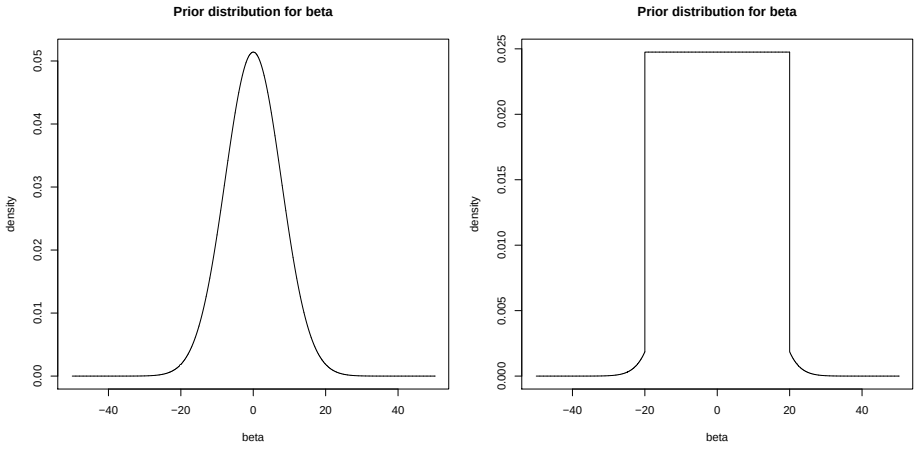

Figure 1. (a) "Common" prior distribution for β. (b) "Dual" prior distribution for β.

seem ideal but has many problems. It seems natural to use a "Dual" prior that combines the above two distributions and integrates to unity. For example, for the β parameter we may say that we are 99% confident that the parameter β should fall between -20 and 20, as above, and we want the probability density function to be the same between -20 and 20. This will give a probability density function as in Figure 1(b) instead of a probability density function as in Figure 1(a). "Dual" priors lead to well-defined Bayes factors. In Sections 3 and 4 we use "Dual" priors for parameters β and T_{foot}. More specifically, for β we use a uniform distribution between -20 and 20 and a normal distribution with $\beta_1 = 0, \beta_2 = 7.76$ for every other value, whereas for T_{foot} we use a uniform distribution between 0 and 3 and a Gamma distribution with $t_1 = 2, t_2 = 1$ for every other value.

2.6. Implementation of the MCMC Method

In generating our posterior distribution samples, the problem of within chain autocorrelation was found to be a significant problem. This problem means that much larger chains are required in order to achieve a representative sample from the target (posterior) distributions. There are several ways to deal with this problem, but we have found the most effective way has been to use the method described by Tierney and Mira (1999). The idea is to use more than one proposal in each step. This means that we start with a proposed parameter value combination. If they are accepted then move to the next step, otherwise propose another parameter value combination. If the second set is accepted then move to the next step, otherwise propose a third parameter value combination and so on. We can stop at any time we like this procedure, keep the current parameter values and move to the next step. The acceptance probability of each stage has to be adjusted in order to preserve a stationary distribution. This method has the advantage that we can test different proposals at each step, which can improve efficiency of mixing. We have used both two- and three-stages in our simulation procedures. At the first stage we propose values from an independent probability density. At the second stage (if needed) we

propose values from a random walk probability density based on the current values. Finally, at the third stage (again if needed) we can propose values as in the second stage but with smaller standard deviation or from a random walk probability density based on the rejected values from the first stage. Furthermore, in order to improve the mixing of the chain, we reparameterize the space as our initial parameters are highly correlated. For the reparameterization and for the independent proposal of the first stage, we have run a pilot chain, *i.e.*, we first run a simple Metropolis algorithm and from the crude estimates of the parameters we get we construct the independent proposal and choose the reparameterization scheme we will follow.

3. Testing Against Simulated Observations

In comparing the observed datasets with the HD simulation, we wish to examine the following four hypotheses:

1. $H_1 : \beta \neq 0$ — that is, heat input is not spatially uniform;
2. $H_2 : \beta = 0$ — heat input is spatially uniform;
3. $H_3 : \beta > 0$ — heat input is footpoint dominant;
4. $H_4 : \beta < 0$ — heat input is apex dominant.

