

# Correction of earthquake location estimation in a small-seismic-array system

YOSHIKO OGATA<sup>1\*</sup>, AKIO KOBAYASHI<sup>2</sup>, NAOYA MIKAMI<sup>2</sup>,  
YASUAKI MURATA<sup>3</sup> and KOICHI KATSURA<sup>1</sup>

<sup>1</sup>*Institute of Statistical Mathematics, Minami-Azabu 4-6-7, Minato-ku, Tokyo 106, Japan*

<sup>2</sup>*Matsushiro Seismological Observatory, Japan Meteorological Agency, 3511 Nishijo, Matsushiro, Nagano-shi, Nagano-ken 381-12, Japan*

<sup>3</sup>*Geological Information Center, Geological Survey of Japan, Higashi 1-1-3, Tsukuba-shi, Ibaraki-ken 305, Japan*

An array of regularly spaced seismic stations can estimate the location of a distant earthquake using arrival times at the stations of seismic waves generated by the earthquakes. However, the accuracy decreases as the distance to the epicentre of the earthquake from the array increases. This paper is concerned with the modification of the estimated location by removing its bias which is locally systematic but globally complex, reflecting the structure of the Earth's interior. Spline surfaces are used to model such biases. Then a Bayesian procedure is carried out not only to tune the smoothness constraints but also to select the best combination among various sums of squares of differently weighted residuals and various roughness penalties for the smoothing. Using the estimated splines of the posterior mode, the newly determined epicentre locations are transformed to confirm its practical utility. Residual distributions show that our procedure improves the modification by the conventional procedure. A spatial pattern of the residuals reveals some geophysical characteristics.

*Keywords:* bias compensation; epicentre; objective Bayesian smoothing; penalized sum of squares; *B*-spline surface

## 1. Introduction

In order to estimate a location of an earthquake accurately, we need a set of densely distributed stations around the source. Currently, more than 1500 stations and observatories in the world are working to report their arrival times and other data to central institutions such as the International Seismological Centre (ISC) and the United States Geological Survey (USGS). These centres collate the data and publish them together with estimated locations. Although it takes longer time to gather and collate the data from many stations, the distribution of the determined epicentres in the world is so accurate that it clearly shows boundaries of tectonic plates on the globe.

The Seismological Observatory of the Japan Meteorological Agency (JMA) at

\*To whom correspondence should be addressed. e-mail: ogata@ism.ac.jp

Matsushiro, central Japan, is one such observatory. It has an array of stations called the Matsushiro Seismic Array System (MSAS) which consists of seven telemetering stations located equidistantly on a circle of radius of about 5 km and at its centre. The MSAS itself locates a focus by a combination of the azimuth of wave approach and the epicentral distance (i.e. distance to the epicentre from the origin of the array). Thus, MSAS plays various roles not only to report the seismic wave arrival times and related data to central institutions in Japan and the world, but also to detect quickly and to determine the locations including those of possible nuclear experiments somewhere in the world (see, for example, Bolt (1976) and Ringdal (1990)). However, except for a range within about  $1^\circ$  (111.16 km) in circular distance from Matsushiro, the distribution of epicentres determined by MSAS alone appears considerably different from that of the true epicentres. There are two types of inevitable estimation error which affect the MSAS epicentre determination procedure (Osada *et al.* 1984; Takayama *et al.* 1985; Kobayashi *et al.* 1993). Namely, errors reading emergent or noisy P arrivals cause an (unbiased) error for epicentre determination, while the heterogeneous velocity structure of the Earth's interior usually causes a biased error depending on the direction of the wave path. For instance, calibration using the Longshot explosion in the Aleutian chain revealed such bias (see, for example, Herrin and Taggart (1968), Carder *et al.* (1967) and Bolt (1976)).

Based on their experience in data processing, Kobayashi *et al.* (1993) constructed a table of transformation for the correction of the biases by gridding and interpolating the data of differences of the MSAS epicentres from the corresponding entries in global earthquake catalogues in Japan and world. Using the table, the modified epicentral locations of the original MSAS locations have been published by Matsushiro Seismological Observatory (1984–1992).

