Geophys. J. Int. (2007) **168,** 1227–1234 doi: 10.1111/j.1365-246X.2006.03257.x

Confidence intervals for earthquake source parameters

G. A. Prieto,[1] D. J. Thomson,[2] F. L. Vernon,[1] P. M. Shearer[1] and R. L. Parker[1]

[1] *Scripps Institution of Oceanography, University of California San Diego, La Jolla,* CA 92093, *USA. E-mail: gprieto@ucsd.edu*
[2] *Mathematics and Statistics Department, Queens University, Kingston, ON, Canada*

Accepted 2006 October 9. Received 2006 August 16; in original form 2006 April 17

SUMMARY

We develop a method to obtain confidence intervals of earthquake source parameters, such as stress drop, seismic moment and corner frequency, from single station measurements. We use the idea of jackknife variance combined with a multitaper spectrum estimation to obtain the confidence regions. The approximately independent spectral estimates provide an ideal case to perform jackknife analysis. Given the particular properties of the problem to solve for source parameters, including high dynamic range, non-negativity, non-linearity, etc., a log transformation is necessary before performing the jackknife analysis. We use a Student's t distribution after transformation to obtain accurate confidence intervals. Even without the distribution assumption, we can generate typical standard deviation confidence regions. We apply this approach to four earthquakes recorded at 1.5 and 2.9 km depth at Cajon Pass, California. It is necessary to propagate the errors from all unknowns to obtain reliable confidence regions. From the example, it is shown that a 50 per cent error in stress drop is not unrealistic, and even higher errors are expected if velocity structure and location errors are present. An extension to multiple station measurement is discussed.

Key words: confidence interval, earthquake source parameters, error, jackknife, stress drop.

1 INTRODUCTION

There is a long-standing controversy on whether stress drop increases with earthquake magnitude or remains constant over a wide range of earthquake sizes (Aki 1967; Archuleta *et al.* 1982; Kanamori *et al.* 1993; Abercrombie 1995; Mayeda & Walter 1996; Ide & Beroza 2001). The behaviour of source parameters including stress drop, corner frequency, radiated seismic energy and apparent stress are of key importance in understanding the physics of earthquakes. However, it is difficult to estimate stress drop reliably from seismograms since it is dependent on the cube of the corner frequency f_c and in turn, f_c is dependent on an accurate account of seismic attenuation, path effects, etc. These factors lead to considerable uncertainty in estimates of stress drop and other source parameters.

Abercrombie (1995) used records from a 2.5 km deep borehole in Cajon Pass and showed that the data supported a constant stress drop, but also an increasing apparent stress with earthquake magnitude. Also from deep borehole data, Prejean & Ellsworth (2001) reported a similar result. A magnitude dependency has been supported by some studies (e.g. Kanamori *et al.* 1993; Mayeda & Walter 1996; Mori *et al.* 2003), while other studies have suggested scale independence (e.g. McGarr 1999; Ide & Beroza 2001; Ide *et al.* 2003), finding no evidence of increasing stress drop or apparent stress with magnitude.

More recently Abercrombie & Rice (2005) revisited some of the Cajon Pass seismograms and using both spectral fitting and Empirical Green's Functions (EGF) concluded that both apparent stress and stress drop may increase with increasing earthquake size, but noted that the uncertainties were still large and scale independence could not be entirely discarded.

So, what are the uncertainties of the estimated source parameters? Error analysis for source parameters has been attempted before (e.g. Archuleta *et al.* 1982; Fletcher *et al.* 1984) but seems to have been neglected more recently. Recently Prieto *et al.* (2006, in press) developed an approach to obtain uncertainties in earthquake source spectral using EGF and applied it to obtain confidence intervals of radiated seismic energy. As pointed out by Tukey (1960):

> 'Probably the greatest ultimate importance, among all types of statistical procedures we now know, belongs to *confidence procedures* which, by making interval estimates, attempt to reach as strong conclusions as are reasonable by pointing out, not single likely values, but rather whole classes (intervals, regions, etc.) of *possible* values, so chosen that there can be high confidence that the 'true' value is *somewhere among them*. Such procedures are clearly quantitative conclusion procedures.'

In this paper, we use the idea of the jackknife variance (Tukey 1958) and follow a similar recipe to the one applied for spectra (Vernon 1989; Thomson & Chave 1991) to construct confidence intervals for earthquake source parameters. This is applied to single station seismograms but can easily be extended to multiple station and spectral ratios and EGF techniques. The confidence intervals are of paramount importance to obtain meaningful scaling relations when different studies, regions, etc. are compared.

We present an example from data recorded at the Cajon Pass Borehole Experiment Phase II with some records also used in

Table 1. Hypocentral parameters for the earthquakes recorded at Cajon Pass used in this study. The relocations by Shearer *et al.* (2005), model SHLK 1.02, are given for SCSN earthquakes.

ID	Year	Month	Day	Hour	Min	Seconds	Latitude (°)	Longitude (°)	Depth (km)	M_L	Dist (km)	CuspID
01	1994	01	01	14	56	42.087	34.2139	−117.4194	14.98	1.1	12.37	3138796
02	1994	01	01	17	47	31.257	34.3863	−117.0185	11.04	3.5	42.86	3138805
03	1994	08	22	21	27	51.366	34.3522	−117.6339	11.03	2.3	14.98	3181641
04	1994	09	21	04	16	51.762	34.2316	−117.4623	13.63	2.3	9.32	3185485

M_L is the local magnitude given by the SHLK catalogue, Dist is the hypocentral distance from the borehole station, CuspID is the SCSN ID for the earthquakes.

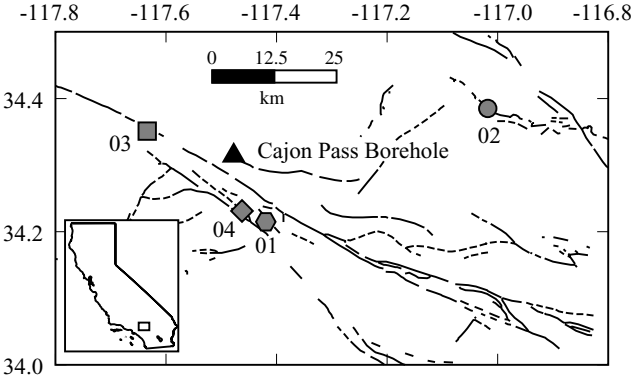

Figure 1. Map showing the Cajon Pass Borehole location and relocated earthquakes considered in this study. Relocations from Shearer *et al.* (2005) model SHLK 1.02.