In what follows, we have generated three datasets (SDS1, SDS2 and SDS3) in order to validate the Bayesian analysis and to explore the effects of changing the size of the error bars. Then, in Section 4, we present the results obtained from PDS and UUDS.

For the three simulated datasets, we have chosen values as follows: $\kappa_0 = 10^{-11}$, $\theta_c = 1$, $\chi_c = 1.97 \times 10^{12}$, $T_c = 1$MK, $\rho_c = 10^{-12}$kgr/m^3, $l = 71.25$Mm (which will give $b = 1$), $T_{foot} = 1$ along with $(\alpha, \beta)^T = (20, 0)^T$ for SDS1, $(\alpha, \beta)^T = (0.68, 10)^T$ for SDS2 and $(\alpha, \beta)^T = (100.68, -10)^T$ for SDS3. These combinations for parameters α and β were chosen so that the total heating input for all the three simulated datasets would be the same. Figure 2 depicts these three heating functions against normalized distance along half the loop. Figure 3 illustrates that difference in thermal structure along the loop for the different values of β. The error bars (0.2MK) were chosen so that we can distinguish between the three different temperature profiles with 13 data points chosen. We assume that the errors between the data and the model come from a Normal distribution and that the error bars represent a "6σ" belief.

3.1. Simulated Dataset 1 ($\beta = 0$)

Under the hypothesis H_1 we present the summary statistics in Table 1, where the first column is the estimated mean value of each parameter, the second column is the parameters' joint probability density mode (the parameters' combination that maximise the posterior distribution), the third column is the standard deviation, the fourth, fifth and sixth columns are the 2.5%, 50% and 97.5% quantiles, while the final column displays the actual value used to generate the simulated dataset. Hence, a 95% credible interval can be constructed from the 2.5% quantile to the 97.5% quantile.

From the joint mode of Table 1 we can see that all the estimates of the parameters are fairly close to the "known" values. The only parameter that is far away from the

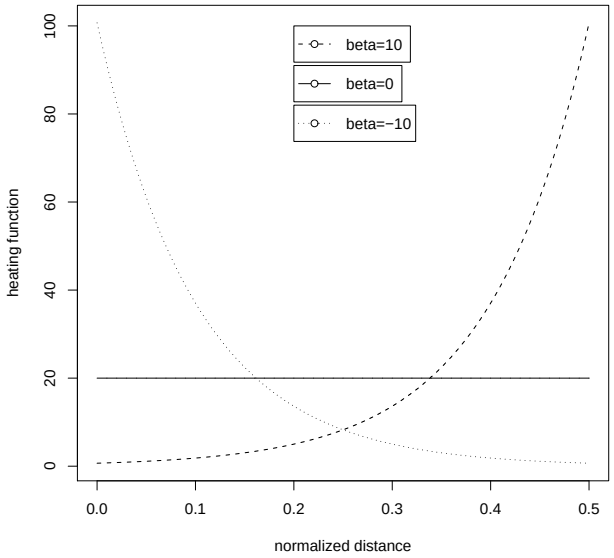

Figure 2. Heating input as a function of normalized distance (apex at 0.0, base at 0.5), for $\alpha = 0.68, \beta = 10$ (footpoint heating — dashed line), $\alpha = 20, \beta = 0$ (uniform heating — solid line), $\alpha = 100.68, \beta = -10$ (apex heating — dotted line).

