Bayesian time-varying AR spectral estimation for ultrasound attenuation measurement in biological tissues

Jérôme Idier†, Jean-François Giovannelli† and Bernard Querleux‡
†Laboratoire des Signaux et Systèmes (CNRS/ESE/UPS)
Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France
‡Laboratoires de recherche de L'Oréal,
93601 Aulnay-Sous-Bois Cedex, France

ABSTRACT: Ultrasonic characterization of biological tissues provides a cheap and harmless support for diagnosis and control. In this area, acoustical attenuation is a relevant physical parameter for skin analysis. The attenuation properties of the investigated tissue can be deduced from series of instantaneous spectra indexed by the investigation depth. Following KITAGAWA & GERSCH [4], a **smoothness prior time-varying AR model** provide robust short-range analysis by adaptive AR modeling. The novelty resides in three points:

- The form of the prior, which accounts for the fact that linear attenuation induces linear variation of the first AR coefficient while the second one remains roughly invariant.
- In order to cope more efficiently with our vectorial context, we propose to substitute the little known INFORMATION FORM for the classical covariance Kalman form.
- In the case of a stationary signal with Gaussian PSD, we devise exact relations between theoretical second-order regression coefficients and the PSD.

Experiments on skin attenuation measurement support the fact that the proposed Bayesian approach provides more reliable results than periodogram and usual least squares autoregressive techniques.

KEY WORDS: attenuation measurement, adaptive spectral estimation, smoothness prior, Kalman smoothing.

1. MOTIVATION

Efforts dedicated to attenuation measurement originate in the idea that attenuation of ultrasounds in biological tissues depends on their nature or state. For instance, Kuc's experiments try to diagnose liver disease. Here we are interested in the characterization of different skin attenuation according to the age of the patient as well as in showing that some treatments induce a noticeable modification of skin attenuation. For sake of statistical evidence, a reliable and systematical method of estimation based upon an echographic data set is necessary.

2. PRINCIPLES

The acoustical attenuation coefficient $A(f)$ of a homogeneous tissue at frequency f characterizes the modification of the spectral content of an ultrasonic wave as it propagates in the medium. From [5], the power spectral density (PSD) $S_d(f)$ when a distance d is covered ($2d$ in reflection mode) is:

$$S_d(f) = S_0(f) \exp -4dA(f). \tag{1}$$

[5] proposes two methods to estimate tissue attenuation.

Spectral difference. According to the first idea developed in [5], $A(f)$ is simply deduced from differences between log-spectra at variable depths.

Spectral shift. Here we are rather interested in the second method which lies on two hypothesis.

In the domain of frequencies lower than 10MHz, assumption of attenuation increasing proportionately with frequency: $A(f) = \beta f$ is valid. Such an assumption reduces the problem of medium characterization to the scalar parameter β estimation. Though the useful frequency band in our case is rather close to 25MHz we retain the working hypothesis of linear attenuation.

The spectral shift method assumes an incident pulse with a Gaussian-shaped PSD:

$$S_0(f) = K_0 \exp -\frac{(f-f_0)^2}{2\sigma_0^2}. \tag{2}$$

From Eq. (1) it is easy to check that:

$$S_d(f) = K_d \exp -\frac{(f-f_d)^2}{2\sigma_d^2},$$

with

$$f_d = f_0 - 4\beta d\sigma_0^2 \tag{3}$$

$$\sigma_d^2 = \sigma_0^2. \tag{4}$$

So the PSD remains Gaussian during its propagation. The spectrum mean decreases of a characteristic quantity of attenuation while its variance remains constant. Equations (3) and (4) are sufficient to deduce β under the following form:

$$\beta = \frac{f_0 - f_d}{4d\sigma_0^2} \text{ nepers/cm/MHz}. \tag{5}$$

The major interest of the spectral shift method consists in the search of a linear evolution of a mean frequency curve preserving a local meaning even if

attenuation behavior differs from linearity within acceptable limits. Therefore, the mean frequency and spectral variance sequence depict the key information proceeding from the PSD sequence, under an easily visualisable shape.