Abercrombie (1997) and show the resultant confidence intervals for stress drop and other source parameters for four small earthquakes (Table 1) that were also recorded by the Southern California Seismic Network (SCSN). Fig. 1 shows a map with the relocated earthquakes and the borehole location.

2 THE JACKKNIFE METHOD

The *jackknife* was first introduced by Quenouille (1949) and then named and extended by Tukey (1958) to estimate variances. It is one of many resampling methods used for statistical inference. One of the great advantages of the jackknife is that one does not need to know the statistical distribution of the parameter in question and that it works on complicated processes reliably (a detailed proof is given in Reeds 1978). In this paper, we will use the so-called delete-one jackknife, which we will refer to simply as the jackknife. A good review can be found in Miller (1974) and Efron (1982).

Assume $X_1, X_2, \ldots, X_K$ are K independent random observations taken from an unknown probability distribution characterized by a parameter θ which is to be estimated. The estimate of θ using all observations is:

$$\hat{\theta} = \hat{\theta}[X_1, X_2, \ldots, X_K] \tag{1}$$

Let

$$\hat{\theta}_{(i)} = \hat{\theta}[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_K] \tag{2}$$

be the delete-one estimate of θ, where the ith observation X_i is not used to estimate $\hat{\theta}_{(i)}$. The data are thus subdivided in K groups of size $(K-1)$ by deleting each entry in turn.

An important application of the jackknife was suggested by Tukey (1958), and is the jackknife estimate of the variance of $\hat{\theta}$

$$\mathrm{var}\{\hat{\theta}\} = \frac{K-1}{K}\sum_{i=1}^{K}\left[\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)}\right]^2, \tag{3}$$

where

$$\hat{\theta}_{(\cdot)} = \frac{1}{K}\sum_{i=1}^{K}\hat{\theta}_{(i)} \tag{4}$$

is the mean of the delete-one estimates (2). Although it has been proposed (Wu, 1986) that deleting an *arbitrary* number of observations might have better convergence properties, we use throughout the paper the delete-one jackknife because of its simplicity, efficiency and independence of an arbitrary chosen subdivision for the groups.

As suggested by Miller (1974) and applied in spectrum estimation (Vernon 1989; Thomson & Chave 1991) it is sometimes necessary to use a transformation that stabilizes the variance, especially when the statistic being investigated is bounded or its distribution is strongly non-Gaussian. This can be important when estimating errors in stress drop $\Delta\tau$, seismic moment M_0, and corner frequency f_c, all with a range $[0, \infty)$.

2.1 Jackknife in regression problems

Consider the regression problem for a basic model

$$\mathbf{Y} = \mathbf{A}\beta + \mathbf{e}, \tag{5}$$

where $\mathbf{Y}$, $\mathbf{e}$ are m sized vectors of the data and the errors, $\mathbf{A}$ is a $m \times p$ matrix from the model, and β is a p size vector of the parameters we wish to find.

Miller (1974) examined the traditional jackknife approach by deleting rows of both $\mathbf{Y}$ and $\mathbf{A}$ simultaneously and showed the asymptotic normality of the jackknife solution vector and its variance under general conditions. The delete-one estimate is given by solving

$$\mathbf{Y}_{(i)} = \mathbf{A}_{(i)}\,\hat{\beta}_{(i)}, \tag{6}$$

where $\mathbf{Y}_{(i)}$ and $\mathbf{A}_{(i)}$ have the ith row removed. As will be clear in the subsequent sections, the problem to solve for source parameters is non-linear and the model (5) is not appropriate. Instead we have

$$y_i = g_i(\beta) + e_I, \tag{7}$$

where g_i is a non-linear smooth function of the parameters in β (Fox *et al.* 1980; Wu 1986). In an analogous way, we want to obtain the delete-one estimates $\hat{\beta}_{(i)}$ that satisfy (7) by means of one of many non-linear parameter estimation techniques (non-linear least-squares, grid search, etc).

3 MULTITAPER SPECTRUM ESTIMATES

The multitaper spectrum algorithm was introduced by Thomson (1982) and has been widely used in the geophysical community (e.g. Park *et al.* 1987; Vernon 1989; Chappellaz *et al.* 1990;

Abercrombie 1995; Lees & Park 1995). The method takes advantage of a family of orthogonal tapers which are resistant to spectral leakage.

Given a time-series $x(t)$ with N contiguous data samples and assuming unit sampling, we multiply the time-series by a sequence $a(t)$ called a taper and apply a DFT

$$y(f) = \sum_{t=0}^{N-1} x(t)a(t)e^{-2\pi i f t} \qquad \text{with} \qquad \sum_{t=0}^{N-1} |a(t)|^2 = 1 \tag{8}$$

to obtain a direct estimate of the true spectrum $S(f)$ of the signal

$$\hat{S}(f) = |y(f)|^2. \tag{9}$$

The question is then what taper to use? Is there a reason to prefer one taper over the other?

Spectral leakage is the bias introduced by energy *leaking* from frequencies different from the frequency f for $S(f)$. Now the question becomes: what taper $a(t)$ has the greatest concentration of energy in its Fourier transform? Spectral properties of the taper can be studied from its DFT

$$A(f) = \sum_{t=0}^{N-1} a(t)e^{-2\pi i f t}. \tag{10}$$

The function $|A(f)|$ for conventional tapers has a broad main lobe and a succession of smaller sidelobes. The larger the sidelobes, the more spectral leakage is biasing $\hat{S}(f)$.

We can express the estimate in eq. (9) as a convolution of the taper transform (10) and the true spectrum $S(f)$ (see Thomson 1982; Park *et al.* 1987, for derivation):

$$\hat{S}(f) = \int_{-1/2}^{1/2} |A(f - f')|^2 S(f')\, df'. \tag{11}$$

The interpretation of this equation is as a convolution describing the *smearing* of the true spectrum as a consequence of the discrete sampling. A good taper will have a spectral window with low amplitudes whenever $|f - f'|$ gets large and large amplitudes whenever $|f - f'|$ is small.