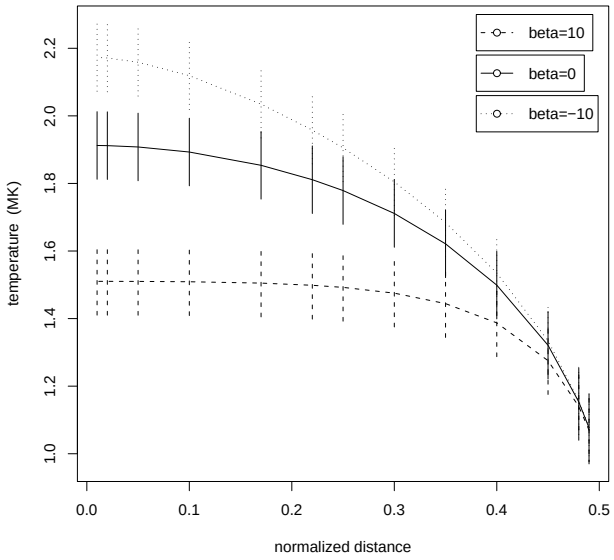

Figure 3. Temperature against normalized distance for SDS1 (uniform heating — solid line), SDS2 (footpoint heating — dashed line) and SDS3 (apex heating — dotted line). The error bars for all the data points are 0.2MK.

Table 1. Summary of the posterior inference for hypothesis H_1 for the SDS1.

	mean	mode	s.d.	2.5%	50%	97.5%	actual value
b	29.83	15.05	22.77	0.98	24.62	79.03	1.00
α	24.40	21.86	4.95	15.27	24.24	34.57	20.00
β	0.15	0.16	1.17	-2.47	0.25	2.16	0.00
T_{foot}	1.01	1.00	0.04	0.94	1.01	1.08	1.00

Table 2. Results for the parameter β from the SDS1 with different error bars and actual value $\beta = 0$.

	mean	mode	s.d.	2.5%	50%	97.5%	$P(\beta < 0)$
E.B. $\times$ 4	-7.22	0.64	6.83	-19.37	-6.55	3.89	0.82
E.B. $\times$ 2	-1.18	0.35	2.95	-8.59	-0.62	3.02	0.60
E.B.	0.15	0.16	1.17	-2.47	0.25	2.16	0.41
E.B./2	0.36	0.14	0.61	-0.99	0.42	1.41	0.26
E.B./4	0.38	-0.04	0.38	-0.39	0.39	1.08	0.17
E.B./500	0.00	0.00	0.01	-0.02	0.00	0.02	0.50

given value is b. A 95% credible interval gives b between 0.98 and 79.03. Of course, this does not imply that the Bayesian approach is inefficient. What this does mean is that b multiplying the radiative loss component appears to have little effect on the thermal profile. Essentially, the value of the isobaric pressure can vary greatly but this has only a small impact on the temperature.

The dispersion of each parameter is correlated with the length of the error bars. This means that the tighter error bars we have, the smaller the range of parameters will be (our confidence about the observations will be projected into the parameter space). This can be seen in Table 2, where in order to test the robustness of this method we have used SDS1 but with various sizes of error bars and examine how the main parameter of interest (β) will react to these changes. As expected the smaller the error bars, the closer we get to the "true" value. This can be seen from the values of the mean, the joint mode and the median. As the error bars are decreasing so does the standard deviation and the length of the 95% credible intervals.

In order to test the posterior hypotheses odds we have constructed the Bayes factors as in Section 2.4. They are computed for our initial size of error bars. Table 3 shows these odds for each of the four hypotheses. It is preferable to use more than one estimate of the marginal density to calculate the Bayes factors. Thus, using the Laplace method with robust posterior covariance matrix, we calculated the odds $25.65 : 2.30 : 1.02 : 1$ for the hypotheses $H_2 : H_1 : H_4 : H_3$ respectively. For example, hypothesis H_2 is 25.65 times more likely to occur than hypothesis H_3, whereas hypothesis H_2 is $25.65/1.02 \approx 25.15$ times more likely to occur than hypothesis H_4. All of the three estimates suggest that most preferable hypothesis is $\beta = 0$, which comes in accordance with our initial assumption. We come to the same conclusion with Table 4, where hypothesis $\beta = 0$ is the one that minimizes the information criteria.

Table 3. Posterior hypotheses odds for the SDS1. Marginal densities have been calculated from: 1: Laplace method with posterior covariance matrix, 2: Laplace method with robust posterior covariance matrix, 3: Monte Carlo estimation with the probability density from stage 1 as the additional probability density.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
1	2.31	15.97	1	1.15
2	2.30	25.65	1	1.02
3	2.62	32.75	1.43	1

Table 4. Information criteria for the SDS1.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
AIC	-56.53	-58.54	-56.53	-56.52
BIC	-54.27	-56.84	-54.27	-54.26

3.2. Simulated Dataset 2 ($\beta = 10$)

Now consider applying the same analysis to SDS2, where this time we have a positive value of β ($\beta = 10$).