From this standpoint and whatever the validity of linear attenuation and Gaussian PSD assumptions the quantities of the spectral shift method have a more immediate and expressive physical meaning than the parameters of the spectral difference method. These considerations explain why we favored the spectral shift method, more restrictive in some sense (it is based upon Gaussian spectrum) but more expressive from a practical standpoint.

From the above definitions attenuation measurement is based upon PSD estimation according to the distance covered by the ultrasonic wave in the medium. Adaptive PSD estimation is in the scope of the next two sections.

3. ADAPTIVE SPECTRAL ESTIMATION

3.1. CLASSIC METHODS

An experimental study has been based on 9 sets of B-mode echographic data from the same context. Fig. 1 shows one of the data sets. The first step of treatment consists of undersampling, alignment and selection of unsatured echogeneous areas. Such pre-processing has been applied before any further treatment. For instance the data set of Fig. 1 provides pre-processed data depicted by Fig. 2.

FFT analysis. Periodogram computation is the basic and most popular method for spectral analysis based on simple Fast Fourier Transforms (FFT). A set of 16 samples periodograms has been calculated and averages over the 75 traces have been performed, one at each depth. Computations of mean frequencies and spectral widths yield the results shown on figure 3 and 4.

These results suffer from the fundamental default of the periodogram *i.e.* strong variability of the estimated spectral parameters. On the one hand high variability from an image to the other is observed. On the other hand, each mean frequency or spectral width sequence is too rough to allow reliable regression.

Autoregressive analysis. The well known least squares AR approach allows better robustness. Yet results of 5 and 6 also show strong variability. They have been obtained using second order AR models, which is a simple way to approximate Gaussian spectra. Coefficients have been estimated using a vectorial form of a 16 samples sliding window least squares.

Both periodogram and AR results lack of robustness because estimation is performed independently at each depth. Reduction of variability is aimed in the next section by accounting for the slowly varying character of the parameters in the medium.

3.2. SMOOTHNESS PRIOR TIME-VARYING AR METHODOLOGY

Following KITAGAWA & GERSCH [4], a **smoothness prior time-varying AR model** is introduced. Short-range analysis is still provided by adaptive second order AR modeling, but the slowly varying character of spectra according to the analysis depth is introduced via a Brownian modeling of the regression coefficients series.

For each measured signal $\{x(n)\}$, let us consider the following state model:

$$\begin{array}{ll} \text{State equation} & \boldsymbol{a}_{n+1} = \boldsymbol{a}_n + \boldsymbol{b}_n, \\ \text{Observation equation} & x(n) = \boldsymbol{x}_n^t \boldsymbol{a}_n + \varepsilon(n). \end{array} \quad (6)$$

Then Kalman smoothing yields optimal Bayesian estimation of series $\{\boldsymbol{a}_n\}$ according to the classical updating equations [1]:

- Prediction of $x(n)$:

$$\boldsymbol{a}_{n|n-1} = \boldsymbol{a}_{n-1|n-1} \quad (7)$$

$$P_{n|n-1} = P_{n-1|n-1} + P_b \quad (8)$$

- Correction step after acquisition of $x(n)$:

$$\boldsymbol{k}_n = P_{n|n-1}\boldsymbol{x}_n \quad (9)$$

$$r_n = \sigma_\varepsilon^2 + \boldsymbol{k}_n^t \boldsymbol{x}_n \quad (10)$$

$$\boldsymbol{a}_{n|n} = \boldsymbol{a}_{n|n-1} + \boldsymbol{k}_n r_n^{-1} \left(x(n) - \boldsymbol{a}_{n|n-1}^t \boldsymbol{x}_n \right) \quad (11)$$

$$P_{n|n} = P_{n|n-1} - \boldsymbol{k}_n r_n^{-1} \boldsymbol{k}_n^t \quad (12)$$