Slepian (1978) suggested choosing a frequency W, where $0 < |W| \leq 1/2$ (unit sampling) and maximizing the fraction of energy of A at frequencies from $(-W, W)$. In mathematical form this is equivalent to:

$$\lambda(N, W) = \frac{\displaystyle\int_{-W}^{W} |A(f)|^2\, df}{\displaystyle\int_{-1/2}^{1/2} |A(f)|^2\, df}. \tag{12}$$

Since no finite time-series can be completely band-limited, $\lambda < 1$. The spectral leakage comes from the sidelobes of $A(f)$ convolved with the spectrum outside the band $(f - W, f + W)$. One can think of $\lambda(N, W)$ as the amount of spectral energy at $\hat{S}(f)$ that comes from $(f - W, f + W)$ and $1 - \lambda$ as the amount that comes from outside the band or as the *bias* from outside the band.

We wish to maximize the value of λ by choosing $A(f)$ appropriately. Substitute (10) into (12) and represent $a(t)$ by an N-vector of coefficients **a**; taking the gradient of λ with respect to **a** and setting to zero leads to the matrix eigenvalue problem:

$$\mathbf{D} \cdot \mathbf{a} - \lambda \mathbf{a} = 0, \tag{13}$$

where **D** is a symmetric matrix

$$D(t, t') = \frac{\sin 2\pi W(t - t')}{\pi(t - t')} \tag{14}$$

with eigenvalues $1 > \lambda_0 > \lambda_1 > \ldots > \lambda_{N-1} > 0$ and associated eigenvectors $v_k(t;N, W)$ called the Slepian sequences (Slepian 1978). From now on we will drop the explicit dependence on N and W.

The eigenvector with the largest eigenvalue is the best possible taper for the suppression of spectral leakage, and in practice we find λ_0 is usually extraordinarily close to one. However, in fact it can be proved that the first $2NW - 1$ eigenvalues are also very close to one, leading to a whole family of excellent tapers. The multitaper method exploits this fact by using all of these tapers rather than merely the first one. Because the eigentapers are orthogonal (both in time and frequency domains), the estimates based on them are statistically independent of each other and can, therefore, be combined together to yield a more reliable overall estimate as we will explain.

In practice, we choose a bandwidth W over which the spectrum is to be smoothed, thus fixing NW, which is called the time-bandwidth product of the system under study. For practical problems we always choose $NW > 1$, because we cannot expect to obtain good concentration into a frequency band narrower than $f_R = 1/N$, the Rayleigh resolution.

Fig. 2 shows the Slepian sequences and their Fourier transforms with corresponding eigenvalues for a time-series with $N = 100$ samples, $NW = 4$, and $W = 0.04$ for unit sampling. The horizontal axis is shown in Rayleigh units, basically equivalent to the

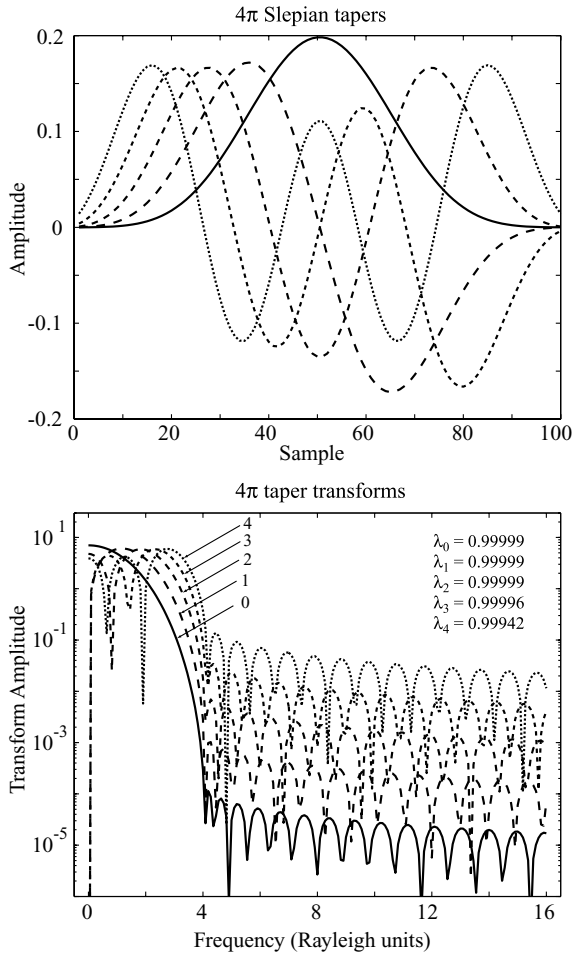

Figure 2. The five lowest order 4π Slepian tapers (top panel) and corresponding Fourier transform amplitudes (lower panel). Solid lines correspond to the zero-order sequence $v_0(t)$, higher-order tapers are plotted with dashed lines. We have used $N = 100$ and $NW = 4$. Estimated eigenvalues are also provided.

frequency sampling. Fig. 2 suggests that $NW = 4$ is equivalent to saying that the smoothing will take place over NW Rayleigh bins around the frequency of interest. Note that here we have assumed unit sampling—if that is not the case, then the time-bandwidth product is actually $\Delta tNW = 4$, where Δt is the sampling rate, in order to maintain the proper units of W.

Turning back to the spectral estimation problem, given a particular bandwidth W, we compute DFTs of the tapered data $y_k(f)$, called the raw eigencoefficients,

$$y_k(f) = \sum_{t=0}^{N-1} x(t) v_k(t) e^{-2\pi i f t}. \tag{15}$$

We generally use $k = 1, \ldots, K$, where $K = 2NW - 1$. As expressed above, the corresponding eigenvalues are $\lambda_k \approx 1$ with good leakage properties.

As suggested by Thomson (1982) we use the adaptive weighting procedure

$$\hat{x}_k(f) = d_k(f) y_k(f) \tag{16}$$

and the corresponding adaptive spectral estimate

$$\hat{S}(f) = \frac{\sum_{k=0}^{K-1} |x_k(f)|^2}{\sum_{k=0}^{K-1} |d_k(f)|^2}, \tag{17}$$

where the weights $d_k(f)$ are chosen to reduce bias from spectral leakage. The frequency dependent weights are useful in the analysis of high dynamic range spectral processes. The weights work as follows. At frequencies where the spectrum is reasonably flat, the weights $d_k(f) \approx 1$, thus reducing the variance of the spectral estimate by averaging over all the eigencoefficients y_k. At frequencies where the spectrum has a large dynamic range the higher-order eigencoefficients might be biased and the weights reduce the contributions from these eigencoefficients.

