

# A Bayesian Analysis of MHD Waves in the Lower Atmosphere

M. S. Marsh, J. Ireland and T. Kucera

NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.  
email: michael.s.marsh@nasa.gov

**Abstract.** Magneto-hydrodynamic wave modes propagating from the solar photosphere into the corona have the potential to be exploited as an observational tool in an analogous way to the use of acoustic waves in helio/terrestrial seismology. In regions of strong magnetic field photospheric p-modes are thought to undergo mode conversion to slow magneto-acoustic waves, and that these slow magnetoacoustic p-modes may be waveguided from the photosphere into the solar corona along the magnetic field. A Bayesian analysis technique is applied to observations which suggests four distinct p-modes may be resolved in the transition region.

**Keywords.** Methods: statistical, MHD, Sun: atmosphere, Sun: photosphere, Sun: chromosphere, Sun: transition region, Sun: Corona, Sun: oscillations, waves

---

## 1. Introduction

The work of Jaynes (1987) and others has shown that the application of Bayesian statistical techniques has many applications to physics. The use of Bayesian methods has great potential with astrophysics and solar physics. Current analysis techniques applied to the problem of wave detection and parameterization in solar physics are not optimal to the problem at hand. It is possible to extract much more information contained within the data by applying Bayesian statistical methods, compared to the traditional Fourier or wavelet analysis currently employed.

We apply the methods described by (Jaynes 1987, and references in Marsh et al. 2007) to the problem of frequency estimation within solar data. A Bayesian code is applied to artificial time series data typical of oscillations within the solar corona, to demonstrate the high precision frequency estimation that can be achieved. It is shown that very closely spaced frequencies can be resolved with the Bayesian code where this is not possible with the Fourier transform. The code is then applied to solar data which suggests p-mode propagation into the transition region.

## 2. Bayesian Oscillation Probability

When applied to the question of oscillation detection, we wish to compute the probability of a particular time series model, given the data and all other prior information (See Marsh et al. 2007, for specific details of the Bayesian method).

### 2.1. The probability density function

The posterior probability density that a general oscillatory model is present within the data is given by:

$$P(\{\omega\}|D, I) \propto \left[1 - \frac{m\overline{h^2}}{N\overline{d^2}}\right]^{\frac{m-N}{2}} \quad (2.1)$$

This probability density has been derived as a function of the frequency parameters only  $\{\omega\}$ , assuming data with an unknown noise variance. Where  $m$  is the number of component model functions,  $N$  is the number of measurements in the data time series and  $\overline{d^2}$  is the mean square value of the data.

It is the  $\overline{h^2}$  function which carries the frequency dependence of the probability density. Given by:

$$\overline{h^2} = \frac{1}{m} \sum_{j=1}^m h_j^2, \quad (2.2)$$

where

$$h_j = \sum_{i=1}^N d_i H_j(t_i), \quad (1 \leq j \leq m). \quad (2.3)$$

The  $h_j$  functions are the projections of the data onto orthonormal model functions, and  $\overline{h^2}$  is the mean square value of these projections as a function of  $\{\omega\}$ . The maximum of this function gives the most probable frequency  $\hat{\omega}$ , supported by the data, for each of the component functions assumed by the model. The corresponding maximum in the probability density function (Eqn. 2.1) is sharply peaked at these frequency values  $\hat{\omega}$ , since the form of the function is similar to an exponential. This allows very precise frequency estimates to be made, at a resolution much higher than can be estimated from the Fourier transform, as described in Sect. 3.

### 3. Parameter Estimation

Although, in the single frequency case, the periodogram, Fourier transform and  $\overline{h^2}$  are similar;  $\overline{h^2}$  has been derived from probability theory and can be understood in a statistical sense. This understanding allows estimates of the model parameters, and their precision, to be derived. These estimates cannot be made from the Fourier transform or periodogram alone without understanding their origin in probability theory.