- Time reversed smoothing step:

$$A_n = P_{n|n} P_{n+1|n}^{-1} \quad (13)$$

$$\boldsymbol{a}_{n|N} = \boldsymbol{a}_{n|n} + A_n \left(\boldsymbol{a}_{n+1|N} - \boldsymbol{a}_{n+1|n} \right) \quad (14)$$

$$P_{n|N} = P_{n|n} + A_n \left(P_{n+1|N} - P_{n+1|n} \right) A_n^t \quad (15)$$

3.3. VECTORIAL SETTING OF KALMAN SMOOTHER

When a set of M signals (such as Fig. 2) is available instead of a single measured signal, estimation of a single series of regression vectors from joint processing remains preferable than *ad hoc* merging of M estimated AR series. In the proper vectorial setting, the state equation remains the same, while the observation vector becomes $\boldsymbol{x}(n) = [x(n,1), \ldots, x(n,M)]^t$ and the corresponding observation equation reads:

$$\boldsymbol{x}(n) = X_n^t \boldsymbol{a}_n + \boldsymbol{e}(n),$$

where $\boldsymbol{e}(n) = [\varepsilon(n,1), \ldots, \varepsilon(n,M)]^t$ is the noise vector with covariance $\sigma_\varepsilon^2 I$. $X_n = [\boldsymbol{x}_n(1), \ldots, \boldsymbol{x}_n(M)]$ is a matrix $N \times M$ of past observations.

Both the prediction step (7)-(8) and the smoothing step (13)-(15) remain unchanged whereas the correction step becomes:

$$K_n = P_{n|n-1} X_n \tag{16}$$

$$R_n = \sigma_\varepsilon^2 I + K_n^t X_n \tag{17}$$

$$\boldsymbol{a}_{n|n} = \boldsymbol{a}_{n|n-1} + K_n R_n^{-1} \left(\boldsymbol{x}(n) - X_n^t \boldsymbol{a}_{n|n-1}\right) \tag{18}$$

$$P_{n|n} = P_{n|n-1} - K_n R_n^{-1} K_n^t \tag{19}$$

The scalar quantity r_n of equation (10) has been converted in (17) into a matrix R_n of size $M \times M$ to be inverted at each step. Such an increase of computational load is clearly intractable for practical values of M. However numerical complexity can be lowered again to very acceptable proportions if we substitute the so called information form of Kalman filtering for the previous covariance form. Such denominations proceed from the statement that the latter form is based upon updating covariance matrices $P_{n|k}$ while the former is based upon the inverse $P_{n|k}^{-1}$, *i.e.* a Fisher information matrix.

Starting from equations (17) and (19), simple applications of the matrix inversion lemma provides the lesser known information form of Kalman filtering correction step:

$$P_{n|n}^{-1} = P_{n|n-1}^{-1} + \sigma_\varepsilon^{-2} X_n X_n^t \tag{20}$$

$$\boldsymbol{a}_{n|n} = \sigma_\varepsilon^{-2} P_{n|n} (X_n \boldsymbol{x}(n) + P_{n|n-1}^{-1} \boldsymbol{a}_{n|n-1}) \tag{21}$$

Calculation of $\boldsymbol{a}_{n|n}$ still requires inversion of the $p \times p$ matrix $P_{n|n}^{-1}$ but it is of low cost ($p = 2$ or 3 in the present context) compared to the inversion of the $M \times M$ matrix R_n ($M = 100$ in practical cases). Let us finally mention that no modification is needed in the smoothing step (13)-(15).

4. APPLICATION TO SHORT-RANGE TRACKING OF A GAUSSIAN PSD

4.1. ESTIMATION OF SPECTRAL MOMENTS

Relevant properties concerning second order AR coefficients are presented now, in order to make the connection between regression coefficients and the mean and variance of a Gaussian PSD for a given stationary signal embedded in additive noise.

Noiseless case. A first estimation method of a Gaussian spectrum from AR coefficients a_1 and a_2 would be to compute the empirical mean and standard deviation of the second order AR spectrum in the frequency domain $\nu \in [0; 1/2]$. This method gives acceptable results but evaluation of the spectrum on a fine frequency grid between 0 an 1/2 is necessary to deduce the spectral moments, in the absence of known analytical results for the corresponding integrals.