The optimal weights $d_k(f)$ can be found by minimizing the misfit between the estimated spectrum and the true spectrum $S(f)$. The approximate optimum weights are

$$d_k(f) \approx \frac{\sqrt{\lambda_k} S(f)}{\lambda_k S(f) + (1 - \lambda_k)\sigma^2}, \tag{18}$$

where σ^2 represents the variance of the time-series. The term $(1 - \lambda_k)\sigma^2$ represents an approximation to the bias from spectral leakage. Since we do not know the true spectrum, we replace $S(f)$ by an estimate $\hat{S}(f)$.

We find the weights and estimated spectrum $\hat{S}(f)$ by iteration. As an initial estimate of $S(f)$ we take the arithmetic average of the first two squared eigencoefficients $|y_0(f)|^2$ and $|y_1(f)|^2$ and substitute in (18) to obtain estimates of $d_k(f)$. The weights are then used in (16–17) to obtain a new spectral estimate $\hat{S}(f)$ and this process is repeated. Convergence is rapid and only a few cycles are necessary. Note that both the tapers and weights are normalized in order to keep the spectrum in physical units.

The kth eigenspectrum is

$$\hat{S}_k(f) = |\hat{x}_k(f)|^2. \tag{19}$$

For the jackknife approach, we will use the $\hat{S}_k$ as the K independent estimates of the spectrum. At each frequency f the multitaper estimate of the log spectrum is given by

$$\ln \hat{S} = \ln \left[\frac{1}{K} \sum_{k=1}^{K} \hat{S}_k \right] \tag{20}$$

and we also define the delete-one spectrum

$$\ln \hat{S}_{\overline{\imath}} = \ln \left[\frac{1}{K-1} \sum_{k=1, k\neq i}^{K} \hat{S}_k \right]. \tag{21}$$

The logarithmic transformation of the spectrum is suggested in Thomson & Chave (1991), providing a more symmetric distribution than the standard χ^2 for spectral estimates.

4 SOURCE PARAMETER JACKKNIFE

A general source model of the displacement spectra of both P and S waves (e.g. Abercrombie 1995) is:

$$u(f) = \frac{\Omega_0 \, e^{-(\pi f t/Q)}}{\left[1 + (f/f_c)^{n\gamma}\right]^{1/\gamma}}, \tag{22}$$

where Ω_0 is the long period amplitude, f is the frequency, f_c is the corner frequency, n the high frequency fall-off rate, γ is a constant, t is the traveltime, and Q a frequency independent quality factor. Modified versions of spectral shapes proposed by Brune (1970) and Boatwright (1980) can be obtained by changing γ. Based on previous studies of data from the Cajon Pass Borehole (Abercrombie 1995, 1997) a value $\gamma = 2$ and a variable fall-off n fits the spectra reasonably well. In this paper, we will use

$$u(f) = \frac{\Omega_0 \, e^{-(\pi f t/Q)}}{\left[1 + (f/f_c)^{2n}\right]^{0.5}}. \tag{23}$$

Following Ide *et al.* (2003) take the logarithm

$$\ln u(f) = g(f; \beta) \tag{24}$$

$$= \ln \Omega_0 - 0.5 \ln \left(1 + (f/f_c)^{2n}\right) - \frac{\pi f t}{Q}, \tag{25}$$

where β is a vector of three components given by the parameters we are searching for, namely Ω_0, f_c, Q. The function $g(f, \beta)$ is clearly non-linear and some kind of non-linear inversion is necessary. It is not the aim of this paper to discuss the difficulties encountered in solving this problem, and we suggest reading Bard (1974) on non-linear parameter estimation, as applied to source physics (see Abercrombie 1995; Ide *et al.* 2003, and references therein).

The estimate $\hat{\beta}$ of β using all observations (as in eq. 1) is given by the solution of (24) using the multitaper spectrum estimate (20). The delete-one $\hat{\beta}_{\overline{\imath}}$ parameter instead uses the delete-one spectrum (21). The result is K delete-one estimates of the long period amplitude, corner frequency and quality factor, denoted, respectively, $\Omega_{0,\overline{\imath}}$, $f_{c,\overline{\imath}}$, $Q_{\overline{\imath}}$.

4.1 Transformations

The use of transformations before performing the jackknife is in some cases necessary. In terms of source parameters a logarithmic transformation should provide more stable estimates of variance. Some reasons for this are as follows:

(i) The three seismic parameters are non-negative. If a simple Gaussian distribution is assumed, the tails give a non-zero probability that a parameter is negative, which is not physical. The jackknife does not constrain variances to be positive.

(ii) As suggested by (Archuleta *et al.* 1982), if no transformation is performed, the arithmetical average will be biased to the larger values, while taking a transformed parameter gives equal weight to all independent estimates.

(iii) A closer to normal distribution of the errors is achieved by such a transformation.

In this respect, we will perform the jackknife on $\theta = \ln \Omega_0$ rather than Ω_0.

4.2 Confidence intervals

As a result of Section 4.1 we obtain an estimate $\hat{\theta} = \ln \hat{\beta}$ and the variance of the transformed variable

$$\tilde{\sigma}^2 = var\{\ln \hat{\beta}\} \tag{26}$$

$$= \frac{K-1}{K} \sum_{i=1}^{K} [\ln \hat{\beta}_{\widehat{i}} - \ln \hat{\beta}_{\widehat{\cdot}}]^2, \tag{27}$$

where $\hat{\beta}_{\widehat{i}}$ is any one of $\hat{\Omega}_{0,\widehat{i}}$, $\hat{f_c}_{,\widehat{i}}$, and $\hat{Q}_{\widehat{i}}$.

Tukey (1958) suggested that $(\ln \hat{\beta}_{\widehat{i}} - \ln \hat{\beta}_{\widehat{\cdot}})/\tilde{\sigma}$ is nearly distributed as Student's t with $K-1$ degrees of freedom for small samples. Hinkley (1977) on the other hand stated that if the data have strongly non-normal distributions, the Student's t approximation can lead to substantial errors. However, if the transformation performed leads to more nearly normal distributions, the approximation is reasonably accurate (Davison & Hinkley 1997). Note that this distribution is very close to the Gaussian distribution and for 30 or more degrees of freedom they are almost indistinguishable. With this in mind, the double-sided $1-\alpha$ confidence interval of the long period amplitude is

$$\hat{\Omega}_0\, e^{-t_{K-1}(1-\alpha/2)\tilde{\sigma}} < \hat{\Omega}_0 \leq \hat{\Omega}_0\, e^{t_{K-1}(1-\alpha/2)\tilde{\sigma}} \tag{28}$$

and similar for $\hat{f_c}$ and $\hat{Q}$. If the Student's t approximation is not appropriate for the particular data, one can always simply plot the $\pm\tilde{\sigma}$ bounds by adjusting (28). Note that because of the transformation, the lower limit is never negative.