From Table 5 we see that the initial error bars are sufficient to distinguish between the different heating function forms. A 95% credible interval for β is between 3.63 and 11.13, whereas the probability of β being positive is one. Again we have fairly reasonable estimates for the mean values of the parameters, apart from b for which we can suggest reasons as with SDS1. Bayes factors (Table 6) and information criteria (Table 7) agree and they both suggest the basal heating model ($\beta > 0$).

3.3. Simulated Dataset 3 ($\beta = -10$)

Consider now a simulated dataset with the same absolute magnitude of β as with SDS2, but with different sign ($\beta = -10$).

The joint mode vector is $(6.09, 96.75, -9.22, 1.00)^T$ for parameters $(b, \alpha, \beta, T_{foot})^T$ (see Table 8), which is close to our initial assumptions. From these summary

Table 5. Summary of the posterior inference for hypothesis H_1 for the SDS2.

	mean	mode	s.d.	2.5%	50%	97.5%	actual value
b	15.05	2.16	10.76	0.74	13.07	39.62	1.00
α	3.27	0.79	2.18	0.52	2.77	8.65	0.68
β	6.73	9.66	1.94	3.63	6.47	11.13	10.00
T_{foot}	1.02	1.00	0.04	0.94	1.02	1.09	1.00

Table 6. Posterior hypotheses odds for the SDS2. Marginal densities have been calculated from: 1: Laplace method with posterior covariance matrix, 2: Laplace method with robust posterior covariance matrix, 3: Monte Carlo estimation with the probability density from stage 1 as the additional probability density.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
1	1.26×10^6	29.07	7.66×10^5	1
2	6.65×10^5	50.43	6.13×10^5	1
3	8.85×10^5	88.31	3.19×10^5	1

Table 7. Information criteria for the SDS2.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
AIC	-56.50	-35.86	-56.50	-33.81
BIC	-54.24	-34.17	-54.24	-31.55

statistics the only parameter that is away from the given values is again b, for which we can assume the same reasoning as with SDS1, whereas the probability for β being negative this time is one. Bayes factors (Table 9) and information criteria (Table 10) both pick the apex heating model ($\beta < 0$), within the given error bars.

4. Application of Observations

4.1. Priest *et al.* Dataset

The data values we have used for this example were extracted from Figure 9*a* of Priest *et al.*, (2000). In this analysis we assume that the error bars reflect high degree of confidence. There are a number of important aspects to be kept in mind. Firstly, the loop under investigation is very long ($\approx$ 700 Mm) yet the hydrostatic code we employ here ignores gravity, which really should be included. Secondly, there are important problems with how the observational results themselves are obtained. The structure widens as one travels from base to apex; therefore one cannot be sure that you are "sitting" on the same loop structure as one travels along the Priest *et al.*,

Table 8. Summary of the posterior inference for hypothesis H_1 for the SDS3.

	mean	mode	s.d.	2.5%	50%	97.5%	actual value
b	24.38	6.09	21.27	0.74	18.11	80.85	1.00
α	107.19	96.75	29.09	66.26	101.36	176.53	100.68
β	-9.61	-9.22	3.80	-18.48	-9.00	-3.84	-10.00
T_{foot}	1.02	1.00	0.03	0.95	1.02	1.08	1.00

Table 9. Posterior hypotheses odds for the SDS3. Marginal densities have been calculated from: 1: Laplace method with posterior covariance matrix, 2: Laplace method with robust posterior covariance matrix, 3: Monte Carlo estimation with the probability density from stage 1 as the additional probability density.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
1	6.28×10^6	135.88	1	1.16×10^7
2	7.59×10^6	156.74	1	7.40×10^6
3	1.70×10^7	295.51	1	1.73×10^7

Table 10. Information criteria for the SDS3.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
AIC	-56.50	-38.13	-35.94	-56.50
BIC	-54.24	-36.43	-33.68	-54.24

chosen data points. Thirdly, other papers question how the background emission has been extracted from the images. Thus, it could be regarded that this dataset is not very good example for this analysis. However, much interest has been generated by this paper and Bayesian analysis methods have never before been applied to this dataset.