A alternate approach would to consider that the spectrum is approximately Gaussian with same position of the maximum and same curvature at this maximum. Practically this method is less expensive because explicit expressions of both the maximum and the curvature are known. It also provides rather reliable estimates of the Gaussian central frequencies. On the other hand, as second order AR spectra are more spiky than Gaussian spectra of same variance, computation of variances from curvatures leads to over estimated attenuation coefficients.

Both previous methods approximate the assumed Gaussian spectrum by the AR DSP yielded by regression. In fact such approximation is not necessary as it is possible to establish a mathematical link between AR coefficients a_1 and a_2 and Gaussian spectral characteristics, via the first correlation lags of the process. The advantage of such an approach is to exploit directly the second order regression coefficients result to estimate the spectrum instead of assuming that the process is a second order AR.

On one hand the well known link between the p first correlation coefficients of a wide sense stationary process (not necessary autoregressive) and the p second order optimal regression coefficients is given by the Yule-Walker equation. For $p = 2$:

$$\begin{bmatrix} r_0 & r_1 \\ r_1 & r_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{22}$$

where $r_k = \mathrm{E}\{x(n)x(n-k)\}$ denotes the kth correlation coefficient. Given r_0, inversion of (22) gives:

$$\begin{cases} r_1 = r_0 a_1/(1-a_2) \\ r_2 = r_1 a_1 + r_0 a_2. \end{cases} \tag{23}$$

On the other hand let us connect the correlation coefficients r_1 and r_2 of the process to the characteristics of the Gaussian PSD:

$$dsp_x(\nu) = \frac{\sigma_x^2}{2} \left[\mathcal{N}(m_f, \sigma_f^2) + \mathcal{N}(-m_f, \sigma_f^2)\right], \tag{24}$$

with mean frequency m_f, spectral width σ_f and power σ_x^2. Inverse Fourier transform gives the continuous time correlation function:

$$r_t = \sigma_x^2 \quad \exp -2\pi^2 \sigma_f^2 t^2 \quad \cos 2\pi m_f t. \tag{25}$$

Hence the three first correlation coefficients r_0, r_1 and r_2 of the sampled process are:

$$\begin{cases} r_0 = \sigma_x^2 \\ r_1 = \sigma_x^2 \alpha r \\ r_2 = \sigma_x^2 \alpha^4 (2r^2 - 1), \end{cases} \tag{26}$$

where $\alpha = \exp -2\pi^2\sigma_f^2$ and $r = \cos 2\pi m_f$. It remains to inverse (26) in order to deduce r and α from r_1 and r_2. Let us note $a = r_1^4/r_0^3 r_2$. Then easy calculation gives the required result

$$\begin{cases} r = \text{sign}(r_1)\sqrt{a + \sqrt{a^2 - a}} \\ \alpha = r_1/r_0 r, \end{cases} \quad (27)$$

under rather complex conditions of existence, fortunately satisfied in practice.

One can derive m_f and σ_f^2 from a_1 and a_2 via (23) and (27). Let us note that r_0 is of no importance: each quantity can be expressed as a function of r_1/r_0 and r_2/r_0 which are deduced from (23) without the knowledge of r_0.

Noisy case. In practice we must cope with several possible sources of noise, at least from quantification effects. The method can be used under a slightly modified form to take noise of known variance into account. Simple accountance for additive white noise of variance σ^2 is obtained by substitution of $r_0 = \sigma_x^2$ for $r_0 = \sigma_x^2 + \sigma^2$ in (25), whereas r_1 and r_2 remain unchanged since the noise is assumed to be uncorrelated. Knowledge of r_0 still is not necessary, but the signal to noise ratio σ_x^2/σ^2 is required.

4.2. AR 2 STATE EQUATION

The previous regularization method tends to line up the series of spectra because it uniformly penalizes first-order differences between successive regression coefficients. As a consequence the decrease rate of mean frequency is artificially lowered so that attenuation is proportionally underestimated.