4.3 Seismic moment, source radius and stress drop

Other important source parameters estimated from the spectrum are the seismic moment (M_0), the source radius (r) and the stress drop ($\Delta\tau$) and are often calculated assuming a circular fault (Brune 1970; Madariaga 1976), in which case

$$M_0 = \frac{4\pi\rho c^3 R\hat{\Omega}_0}{U_{\theta\phi}}, \tag{29}$$

$$r = \frac{k\beta}{\hat{f_c}}, \tag{30}$$

where a constant rupture velocity is assumed. From the mean estimates of the previous two equations,

$$\Delta\tau = \frac{7M_0}{16r^3} \tag{31}$$

where ρ, c, R, $U_{\theta\phi}$, β are density, wave velocity, hypocentral distance, the mean radiation pattern (0.52 and 0.63 for P and S waves) and the shear wave velocity at the source. k is 0.32 and 0.21 for P and S waves, respectively, assuming the rupture velocity is 0.9β (Madariaga 1976).

We will assume that the parameters not associated with the source (shear wave speed, density, etc.) are known exactly, that is, do not contribute to the uncertainties of seismic moment, source radius and stress drop. We will use the idea of propagation of errors (Taylor 1997) to obtain confidence limits of these parameters.

We perform the propagation of errors in the log domain, since it is where we have variance estimates of $\ln \Omega_0$ and $\ln f_c$, denoted, respectively, $\sigma^2_{\Omega_0}$ and $\sigma^2_{f_c}$. The idea is to obtain the variance of seismic moment $\sigma^2_{M_0} = var\{\ln M_0\}$, source radius $\sigma^2_r = var\{\ln r\}$, and stress drop $\sigma^2_{\Delta\tau} = var\{\ln \Delta\tau\}$. After this, eq. (28) can be used to obtain confidence intervals. Some rules of propagation of errors are shown in Appendix A.

The relation of errors between the source and spectral parameters are

$$\sigma^2_{M_0} = \sigma^2_{\Omega_0} \tag{32}$$

$$\sigma^2_r = \sigma^2_{f_c} \tag{33}$$

and a more complicated relation is obtained for the stress drop, since it depends on two variables

$$\sigma^2_{\Delta\tau} = \sigma^2_{M_0} + 9\,\sigma^2_r = \sigma^2_{\Omega_0} + 9\,\sigma^2_{f_c}, \tag{34}$$

where it is assumed that the covariance of Ω_0 and f_c is negligible. This relation was used by Fletcher *et al.* (1984) to estimate uncertainties of stress drop using a multiplicative error. Again, the bounds (either $\pm\sigma$ or confidence intervals using the Student's t approximation) can be transformed back to the linear domain using (28).

5 APPLICATION TO CAJON PASS DATA

Because attenuation can also cause fall-off at high-frequencies it is important to correct observed spectra for Q effects. Full consideration of these effects is beyond our focus here, therefore, as a demonstration of the jackknife procedure to obtain variance and confidence intervals we choose seismograms recorded at two different depths (1.5 and 2.9 km) at the Cajon Pass Borehole, where attenuation effects are relatively small and have been previously modelled by Abercrombie (1995). The seismometers that recorded this data set are 10-Hz L-15LA high temperature geophones, with sample rates of 1000 samples s^{-1}.

Fig. 3 shows displacement seismograms recorded at the 2.9 km depth sensor for a M_L3.5 earthquake 43 km away (ID 02 in Table 1). The spectrum is computed for a 1 s window, starting 0.15 s before the P pick at the station, similar to windows used in previous work (e.g. Abercrombie 1995; Prieto *et al.* 2004; Abercrombie & Rice 2005) for small earthquakes.

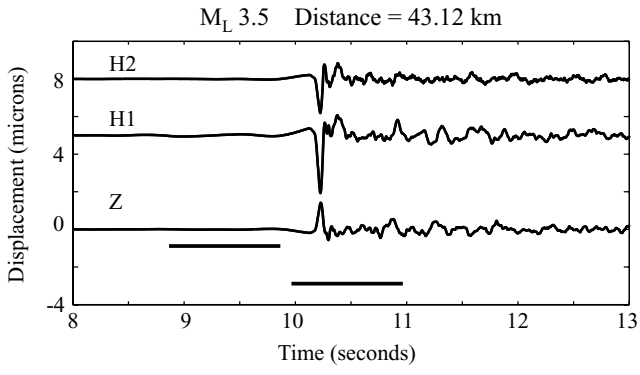

Figure 3. Example seismogram of the largest event used in this study, M_L3.5 recorded at the deepest borehole sensor 2.9 km. Seismograms have been corrected for instrument response and are flat to displacement between 2 and 300 Hz. Horizontal bars show the choice of noise and P-wave window.

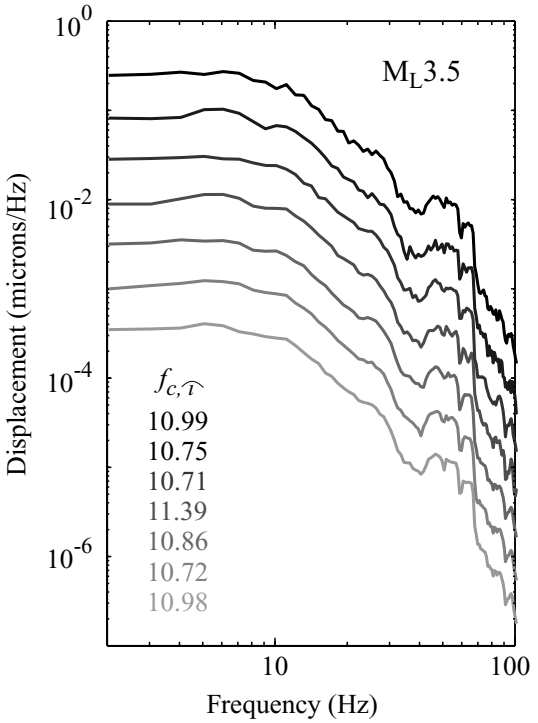

Figure 4. The delete-one spectrum for the M_L 3.5 earthquake, from signal in Fig. 3 at the 2.9 km sensor. The spectra have been shifted for comparison purposes. All spectra show a similar behaviour and slight differences are seen. For each spectrum, the delete-one corner frequency estimates are listed.