However, it seems difficult to make a satisfactory correction in areas where the events are dense. Another difficulty is the task of interpolation to regions where the MSAS locations of events are sparse or non-existent. Kobayashi *et al.* (1993) recommended update of the transformation at each stage of data accumulation. However, it is a very elaborate and time-consuming task to make a better partition for gridding the data by trial and error.

To overcome such technical difficulties, we model a smooth transformation on the plane to estimate and compensate the biases using two-dimensional *B*-spline functions. Then, very many spline coefficients have to be estimated by minimizing the penalized least squares with automatically determined penalties and their weights by an objective Bayesian method. If this procedure is feasible, then the elaborate update of the transformation for epicentre compensation is reduced to a mere routine of the statistical computation. In this paper we shall implement this procedure and demonstrate the performance using the MSAS location data of earthquakes in the world relative to the corresponding accurate data from a global catalogue.

## 2. Data

The three-dimensional location of an earthquake (i.e. the location of its hypocentre) is usually estimated by minimizing the sum of squares of weighted residuals:

$$\sum_i \frac{\{t_i - t_0 - T(\Delta_i, h)\}^2}{\sigma_i^2}, \quad (1)$$

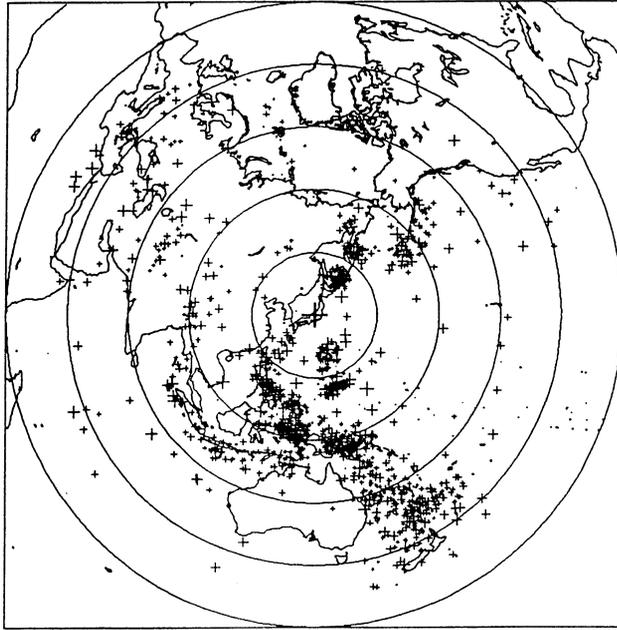
where the sum is taken over the different stations selected for the purpose,  $\sigma_i$  stands for an error factor of both the modelling and measurement,  $t_i$  is the measured arrival time of a seismic wave group (usually P waves) at the  $i$ th station,  $t_0$  is the origin time,  $\Delta_i$  is the distance between the epicentre (projected hypocentre to the global surface from the Earth's center) and the  $i$ th station,  $h$  is the focal depth, and the nonlinear function  $T(\Delta, h)$  is the predicted travel time obtained from standard travel time tables. For instance, the global travel time table by Jeffreys and Bullen (1940) is still used for compiling the world-wide hypocentre catalogues of earthquakes. For more detail of earthquake location, the reader is referred to Vere-Jones and Smith (1981) who provide an excellent introduction to the statistical aspect of a range of problems in seismology.

However, unlike the ordinary hypocentre determination procedure described in (1), the data determined by MSAS have, in principle, been produced according to the method by Mikumo (1965) and Otsuka (1966) in the following manner. A certain location within MSAS is considered as the origin of a polar coordinate system of epicentral distance and azimuth which are distance to the epicentre from the origin of the array and angle around the origin measured clockwise from the north direction to the epicentre, respectively.