It is easy to correct this deficiency in Kalman filtering framework by means of a proper change in the state model. First, let us study the respective role of the two estimated regression coefficients a_1 and a_2. Up to a first order analysis when mean frequency becomes close to $\pi/2$, from (22) and (25), it is easy to check that a_1 and a_2 are proportional repectively to the position of the maximum *i.e.* the mean frequency and to the spectral width. Thus we expect rather constant coefficients a_2 and a linearly varying coefficient a_1.

In the Kalman filter framework, correction goes through a suited modification in the state equation (6). More precisely let us keep a first order model for a_2 and introduce a second order model for a_1:

$$\begin{aligned} a_1(n+1) - a_1(n) &= a_1(n) - a_1(n-1) + b_1(n) \\ a_2(n+1) &= a_2(n) + b_2(n), \end{aligned}$$

These two equations can be rewritten in a new state equation instead of (6):

$$\boldsymbol{h}_{n+1} = F\boldsymbol{h}_n + \boldsymbol{b}_n,$$

with

$$F = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{b}_n = \begin{bmatrix} 0 \\ b_1(n) \\ b_2(n), \end{bmatrix}$$

where $\boldsymbol{h}_n = [a_1(n-1), a_1(n), a_2(n)]^t$ is the new state vector for $\boldsymbol{a}_n = [a_1(n), a_2(n)]^t$. Adaptation of the Kalman smoother in its information form to the new state model yields the following algorithm:

- Prediction of $\boldsymbol{x}(n)$ step:

$$\begin{aligned} \boldsymbol{h}_{n|n-1} &= F\boldsymbol{h}_{n-1|n-1} \\ P_{n|n-1} &= FP_{n-1|n-1}F^t + P_b \end{aligned}$$

- Correction step after acquisition of $\boldsymbol{x}(n)$

$$P_{n|n}^{-1} = P_{n|n-1}^{-1} + \sigma_\varepsilon^{-2} \begin{bmatrix} \boldsymbol{O}_M^t \\ X_n \end{bmatrix} [\boldsymbol{O}_M \; X_n^t]$$

$$\boldsymbol{h}_{n|n} = \sigma_\varepsilon^{-2} P_{n|n} (\begin{bmatrix} \boldsymbol{O}_M^t \\ X_n \end{bmatrix} \boldsymbol{x}(n) + P_{n|n-1}^{-1} \boldsymbol{h}_{n|n-1})$$

- Smoothing step:

$$\begin{aligned} A_n &= P_{n|n}F^t P_{n+1|n}^{-1} \\ \boldsymbol{h}_{n|N} &= \boldsymbol{h}_{n|n} + A_n \left(\boldsymbol{h}_{n+1|N} - \boldsymbol{h}_{n+1|n}\right) \\ P_{n|N} &= P_{n|n} + A_n \left(P_{n+1|N} - P_{n+1|n}\right) A_n^t \end{aligned}$$

where $\boldsymbol{O}_N$ is the N length nul vector. As the state vector size increases from 2 to 3, computation is slightly more expensive but it remains reasonable and still allows quasi-instantaneous data analysis.

5. EXPERIMENTAL COMPARISON

Figure 7 and 8 have been computed using the proposed method. More reliable results are obtained in two senses. First each trace shows better regularity since smoothness *prior* is taken into account. Second we have reduced variability from an image to the other. Since all images proceed from the same set of data, this is an good argument in favor of the proposed method.

Via linear regression on the mean frequencies and the spectral width we have obtained 9 values of β. The following table gives the Mean (M) values of these 9 β and the Standard Deviation (SD):

This results gives quantitative evaluation of variability reduction.

Moreover the method accounts for noise (a SNR level of 20dB has been assumed) whereas classical methods do not, which yields overestimated spectral widths and consequently underestimates attenuation coefficients.