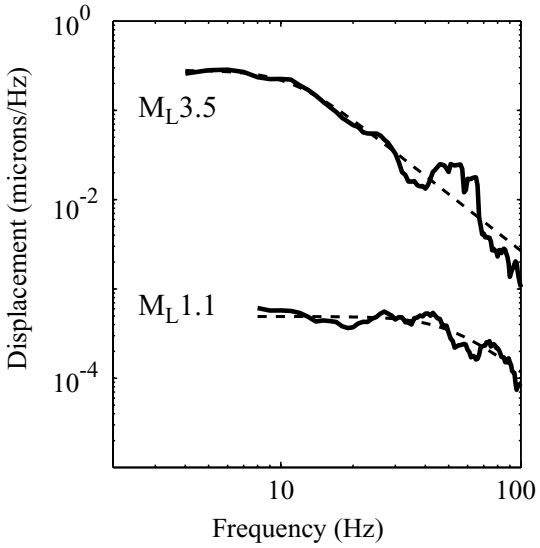

Figure 5. Spectral modeling of *P* waves for two different sized earthquakes. The bandwidth used to fit the source model varies depending on event size. An attenuation correction was previously performed, using $Q = 1000$ as suggested by Abercrombie (1995).

For each of the three components (Z, H1 and H2) we estimate the amplitude spectrum using a time-bandwidth product $NW = 4$ and work with $K = 7$ tapers. This means we also compute for each component 7 delete-one spectra as in eq. (21). The final amplitude spectrum is then computed by vector summation of the three component spectra, and similarly for the delete-one spectra.

Fig. 4 shows the complete set of delete-one spectra for the M_L3.5 earthquake recorded at the 2.9 km sensor. The spectra have been shifted for comparison. In general the spectral shapes are very similar and only slight differences at very low frequencies and roughness at higher frequencies are visible. Note that in Fig. 4 the delete-one spectra are plotted, which are not independent estimates. Only $\hat{S}_k(f)$ are treated as independent, given the orthogonality properties of the Slepian tapers.

Here, we have a good example of the properties of the multitaper algorithm. The data from the Cajon Pass have a very strong 60 Hz signal. In this work, we have $N = 1000$ samples, $dt = 0.001$, we chose $NW = 4$, the band $W = 4$ Hz and we use $K = 7$ tapers. This means that the 60 Hz peak will be smoothed over the band between 56 and 64 Hz. In Fig. 4, we can see a very sharp discontinuity at 64 Hz due to the fact that outside the band, very little energy is leaked to frequencies $f > 64$ Hz.

Following Abercrombie (1995) and Abercrombie & Rice (2005) we use $Q = 1000$ and correct the spectra before performing spectral fitting. We vary the fitting bandwidth to obtain optimal fits, but the same bandwidth is used for all delete-one spectra and the average spectra. An example fit to the model (25) is shown in Fig. 5 for the largest and smallest earthquake in this study. Note, however that the 1.5 km sensor data are used for the small earthquake, due to complicated resonances present at the 2.9 km sensor spectra that may have affected the results.

As in Abercrombie & Rice (2005) we use density $\rho = 2700\,\mathrm{kg\,m^{-3}}$, $\alpha = 6000\,\mathrm{m\,s^{-1}}$ and $\beta = \alpha/\sqrt{3}$ and using eqs (29–31) we estimate the source parameters M_0, f_c, r, $\Delta\tau$ from the average spectra and the jackknife parameters (e.g. $M_{0,\widehat{\tau}}$) from the delete-one spectra to get the jackknife variance and confidence intervals. Table 2 shows the source parameters and 5–95 per cent confidence limits. Fig. 6 shows plots of seismic moment and corner frequencies and seismic moment and stress drop for the data used in this paper.

It is important to note the assumptions and unknowns in the calculations. For example, we have assumed that the wave speeds (α, β) are known exactly. If there are errors (and certainly there are) associated with the wave speed, errors will propagate to M_0, r and subsequently to the stress drop. Assuming a 5 per cent error in the *S*-wave speed, thus affecting the radius uncertainties (and rupture speed), the confidence region for the stress drop for the M_L3.5 earthquake recorded at 2.9 km sensor would be (47, 108), a change of about 10 per cent. Other sources of errors for this example include the attenuation correction, the constant Q assumption used, earthquake location errors, radiation pattern and directivity, etc. Perhaps most importantly, we assume the validity of the source model; our method provides an estimate of the errors in $\Delta\tau$ with respect to random fluctuations in the data but is not a test of the validity of the model itself.

5.1 Extension to multiple stations

A generalization of the jackknife to multiple stations is desirable. A major source of uncertainty would be directivity, because as expected from directivity, the pulse width of the source time functions is narrower in the direction of rupture and broader in the opposite direction, also changing the corner frequency (e.g. McGuire 2004).

One approach is to treat the different station estimates of corner frequencies f_c and seismic moment M_0 as independent, and, after suitable transformations, compute confidence intervals as explained in Sections 2 and 4. If the earthquake source spectrum is simple (no directivity effects, radiation pattern correctly accounted for, etc.) all stations would return a similar estimate, thus having small uncer-

Table 2. Source parameters and confidence intervals obtained by spectral fitting and jackknife analysis.

ID	M_L	Sensor	M_0 (Nm)	M_0 conf	f_c (Hz)	f_c conf	n	r(m)	$\Delta\tau$ (MPa)	$\Delta\tau$ conf
01	1.1	1.5	0.96 e + 11	(0.78 1.17) e + 11	43.80	(33.29 57.63)	1.72	25	2.60	(1.11 6.07)
02	3.5	1.5	1.73 e + 14	(1.11 2.68) e + 14	6.92	(4.77 10.03)	1.72	160	18.34	(5.75 58.47)
02	3.5	2.9	1.72 e + 14	(1.35 2.19) e + 14	10.91	(9.91 12.00)	2.09	101	71.55	(52.78 96.98)
03	2.3	2.9	1.53 e + 12	(1.31 1.79) e + 12	20.74	(17.46 24.63)	1.77	53	4.55	(2.72 7.64)
04	2.3	2.9	1.36 e + 12	(1.00 1.85) e + 12	21.66	(18.74 25.02)	2.03	51	4.66	(2.90 7.50)

M_0 conf, f_c conf and $\Delta\tau$ conf are the 5–95 per cent confidence intervals for the seismic moment, corner frequency and stress drop, respectively. n is the high-frequency fall-off for the P waves. $Q = 1000$ is assumed. The column named Sensor represents the depth in km of the sensor recording the earthquake.