#### 3.1. The expected noise variance $\langle \sigma^2 \rangle$

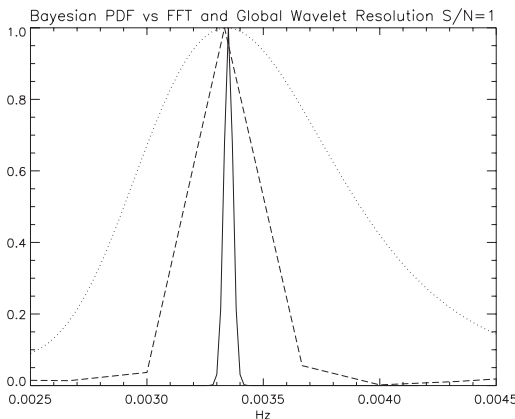
The Bayesian analysis allows the expectation value of the noise variance within the data to be calculated as:

$$\langle \sigma^2 \rangle = \frac{1}{N - m - 2} \left[ \sum_{i=1}^N d_i^2 - \sum_{j=1}^m h_j^2 \right]. \quad (3.1)$$

The expectation value of the noise variance is essentially the difference between: the total square value of the data, and the total square value of the data projected onto the orthonormal model functions. It is implicit in the Bayesian model that everything within the data that is not fitted by the model is assumed to be noise. A method to determine the point at which the model best represents the signal is described in Sect. 5.1

#### 3.2. The frequency parameters $\{\omega\}$

The most probable frequencies, of the applied model, are given by the location of the maximum within the probability density function as described in Sect. 2.1. The accuracy of the frequency parameters can be estimated by expanding  $\overline{h^2}$  (Eqn. 2.2) in a Taylor



**Figure 1.** A comparison of the frequency resolution obtained for the single frequency data with a S/N=1, with the normalized Bayesian probability density function (solid line), FFT (dashed) and global wavelet transform (dotted).

**Table 1.** Obtained frequency resolutions for the single frequency data with a S/N=1.

Analysis method	$\delta f$ (mHz)
Bayesian PDF	0.02
FFT HWHM	0.2
Wavelet HWHM	0.5

series. This accuracy is dependent on the matrix  $b_{jk}$ , related to the Hessian matrix of  $\overline{h^2}$  evaluated at the most probable model frequencies  $\hat{\omega}$ :

The estimated frequency resolution is given by the variance of the probability density function for  $\omega_k$ :

$$\sigma_{\omega_k}^2 = \langle \sigma^2 \rangle \sum_{j=1}^r \frac{u_{jk}^2}{v_j}. \quad (3.2)$$

Where  $u_{jk}$  is the  $k$ th component of the  $j$ th eigen vector of  $b_{jk}$  with a corresponding eigen value  $v_j$ , and the expected noise variance  $\langle \sigma^2 \rangle$  is evaluated at the most probable frequencies. These most probable frequencies are then equal to  $\hat{\omega} \pm \sigma_{\omega_k}$ .

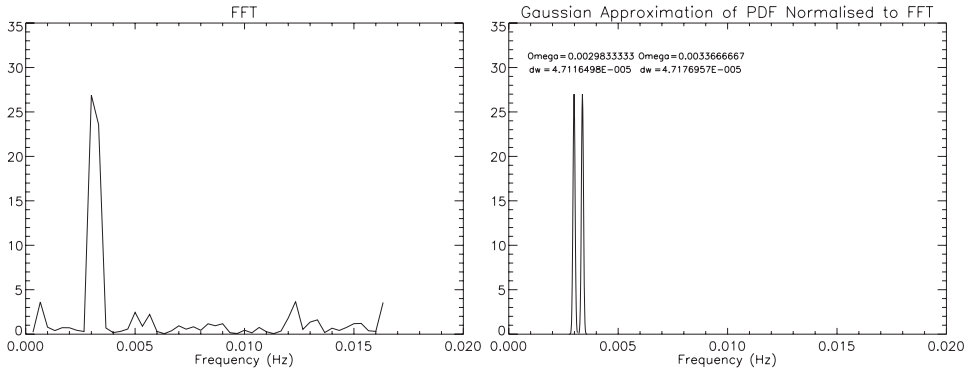
Equation 3.2 shows that the obtainable frequency resolution is related to how sharply the probability distribution is peaked around the most probable frequencies and the magnitude of the noise variance within the data.