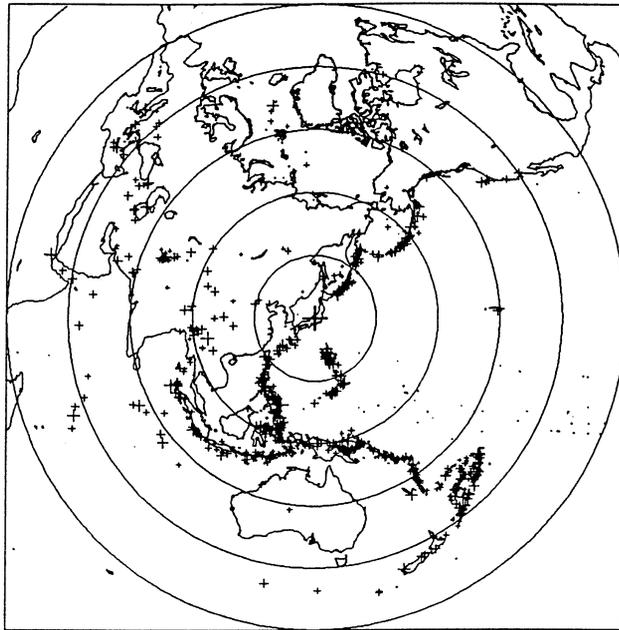
Using P-wave arrival times at the seven stations of the array, the azimuth  $\Psi$  of the wave approach is calculated in association with an estimate of its standard error  $\sigma_\Psi$ , whereas the epicentral distance  $\Delta$  from the MSAS origin to an epicentre, obtained by measuring the time span between P- and S-wave arrivals, is determined only by a certain single station of MSAS, so that no error assessment is associated with the estimate of  $\Delta$ . It would be more favourable to compute these parameters through modelling the time series of seismic waves from the array (see, for example, Brillinger (1985)), but the number of earthquakes to be detected here is prohibitive.

For each event, the location in the MSAS data is compared with that in Earthquake Data Reports compiled by USGS for the world which provide much more accurate locations since they were determined using arrival times observed at a sufficiently large number of stations around epicentres in most cases. Both data sets are chronologically divided into two parts. The first part for the term 1984–1988 will be used for the estimation of models, and the remainder 1989–1992 will be used to evaluate the estimated models. Figure 1a shows the epicentre locations of the latter period as determined by MSAS. Here, we should note that the earthquakes detected and located by MSAS are confined to a distance of about  $100^\circ$  (i.e. 11116 km) from the MSAS origin, because direct S waves from an earthquake occurring beyond this distance do not reach the MSAS owing to the refraction of waves by the Earth's core.

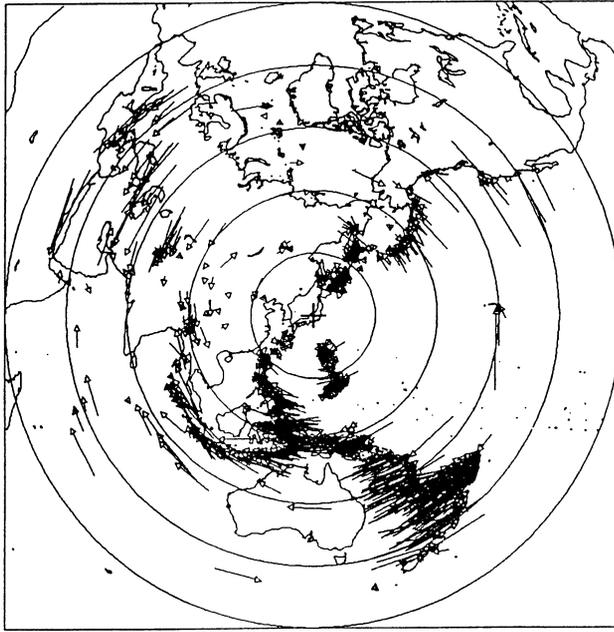
Comparing these epicentre locations (Figure 1a) with the corresponding locations by USGS (Figure 1b), the epicentre distribution of MSAS along the Circum-Pacific seismic zone does not appear to be well focused. Most of the events actually occur along the plate boundaries. The investigations by Osada *et al.* (1984) and later related work by some JMA staff suggested that the location error included a significant bias component. For instance, differences of the locations of the same events are illustrated in Figure 1c, where an arrow