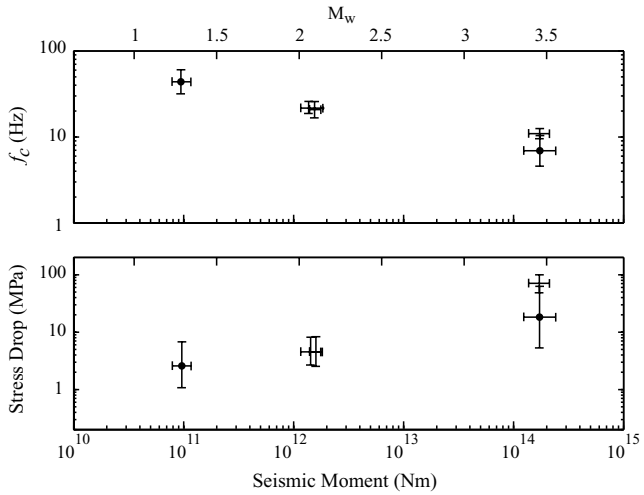

Figure 6. Source parameters determined in this study and 5–95 per cent confidence intervals using a Student's *t* approximation. Symbols with dot at the centre have parameters determined at the 1.5 km deep sensor. Note that the confidence regions may vary from event to event.

tainties, while if there is strong directivity, certain stations will have considerably different corner frequencies, increasing uncertainties. A recipe for a multiple station jackknife of f_c is as follows:

(1) Compute f_c for every station that recorded the earthquake.
(2) Use the log transformation.
(3) Compute a mean $\ln \hat{f_c}$ (eq. 1).
(4) Compute delete-one $\ln \hat{f}_{c,\widehat{i}}$ (eq. 2).
(5) Compute variance $var\{\ln f_c\} = \sigma^2_{f_c}$ (eq. 3).
(6) Obtain confidence intervals (eq. 28).

A similar approach could be used for other parameters such as long period amplitude Ω_0, Q, etc. Propagation of errors (eqs 32–34) is necessary to obtain confidence intervals on seismic moment, source radii and stress drop.

6 CONCLUSIONS

In estimating source parameters from the seismic spectrum, it is important not only to obtain a measure of the source parameters but also to obtain a measure of the uncertainties, by means of confidence intervals. The jackknife is a reliable way of estimating the variance of source parameters, and, given suitable transformations, confidence intervals. It should be easy to extend this approach to multiple station studies, where other sources of error include the radiation pattern and directivity effects which might generate different corner frequencies and radiated energy.

We calculate source parameters and confidence intervals for four small earthquakes as an example of the use of the jackknife approach. The error analysis is necessary if the data are to be used to constrain rupture models (Abercrombie & Rice 2005), examine scaling relations and the size dependence of earthquake parameters, in order to conclude, within a reasonable reliability, something about the physics of earthquakes.

From Fig. 6 there appears to be a slight increase of stress drop with earthquake magnitude, which, unless the errors are kept small, would pass unnoticed. Note also that the larger uncertainties are associated with the 1.5 km sensor, compared to at 2.9 km. If some of the assumptions such as radiation pattern and earthquake location contribute to the errors, the stress drop scaling would be less apparent, suggesting the need to find ways of reducing uncertainties.

The M3.5 earthquake (ID02) was recorded at two different depths and the corner frequency confidence regions barely overlap. This variation is likely explained by other sources of error such as near site effects at the shallower station. This also shows that even stations close to each other may have very different estimates of source parameters and uncertainties are needed to address the significance of these estimates.

ACKNOWLEDGMENTS

We thank R.E. Abercrombie for help on using the Cajon Pass data, Ralph Archuleta and one anonymous reviewer for their constructive comments. Funding for this research was provided by NSF Grant number EAR0417983. This research was also supported by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 979. We also thank IRIS and the staff at the Data Management Center for access to the Cajon Pass data.

REFERENCES

Abercrombie, R.E., 1995. Earthquake source scaling relationships from −1 to 5 M_L using seismogram recorded at 2.5-km depth, *J. geophys. Res.,* **100,** 24 015–24 036.

Abercrombie, R.E., 1997. Near-Surface attenuation and site effects from comparison of surface and deep borehole recordings, *Bull. seism. Soc. Am.,* **87**(3), 731–744.

Abercrombie, R.E. & Rice, J.R., 2005. Can observations of earthquake scaling constrain slip weakening?, *Geophys. J. Int.,* **162,** 406–424.

Aki, K., 1967. Scaling law of seismic spectrum, *J. geophys. Res.,* **72,** 1217–1231.

Archuleta, R.J., Cranswick, E., Mueller, C. & Spudich, P., 1982. Source parameters of the 1980 Mammoth Lakes, California, Earthquake sequence, *J. geophys. Res.,* **87**(B6), 4595–4607.

Bard, Y., 1974. *Nonlinear Parameter Estimation,* Academic Press. New York, p. 341.

Boatwright, J., 1980. A spectral theory for circular seismic sources: simple estimates of source dimension, dynamic stress drop and radiated energy, *Bull. seism. Soc. Am.,* **70,** 1–27.

Brune, J.N., 1970. Tectonic stress and seismic shear waves from earthquakes, *J. geophys. Res.,* **75,** 4997–5009.

Chappellaz, J., Bernola, J.M., Raynaud, D., Korotkevich, Y.S. & Lorius, C., 1990. Ice-core record of atmospheric methane over the past 160 000 years, *Nature,* **345,** 127–131.

Davison, A.C. & Hinkley, D.V., 1997. *Bootstrap Methods and their Application,* Cambridge University Press, New York.

Efron, B., 1982. *The Jackknife, the Bootstrap, and Other Resampling Plans,* SIAM, Philadelphia.