## 4. Frequency Resolution

In this section, we apply the Bayesian model to artificial test data, typical of the type of oscillations that are observed within the solar corona using current instrumentation.

### 4.1. The single frequency case

Here we compare the results obtained from the Bayesian model with those obtained by performing a Fourier, or wavelet analysis. The analyzed time series is of the form:  $d_i = [\text{Cos}(\omega t_i) + \text{Sin}(\omega t_i)] + e_i$ , where  $\omega=3.3$  mHz and  $e_i$  is Gaussian distributed noise. This is typical of the ‘5-minute’ like oscillations observed in coronal loops with 100 samples of 30 s cadence and has a low signal to noise ratio with a RMS S/N=1. The Bayesian



**Figure 2.** Left: FFT of the two frequency time series separated by 1 Fourier Nyquist step with a S/N=1. Right: Gaussian approximations of the Bayesian probability density function for a two harmonic frequency model, normalized to the FFT.

**Table 2.** Estimated frequencies and  $1\sigma$  errors obtained from the Bayesian model applied to the two frequency time series separated by one Fourier Nyquist step and a S/N=1.

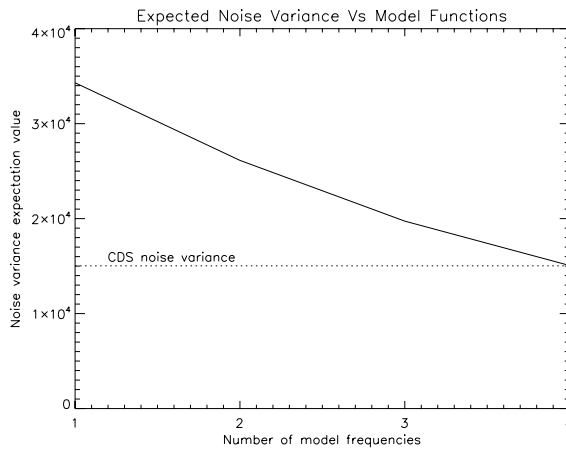
Resolved frequencies (mHz)	True value (mHz)
$2.98 \pm 0.05$	3.0
$3.37 \pm 0.05$	3.3

model is applied to calculate the probability density function for a single harmonic frequency model of the time series, using Eqn. 2.1. The peak within the probability density function, representing the most probable frequency, can be described by a Gaussian of width  $\sigma_{\omega_k}$  given by Eqn. 3.2. This width determines the theoretical frequency resolution obtainable with the Bayesian model; this resolution is much less than the estimated frequency resolution obtained using the Fourier transform.

Figure 1 shows the obtainable resolving power of the Bayesian model, normalized to that obtained by the Fourier and wavelet transforms. The solid line indicates the probability density function for the single frequency model, the dashed line shows the FFT and the dotted line is the global wavelet transform using the Morlet wavelet function. We can see that the Bayesian model gives a significant increase in the resolution of the estimated frequency. Table 1 lists the obtained frequency resolution for each method, if we estimate the resolution of the Fourier and global wavelet transforms using the half width half maximum (HWHM). We see that, for a S/N=1, the  $1\sigma$  error on the estimated frequency from the Bayesian model gives an order of magnitude increase in resolution over the FFT. The global wavelet has an even lower resolution due to the smoothing effect of the transform on the wavelet scale. In fact, if we are interested in high precision frequency measurements of stationary frequencies, or closely separated frequencies, then a wavelet analysis is one of the worst methods we can apply.