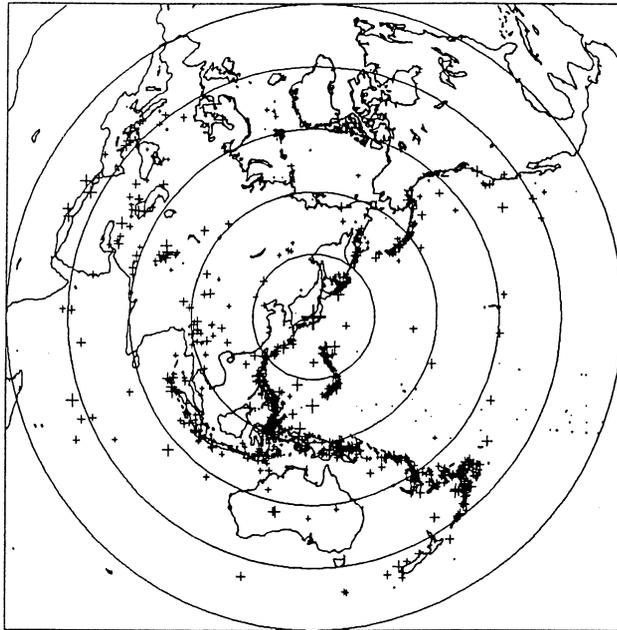
(a)



(b)



(c)



(d)

means difference from the MSAS location to the USGS location. Thus, the locally similar directions of the arrows clearly indicate such a bias in the area. In particular, the biases with respect to the azimuth are conspicuous since the directions are mostly parallel with the circles centred at the MSAS origin. In the following sections, we analyse such biases in detail for the practical purpose of predicting more accurate epicentres.

### 3. Models for estimating biases and Bayesian model selection

Let an epicentre location estimated by MSAS be  $(\Delta_i, \Psi_i)$  in the polar coordinates centred at the MSAS origin, and the corresponding true location be  $(\Delta_i^0, \Psi_i^0)$ . We may then analyse the structure of the bias with

$$\Delta_i^0 = \Delta_i + f(\Delta_i, \Psi_i) + \epsilon_i, \quad (2)$$

$$\Psi_i^0 = \Psi_i + g(\Delta_i, \Psi_i) + \eta_i, \quad (3)$$

for  $i = 1, 2, \dots, N$ , where  $\epsilon_i$  and  $\eta_i$  are unbiased errors probably owing to the reading errors, while the functions  $f$  and  $g$  are the bias functions which are dependent upon the location  $(\Delta_i, \Psi_i)$  determined by MSAS. For simplicity, we assume that  $\{\epsilon_i\}$  and  $\{\eta_i\}$  are mutually independent and Gaussian. One of the reasons for using the polar coordinates for the models is the convenience for easy understanding of the biases of both epicentral distance and azimuth as anomalies in the seismic wave velocities through the media between an epicentre and the observatory. Another reason is that, in the original MSAS data, a set of standard errors  $\{\sigma_i\}$  for azimuth measurement is available as additional data. These errors are obtained as byproducts in estimating the azimuth by the arrival times at the seven MSAS stations.

We shall examine whether or not these error data are useful for the estimation of the standard deviation of the error  $\{\eta_i\}$  in (3). That is to say, whether  $\eta_i \sim N(0, s^2\sigma_i^2)$  for some constant  $s^2$ . On the other hand, for the other error  $\epsilon_i$  in (2), it has been empirically determined that the estimation error of the epicentral distance  $\Delta$  may be roughly proportional to  $\Delta^{1/2}$ . We shall also examine whether this is useful, i.e. whether  $\epsilon_i \sim N(0, s^2\Delta_i)$  can improve the fit to the data.