Fletcher, J., Boatwright, J., Haar, L., Hanks, T. & McGarr, A., 1984. Source parameters for aftershocks of the Oroville California Earthquake, *Bull. seism. Soc. Am.,* **74**(4), 1101–1123.

Fox, T., Hinkley, D.V. & Larntz, K., 1980. Jackknifing in Nonlinear regresion, *Technometrics,* **22,** 29–33.

Hinkley, D.V., 1977. Jackknife Confidence Limits Using Student *t* Approximations, *Biometrika,* **64,** 21–28.

Ide, S. & Beroza, G.C., 2001. Does apparent stress vary with earthquake size?, *Geophys. Res. Lett.,* **28,** 3349–3352.

Ide, S., Beroza, G.C., Prejean, S.G. & Ellworth, W.L., 2003. Apparent break in earthquake scaling due to path and site effects on deep borehole recordings, *J. geophys. Res.,* **108**(B5), 2271, doi:10.1029/2001 JB001617.

Kanamori, H., Mori, J., Hauksson, E., Heaton, T.H., Hutton, L.K. & Jones, L., 1993. Determination of earthquake energy release and M_L using TERRAscope, *Bull. seism. Soc. Am.,* **83,** 330–346.

Lees, J. & Park, J., 1995. Multiple-taper spectral analysis: a stand-alone C-subroutine, *Comput. Geosci.,* **21**(2), 199–236.

Madariaga, R., 1976. Dynamics of an expanding circular fault, *Bull. seism. Soc. Am.,* **66,** 639–666.

Mayeda, K. & Walter, W.R., 1996. Moment, energy, stress drop, and source spectra of western United States earthquakes from regional coda envelopes, *J. geophys. Res.,* **101,** 11 195–11 208.

McGarr, A., 1999. On relating apparent stress to the stress causing earthquake fault slip, *J. geophys. Res.,* **104,** 3001–3003.

McGuire, J.J., 2004. Estimating Finite Source Properties of Small Earthquake Ruptures, *Bull. seism. Soc. Am.,* **94,** 377–393.

Miller, R.G., 1974. The jackknife—a review, *Biometrika,* **61,** 1–15.

Mori, J., Abercrombie, R.E. & Kanamori, H., 2003. Stress drops and radiated energies of aftershocks of the 1994 Northridge, California, earthquake, *J. geophys. Res.,* **108**(B11), 2545, doi:10.1029/2001JB000474.

Park, J., 1992. Envelope estimation for quasi-periodic geophysical signals in noise: a multitaper approach, in *Statistics in the Enviromental and Earth Sciences* pp. 189–219, eds Walden, A.T. & Guttorp, P., Edward Arnold, London.

Park, J., Lindberg, C., Vernon, F.L., 1987. Multitaper spectral analysis of high-frequency seismigrams, *J. geophys. Res.,* **92**(B12), 12 675–12 684.

Prejean, S.G. & Ellsworth, W.L., 2001. Observations of earthquake source parameters and attenuation at 2 km depth in the Long Valley Caldera, eastern California, *Bull. seism. Soc. Am.,* **91,** 165–177.

Prieto, G.A., Shearer, P.M., Vernon, F.L. & Kilb, D., 2004. Earthquake source scaling and self-similarity estimation from stacking *P* and *S* spectra, *J. geophys. Res.,* **109,** B08310, doi:10.1029/2004JB003084.

Prieto, G.A., Parker, R.L., Vernon, F.L., Shearer, P.M. & Thomson, D.J., 2006. Uncertainties in earthquake source spectrum estimation using empirical Green functions in AGU *Monograph on Radiated Energy and the Physics of Earthquake Faulting,* in press.

Quenouille, M., 1949. Approximate tests of correlation in time series, *J. R. Stat. Soc.,* **B11,** 18–84.

Reeds, J., 1978. Jackknifing Maximum-Likelihood Estimates, *Ann. Stat,* **6**(4), 727–739.

Shearer, P.M., Hauksson, E. & Lin, G., 2005. Southern California hypocenter relocation with waveform cross-correlation, Part 2: results using source-specific station terms and cluster analysis, *Bull. seism. Soc. Am.,* **95,** 904–915.

Slepian, D., 1978. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V—the discrete case, *Bell Sys. Tech. J.,* **57,** 1371–1430.

Taylor, J.R., 1997. *An introduction to Error Analysis*, 2nd edn. University Science Books, Sausalito, California.

Thomson, D.J., 1982. Spectrum estimation and harmonic analysis, *Proc. IEEE,* **70,** 1055–1096.

Thomson, D.J. & Chave, A.D., 1991. Jackknife error estimates for spectra, coherences, and transfer functions, in *Advances in Spectrum Analysis and Array Processing,* Vol. 1, Chapter 2, pp. 58–113, ed. Haykin, S., Prentice Hall, Englewood Cliffs, New Jersey.

Tukey, J.W., 1958. Bias and Confidence in not-quite large samples, *Ann. Math. Stat,* **29,** 614.

Tukey, J.W., 1960. Conclusions vs Decisions, *Technometrics,* **2**(4), 423–433.

Vernon, F.L., 1989. Analysis of data recorded on the ANZA seismic network, in *PhD thesis,* ed. Harris, B., Spectral Analysis of Time Series, University of California, San Diego.

Wu, C.F.J., 1986. Jackknife, Bootstrap and other resampling methods in regression analysis (with Discussion), *Ann. Stat.,* **14,** 1261–1350.

APPENDIX A: PROPAGATION OF ERRORS

Some rules of propagation of errors (Taylor 1997) are listed here. Assume u, v are random variables with associated variance σ_u^2, σ_v^2, and covariance σ_{uv}^2. Constants a, b do not contribute to uncertainties. Then we have:

$$x = au \pm bv \qquad \sigma^{2x} = a^2\sigma_{u^2} + b^2\sigma_v^2 + 2ab\sigma_{uv}^2$$

$$x = auv \qquad \frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} + 2\frac{\sigma_{uv}^2}{uv}$$

$$x = \frac{au}{v} \qquad \frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} - 2\frac{\sigma_{uv}^2}{uv}$$

$$x = au^{\pm b} \qquad \frac{\sigma_x}{x} = b\frac{\sigma_u}{u}$$

$$x = ae^{\pm bu} \qquad \frac{\sigma_x}{x} = b\sigma_u$$

$$x = a\ln(\pm bu) \qquad \sigma_x = a\frac{\sigma_u}{u}.$$