	Periodo. method	usual LS AR method	Smooth. *prior* AR method
M	0.73	0.83	1.54
SD	0.16	0.12	0.14
SD %	22	14	8.8

Table 1: Comparison of β estimation performances. Mean values, standard deviations and percents errors, for the three methods.

6. CONCLUSIONS

We have introduced a particular form of time-varying second order AR *prior* model, which accounts for the fact that linear attenuation induces linear variation of the first AR coefficient and a roughly constant second coefficient. In order to cope more efficiently with a vectorial context, we have proposed a costless vectorial information form of Kalman smoother. Finally exact relations between theoretical second-order regression coefficients and Gaussian PSD parameters have been brought out.

Experiments on skin attenuation measurements support the fact that the proposed Bayesian approach provides better repeatability than classic periodogram or autoregressive techniques.

7. ACKNOWLEDGMENTS

Authors wish to thank N. Duclos for editorial assistance.

Partial support for this work was provided by L'ORÉAL under contract number 770566, 1993.

8. REFERENCES

[1] Anderson B.D.O. et Moore J.B., 1979, *Optimal filtering*, Prentice Hall.

[2] Kay S.M., 1988, *Modern spectral estimation, theory and applications*, Prentice Hall.

[3] Kitagawa G. et Gersch W., 1985, "A smoothness priors long AR model method for spectral estimation", *IEEE Trans. on Automatic Control*, **AC-30**, pp. 57-65.

[4] Kitagawa G. et Gersch W., 1985, "A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series", *IEEE Trans. on Automatic Control*, **AC-30**, pp. 48-56.

[5] Kuc R., 1986, "Employing spectral estimation for characterizing diffuse liver desease", *Tissue characterization with ultrasound*, **Vol.2**, pp. 147-167.

[6] Sorenson H.W., 1970, "Least-squares estimation: from Gauss to Kalman", *IEEE Spectrum July*, pp. 63-68.

9. ILLUSTRATIONS

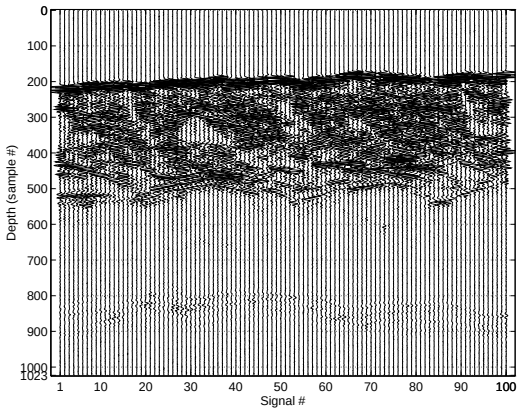

Figure 1: Example of raw data, 400 MHz, 1024 $\times$ 100 samples.

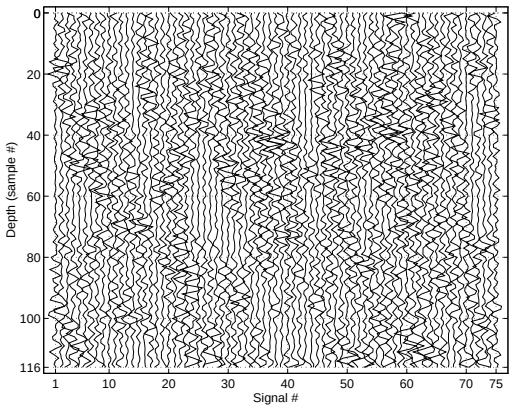

Figure 2: Undersampling, alignment and selection of unsatured areas. 200 MHz, 117 $\times$ 75 samples.

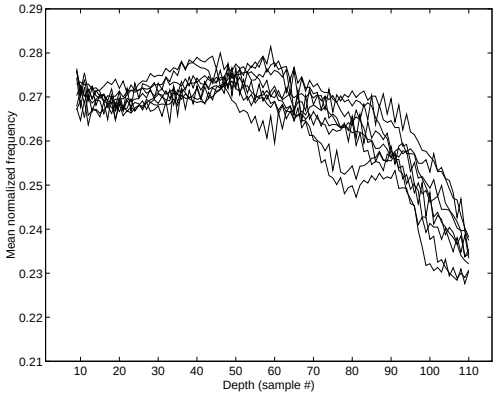

Figure 3: Mean frequencies and periodogram.