Since the only information of the data that we have are the error bars, we assume a Normal distribution for the data with $pr_i = 0.98$, $i = 1, \ldots 74$. The summary statistics for the four parameters are presented in Tables 11 and 12. From Table 11 we can see that a 95% credible interval for β is between 1.58 and 3.37, which excludes negative values. In fact the probability of β being negative is $P(\beta < 0) = 0$. Thence, we would expect that H_1 and H_3 will give almost the same results. Since there is that high belief that $\beta > 0$, it might seem pointless to construct Bayes factors. However, for the sake of completeness, we have calculated the values which will be useful for model comparison. Thus, according to Table 12 and using the Monte Carlo estimation with the probability density from stage 1 of the delayed rejection algorithm of Mira (2001) as the additional probability density, we obtain the posterior hypotheses odds $2.13 \times 10^7 : 6.17 \times 10^6 : 390.96 : 1$ for the hypotheses $H_3 : H_1 : H_2 : H_4$ respectively. AIC and BIC (see Table 13) agree with the Bayes factors estimates and suggest that H_1 and H_3 are the best hypotheses. Therefore, we conclude that we have basal heating for this loop. This comes in contradiction with the Mackay *et al.*, (2000) conclusion for this specific example. The three fitted curves of the mean, joint mode and median of hypotheses H_1 are depicted in Figure 4.

4.2. Ugarte-Urra Dataset

Using spectral line ratio diagnostic techniques employed upon SOHO/CDS observations, Ugarte-Urra *et al.*, (2005) determined the electron density along an off-limb

Table 11. Summary of the posterior inference for H_1 for the PDS.

	mean	mode	s.d.	2.5%	50%	97.5%
b	41.38	3.11	32.68	1.32	33.34	117.73
α	15.74	12.17	3.33	10.77	15.21	23.43
β	2.47	2.73	0.46	1.58	2.46	3.37
T_{foot}	1.61	1.61	0.00	1.60	1.61	1.61

Table 12. Posterior hypotheses odds for the PDS. 1: Laplace method with posterior covariance matrix, 2: Laplace method with robust posterior covariance matrix, 3: Monte Carlo estimation with the probability density from stage 1 as the additional probability density.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
1	4.64×10^6	182.51	4.47×10^6	1
2	3.99×10^6	234.04	3.72×10^6	1
3	6.37×10^6	390.96	2.13×10^7	1

coronal loop (see Figure 1 in Ugarte-Urra *et al.*, , 2005 for the specific CDS observation). Those authors used a 1D hydrostatic model similar to the one utilised in this paper to generate theoretical density profiles for comparison with the observations. Using a minimum chi-squared analysis, they concluded that the best fit, minimum chi-squared case resulted from a heating function that was weighted preferentially towards the loop base (see Figure 8 in Ugarte-Urra *et al.*, , 2005). This density profile is reanalysed here but now the model comparison step is undertaken using the Bayesian analysis method outlined in Section 2.4.

Since the error bars are not symmetric we assume a Gamma distribution for the data with $pr_i = 0.9973$, $i = 1 \ldots 9$. From the summary statistics in Table 14 we can conclude that the model prefers the negative values of β. To support this we have calculated the probability of β being negative, *i.e.* $P(\beta < 0) \approx 0.90$. Table 15 shows the posterior hypotheses odds for each of the four hypotheses. All of the estimates suggest that the most probable hypothesis is H_4. For example, if we use the Laplace method using the posterior covariance matrix, the posterior hypotheses odds will be $17.06 : 13.72 : 2.13 : 1$ for the hypotheses $H_4 : H_1 : H_2 : H_3$ respectively. The information criteria (see Table 16) suggest the hypotheses (in preference order): H_2, H_3, H_4. Note that hypothesis H_3 is preferable than hypothesis