#### 4.2. Two closely separated frequencies

We now compare the results obtained from a Bayesian model and a Fourier analysis of two closely separated frequencies within a simulated time series. Again we generate a time series typical of coronal loop oscillations of the form:  $d_i = [A_1 \cos(\omega_1 t_i) + B_1 \sin(\omega_1 t_i)] + [A_2 \cos(\omega_2 t_i) + B_2 \sin(\omega_2 t_i)] + e_i$ , with 100 samples at 30 s cadence, a harmonic frequency of  $\omega_1=3.3$  mHz, an additional frequency of  $\omega_2=3.0$  mHz, Gaussian distributed noise  $e_i$



**Figure 3.** Change in the expectation value of the noise variance for an increasing number of component frequencies within the Bayesian model. The dotted line indicates the known noise variance within the CDS data.

and a RMS  $S/N=1$ . These two frequencies separated by only one Nyquist step in the Fourier transform, so in principle their frequencies are directly adjacent in the FFT.

Figure 2 shows the FFT for the two frequency time series and the Gaussian representation of the resolution obtained from the Bayesian probability density function which has been normalized to the FFT for comparison. As expected the FFT is unable to resolve two such closely separated frequencies. A single broad peak is observed, with a large HWHM, suggesting the possibility that more than one frequency may be present. However, the result from the Bayesian model resolves the two frequencies independently and to a very high precision even with a  $S/N=1$ . Table 2 lists the resolved frequencies and their  $1\sigma$  errors, estimated from the probability density function for a two harmonic frequency Bayesian model. We see that the Bayesian model not only resolves the two frequencies but does so to a very high precision with  $1\sigma$  errors of 0.05 mHz.

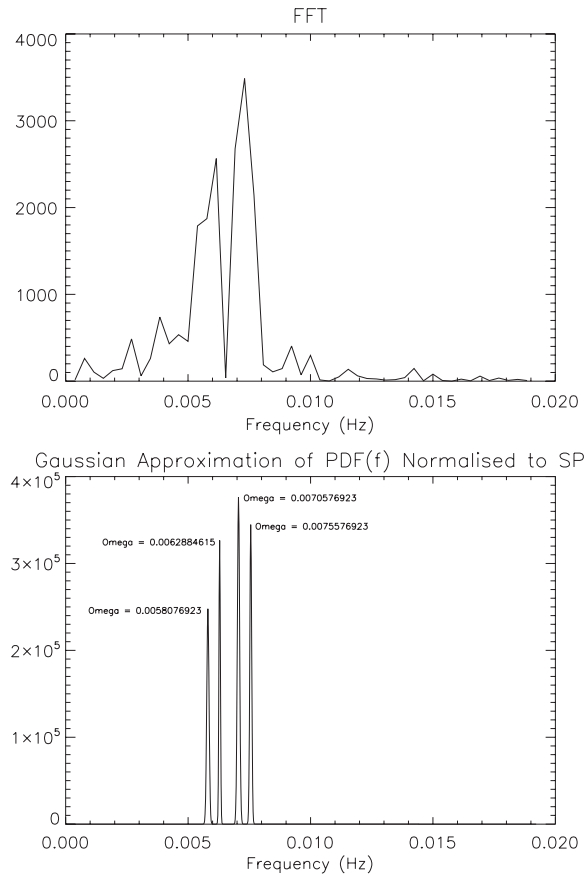
## 5. Application to Solar Data

The Bayesian model is now applied to real observations of solar oscillations. Marsh & Walsh (2006) observe the apparent propagation of slow-magnetoacoustic waves within a sunspot region. These waves are observed to propagate from the transition region into the coronal loop system emerging from the sunspot and are interpreted as the propagation of photospheric p-modes waveguided along the magnetic field.

The original analysis applied Fourier techniques to the time series; here we apply the Bayesian model to the O V data described in Marsh & Walsh (2006). The results presented in Marsh & Walsh (2006) show the presence of two ‘3-min’ frequencies observed in the transition region above the sunspot umbra. The data consist of 100 samples observed at 30 s cadence obtained with the Coronal Diagnostic Spectrometer (CDS, see Harrison et al. 1995). The Bayesian model is applied iteratively, with the addition of further model functions to increase the complexity of the applied model.