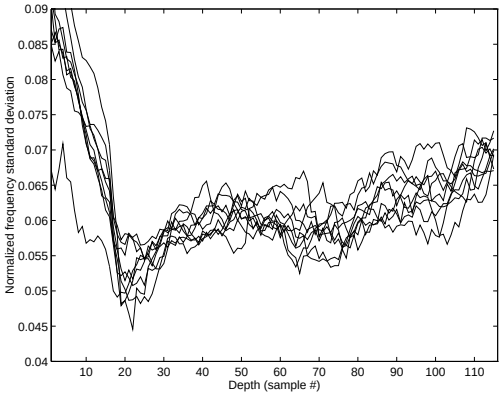

Figure 6: Spectral widths and sliding windows least squares second order autoregressive model.

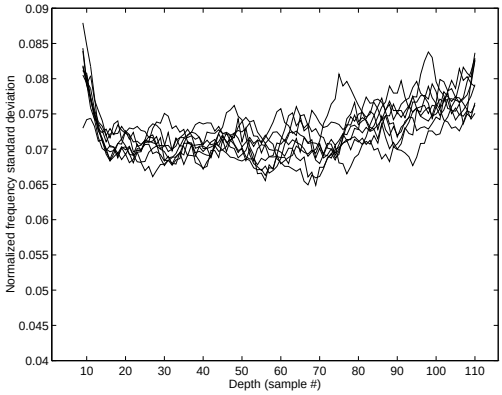

Figure 4: Spectral widths and periodogram.

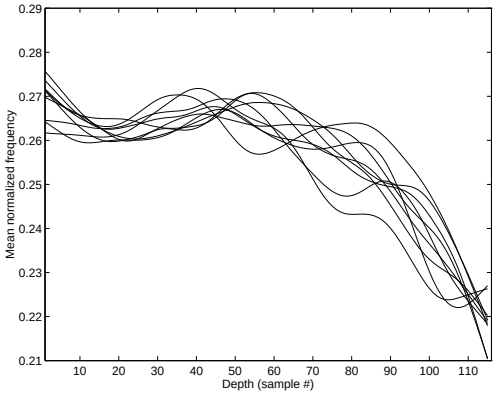

Figure 7: Mean frequencies and smoothness *prior* second order autoregressive model

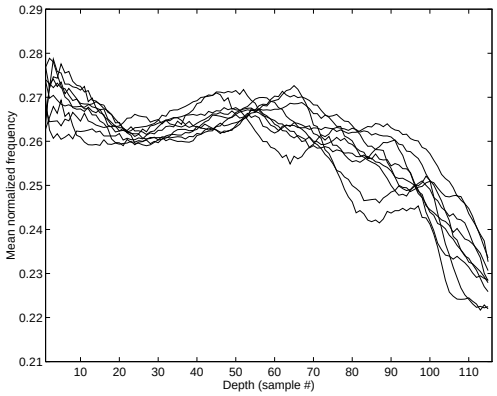

Figure 5: Mean frequencies and sliding windows least squares second order autoregressive model.

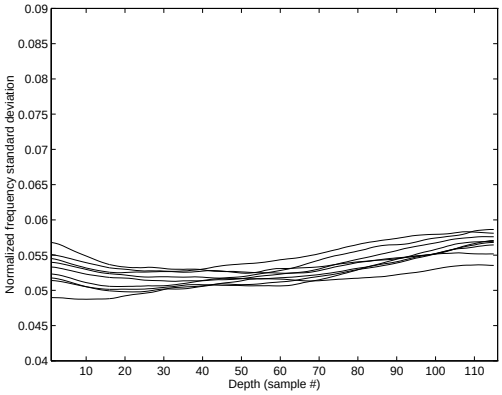

Figure 8: Spectral widths and smoothness *prior* second order autoregressive model.