Table 13. Information criteria for the PDS.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
AIC	269.50	292.15	269.50	294.44
BIC	278.71	299.07	278.71	303.66

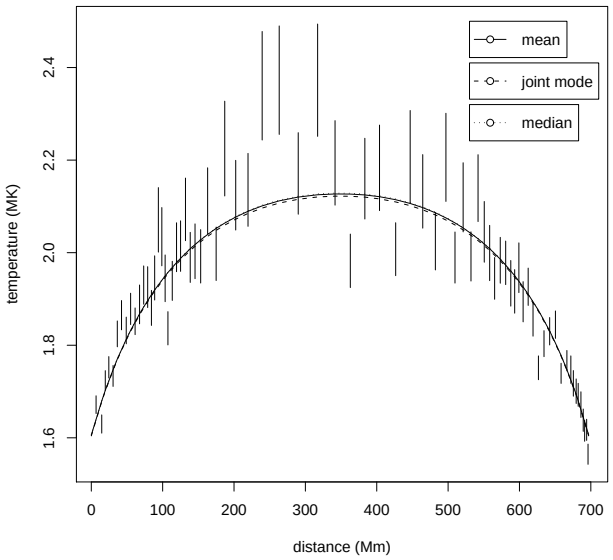

Figure 4. Observational temperature values and fitted temperature profiles against distance along the loop for the PDS. The fitted temperature profiles are constructed using the mean (solid curve), joint mode (dashed curve) and median (dotted curve) values of the parameters taken from Table 11.

Table 14. Summary of the posterior inference for H_1 for the UUDS.

	mean	mode	s.d.	2.5%	50%	97.5%
b	9.02	4.42	7.50	0.31	7.11	27.52
α	33.61	2.47	19.16	4.15	32.40	76.42
β	-10.34	5.16	6.97	-19.85	-11.44	3.60
T_{foot}	1.15	1.05	0.06	1.00	1.16	1.26

H_4. This is because information criteria are focused only in the natural logarithm of the likelihood (plus the correction factor), ignoring the parameters dispersion under each hypothesis. Thus, this result is at odds with the conclusion reached in Ugarte-Urra *et al.*, (2005) as we shall discuss in the following section. If we consider only the hypothesis that maximises the likelihood of the data, then hypothesis H_3 is the best, but when we add the correction factor (the factor that has to do with the number of parameters) hypothesis H_2 is much preferable! Figure 5 illustrates the three fitted curves of the mean, joint mode and median of hypothesis H_1 with the observed data points and error bars.

Table 15. Posterior hypotheses odds for the UUDS. 1: Laplace method with posterior covariance matrix, 2: Laplace method with robust posterior covariance matrix, 3: Monte Carlo estimation with the probability density from stage 1 as the additional probability density.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
1	13.72	2.13	1	17.06
2	13.54	5.05	1	20.29
3	7.02	8.80	1	21.88

Table 16. Information criteria for the UUDS.

	$H_1 : \beta \neq 0$	$H_2 : \beta = 0$	$H_3 : \beta > 0$	$H_4 : \beta < 0$
AIC	-17.30	-18.27	-17.30	-15.79
BIC	-16.51	-17.68	-16.51	-15.00

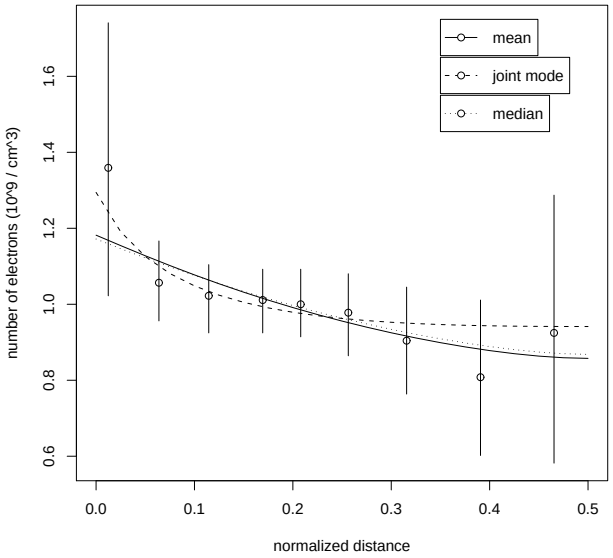

Figure 5. Observational temperature values and fitted temperature profiles against distance along the loop for the UUDS. The fitted temperature profiles are constructed using the mean (solid curve), joint mode (dashed curve) and median (dotted curve) values of the parameters taken from Table 14.