### 5.1. *What is a good model*

In addition to the problem of fitting a model to the data, it needs to be determined what is a good model. If we already have a knowledge of the noise variance within the data,



**Figure 4.** Top: The FFT of the O V data originally presented in Marsh & Walsh (2006) showing two frequencies in the 3-min range. Bottom: The Gaussian approximations of the probability density function for the four harmonic frequency model, normalized to the Schuster periodogram.

then the calculated expectation value of the noise variance (Eqn 3.1) may be used to determine when the model has accounted for the real signal.

The noise properties of the CDS instrument are described by Thompson (2000). We can use this known value of the noise variance to determine when the model has reached sufficient complexity to account for the real signal and the addition of further model functions will start to fit the noise.

Figure 3 shows the change in the expectation value of the noise variance for increasingly complex models. The dotted line shows the noise variance within the data due to the photon statistics and the noise properties of the CDS detector. The expected noise variance reaches the level of the CDS data with a model containing four frequencies, thus the data best supports this model. We are then able to derive estimates of these frequencies and their associated errors from the probability density function of the four frequency model.

Figure 4 shows the FFT of the O V data presented in Marsh & Walsh (2006) and the Gaussian approximation of the four frequencies resolved by the Bayesian model, listed in Table 3. As demonstrated in Sect. 4.2, the Bayesian model is able to resolve closely spaced frequencies to a much higher resolution than is possible with the Fourier

**Table 3.** The four frequencies resolved by the Bayesian model applied to the O V data from Marsh & Walsh (2006), and their associated  $1\sigma$  errors.

Resolved frequencies (mHz)	$\sigma_\omega$ (mHz)
5.81	0.05
6.29	0.03
7.06	0.05
7.56	0.04

transform; where the FFT can resolve only two frequencies, the Bayesian model is able to resolve four independent frequencies within the data, to a very high precision.

## 6. Conclusions

Considering the problem of frequency estimation within a time series, the Bayesian method returns very precise estimates and a consistent statistical error analysis. It is not clear how to determine frequency error estimates from the Fourier transform, without understanding its relation to probability theory. It is often misconceived that the Fourier Nyquist spacing is the limit to the resolution with which a frequency can be resolved within a time series. This is not the case, as the resolution limit is principally determined by the signal to noise ratio. This is not surprising as the Bayesian method has similarities with least squares fitting of a function. In an analogous way, the limiting resolution with which a sinusoidal function can be fitted to oscillating data is not equal to the time series cadence; it is largely determined by the signal to noise ratio within the data.

We apply the Bayesian model to the O V data presented in Marsh & Walsh (2006), and resolve four closely spaced independent frequencies within the 3-min period range. The observations presented in Marsh & Walsh (2006) are interpreted as the conversion and propagation of photospheric p-mode oscillations along the magnetic field into the corona. The results presented here suggest that we may now be able to resolve these oscillations into four closely spaced p-mode frequencies. Zhukov (2002) calculate the spectrum of eigen modes within the vertical magnetic field of the sunspot umbra, and finds the 3-min umbral oscillations are due to p-modes modified by the strong magnetic field within the sunspot. The frequencies detected here, and their spacing, are consistent with the results of Zhukov (2002) and may represent the detection of the  $P_5$ ,  $P_6$ ,  $P_7$  and  $P_8$  photospheric p-modes in the transition region.

## Acknowledgements

M.S. Marsh is supported by the NASA/ORAU postdoctoral program and wishes to acknowledge the useful comments of L.E. Pickard.

## References

- Harrison R.A., Sawyer E.C., Carter M.K., et al., 1995, *Sol.Phys.*, 162, 233.  
 Jaynes E.T., 1987, In: Jaynes, E. T. (1987) *Bayesian Spectrum and Chirp Analysis*, in Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems, C. R. Smith & G. J. Erickson (editors), D. Reidel Publishing Company, Holland; pp. 1–37.  
 Marsh M.S., Walsh R.W., May 2006, *ApJ*, 643, 540.  
 Marsh M.S., Ireland J., Kucera T., 2007, *ApJ*, *Submitted*.  
 Thompson W.T., 2000, *CDS Software Note*, #49.  
 Zhukov V.I., May 2002, *A&A*, 386, 653.