5. Discussion and Future Work

In this paper we have presented a new method for comparing observations with theoretical models for solar astrophysical datasets. Bayesian statistics are generally more powerful than a χ^2 test for the reasons discussed in Section 1. We have examined the robustness of this method with simulated datasets (Section 3), and demonstrated how this can be applied in real datasets (Section 4).

The results from our analysis of the PDS (Section 4.1) show conclusively that we have basal heating for that loop. The combination of 74 observations and narrower error bars compared with the UUDS mean that all diagnostic assessments strongly and unequivocally indicate this form of heat input. However, as it has been mentioned earlier in this paper, different authors have found different spatial forms of the heating function for this same loop. This is likely to occur because different analysis techniques could give different temperature profiles that do not resemble to each other.

The results of our stochastic analysis of the UUDS (Section 4.2) are inconclusive. On the one hand, we have found that those techniques which use maximisation to locate or estimate parameters (AIC or BIC) suggest a uniform heating mechanism ($\beta = 0$). However, if we use an integral approach, *e.g.* Bayes factors, a negative value of β is found.

The resolution of this contradiction is probably found in the examination of the 95% credible interval for β which straddles 0. The marginal distribution of β is negatively skewed, allowing the mode to positive, whilst the majority of the values are negative (the median value is negative). Essentially this is telling us is that we cannot make a firm decision as to the heating mechanism (*i.e.* β negative or positive) from the UUDS. The error bars are simply too large and/or we have insufficient number of observations to allow us to make a heating mechanism determination even with the methodology we have described in this paper.

This leads onto another important point: we need more than one of these diagnostic assessment techniques described in this paper in order to be able to make a rounded reliable judgment on the nature of the heating mechanism. Our recommendation is to use the 95% credible intervals, together with a maximisation and integration based method. Use of say, the χ^2 method alone may lead to a false conclusion, *e.g.* with the UUDS we may have decided that a basal heating mechanism is appropriate for the loop being observed, when in actual fact we can draw no conclusions as to the mechanism.

The fact that information criteria seem to select $\beta = 0$ (uniform heating) as the hypothesis of choice for UUDS may simply reflect a position of the error bars being too wide in comparison with the magnitude of β, which in turn determines the strength of basal or apex heating. Therefore, even if a loop is basally or apex heated, the size of β may simply be too small to be "detected" by the available data, and, rather like the null hypothesis in classical statistical testing, a conclusion of "uniform heating" may be decided upon. Improved data may then come to a very different conclusion on the same loop!

It is clear from the preceding discussion that the Bayesian MCMC methodology described in this paper is a significant development in the assessment of solar coronal loop data. Simply choosing the model paradigm which happens to furnish the minimum χ^2 value can be problematic. In particular, this Bayesian MCMC approach allows for a quantitative assessment of the parameters simultaneously.

However, it must be kept in mind that this model comparison is only as good as the analysed data to which it is applied. In determining any spatial variation in the thermal/density structure along a loop and relating this to a model form, the main drivers are the number of observed data points along the structure under observation and the size of the error bar associated with each data point. Of course, it is the case that you would want to maximise one (the number of observations obtained) and minimise the other (to produce the smallest error bar).

Given that the spatial resolution of new (future) instrumentation are (will be) an improvement on that considered in this paper, it is likely that the number of data points along a structure will not be a vital issue, assuming one is dealing with a loop of reasonable lengh ($>$ 100Mm say). For example, the spatial resolution of Hinode/EIS is over twice that of SOHO/CDS. Similarly, a decrease in the size of the associated error bars should occur with greater instrument sensitivity — however, it is possible that with greater spatial resolution, longer exposure times may be required.

With this in mind, future work in this area will include examining the density structure along many loop examples observed by Hinode/EIS. Also, the numerical scheme will be extended to include gravity, relevant to longer loops.

Acknowledgements SA has been supported by an STFC grant.

References

Adamakis, S., Morton-Jones, T., Walsh, R.: 2007, Applications of Metropolis-Hastings Algorithm to Solar Astrophysical Data. In: Babu, G., Feigelson, E. (eds.) *Statistical Challenges in Modern Astronomy IV*. Astronomical Society of the Pacific Conference Series, 401 – 402.

Akaike, H.: 1974, A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control* **19**, 716 – 723.

Aschwanden, M.: 2001, Revisiting the Determination of the Coronal Heating Function from Yohkoh Data. *The Astrophysical Journal* **559**, 171 – 174.

Bartlett, M.S.: 1957, A Comment on D. V. Lindley's Statistical Paradox. *Biometrika* **44**, 533 – 534.

Gelfand, A., Dey, D.: 1994, Bayesian Model Choice: Asymptotics and Exact Calculations. *Journal of the Royal Statistical Society. Series B (Methodological)* **56**, 501 – 514.

Gelman, A., Carlin, J., Stern, H., Rubin, D.: 2003, Bayesian Data Analysis. Chapman and Hall.

Hastings, W.K.: 1970, Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika* **57**, 97 – 109.

Hildner, E.: 1974, The Formation of Solar Quiescent Prominences by Condensation. *Solar Physics* **35**, 123 – 136.

Kass, R.E.: 1993, Bayes Factors in Practice. *Statistician* **42**, 551 – 560.

Kass, R.E., Greenhouse, J.: 1989, Comment on Investigating Therapies of Potentially Great Benefit: ECMO. *Statistical Science* **4**, 310 – 317.

Kass, R.E., Raftery, A.: 1995, Bayes Factors. *Journal of the American Statistical Association* **90**, 773 – 795.

Lindley, D.V.: 1957, A Statistical Paradox. *Biometrika* **44**, 187 – 192.

Mackay, D.H., Galsgaard, K., Priest, E., Foley, C.: 2000, How Accurately can we Determine the Coronal Heating Mechanism in the Large-Scale Solar Corona? *Solar Physics* **193**, 93 – 116.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A.: 1953, Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **21**, 1087 – 1092.

Mira, A.: 2001, On Metropolis-Hastings Algorithms with Delayed Rejection. *Metron* **59**, 231 – 241.

Neal, R.M.: 2003, Slice Sampling. *The Annals of Statistcs* **31**, 705 – 767.

Peres, G.: 2000, Models of Dynamic Coronal Loops. *Solar Physics* **193**, 33 – 52.
Priest, E., Foley, C., Heyvaerts, T., Arber, T., Mackay, D., Culhane, J., Acton, L.: 2000, A Method to Determine the Heating Mechanisms of the Solar Corona. *The Astrophysical Journal* **539**, 1002 – 1022.
Raftery, A.E.: 1996, Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models. *Biometrika* **83**, 251 – 266.
Raftery, A.E.: 1996, Hypothesis Testing and Model Selection. In: Gilks, W., Richardson, S., Spiegelhalter, D. (eds.) *Markov Chain Monte Carlo in Practice*. Chapman and Hall, 163 – 188.
Reale, F.: 2002, More on the Determination of the Coronal Heating Function from Yohkoh Data. *The Astrophysical Journal* **580**, 566 – 573.
Schwarz, G.: 1978, Estimating the Dimension of a Model. *Annals of Statistics* **6**, 461 – 464.
Tierney, L., Mira, A.: 1999, Some Adaptive Monte Carlo Methods for Bayesian Inference. *Statistics in Medicine* **18**, 2507 – 2515.
Ugarte-Urra, I., Doyle, J.G., Walsh, R.W., Madjarska, M.S.: 2005, Electron Density Along a Coronal Loop Observed with CDS/SOHO. *Astronomy & Astrophysics* **439**, 351 – 359.
Walsh, R., Bell, G.E., Hood, A.W.: 1995, Time-Dependent Heating of the Solar Corona. *Solar Physics* **161**, 83 – 101.