We parametrize each of the spatial functions  $f$  and  $g$  by cubic two-dimensional  $B$ -spline functions in the following way; a disc  $A$  with radius  $R$  of epicentral distance from the

**Fig. 1.** Epicentre locations (plus signs) for the earthquakes which occurred in 1989–1992 determined by MSAS. The geographical world map is drawn by a polar coordinate system whose centre is the location of the MSAS origin (the large plus sign) in central Japan. The interval between radii of the concentric circles is  $20^\circ$  (about 2222 km) in the angular distance. The sizes of the plus signs are proportional to the standard error of the azimuth estimate. (a) Original MSAS locations. (b) USGS locations. (c) Difference of epicentre locations of the same earthquake. An arrow indicates the shift from the MSAS location to the USGS location, provided that the differences of epicentral distance and azimuth are within  $2^\circ$  and  $25^\circ$ , respectively. (d) Modified MSAS locations (referred to in Section 4.1) compensated by the mapping  $\varphi$  in (8) using the estimated functions shown in Figure 2a and b.

origin in polar coordinates is regarded as a rectangle  $[0, R] \times [0^\circ, 360^\circ)$  in the ordinary Cartesian coordinates and is equally divided into  $M_\Delta \times M_\Psi$  rectangular subregions. For each of  $f$  and  $g$ , we estimate the spline coefficients  $\theta$  whose number is  $M = (M_\Delta + 3)(M_\Psi + 3)$ . For the explicit definition of the bi-cubic  $B$ -spline function, see Inoue (1986), Ogata and Katsura (1988, 1993) and Ogata *et al.* (1991).

We assume that the functions  $f$  and  $g$  for the biases are smooth enough for the stable estimation of a large number of coefficients. We consider the following penalties for the function  $f$ :

$$\begin{aligned}\Phi_1(f) &= \iint_A \left\{ \left( \frac{\partial f}{\partial \Delta} \right)^2 + \left( \frac{\partial f}{\partial \Psi} \right)^2 \right\} \xi(d\Delta, d\Psi), \\ \Phi_2(f) &= \iint_A \left\{ \left( \frac{\partial^2 f}{\partial \Delta^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial \Delta \partial \Psi} \right)^2 + \left( \frac{\partial^2 f}{\partial \Psi^2} \right)^2 \right\} \xi(d\Delta, d\Psi)\end{aligned}\quad (4)$$

to measure the roughness of the function, where  $\xi(d\Delta, d\Psi)$  is either  $d\Delta d\Psi$  or  $\Delta d\Delta d\Psi$ , discriminating whether or not the penalty is proportional to the area in the ordinary Cartesian coordinates. The same penalties are considered for the function  $g$ . Then, consider penalized sums of squares of residuals given by

$$Q(\theta; w_1, w_2) = \sum_{p=1}^N \left( \frac{\Delta_p^0 - f(\Delta_p, \Psi_p)}{\sigma(\Delta_p)} \right)^2 + w_1 \Phi_1(f) + w_2 \Phi_2(f) \quad (5)$$

and

$$Q(\theta; w_3, w_4) = \sum_{p=1}^N \left( \frac{\Psi_p^0 - g(\Delta_p, \Psi_p)}{\sigma(\Psi_p)} \right)^2 + w_3 \Phi_1(g) + w_4 \Phi_2(g). \quad (6)$$

Minimization of these (Good and Gaskins 1971) is carried out independently for given pairs of weights  $(w_1, w_2)$  or  $(w_3, w_4)$  to be determined by our Bayesian procedure described later. Here,  $\sigma(\Delta_p)$  is either  $\Delta_p^{1/2}$  or constant, and  $\sigma(\Psi_p)$  is either  $\sigma_p$  or constant, the reasons for which were stated above.

The models to be compared are summarized in the following. Firstly, we examine whether the standard error components  $\sigma(\Delta_p)$  of the sum of squares of the epicentral residuals in (5) is proportional either to  $\Delta_p^{1/2}$  or to a constant, and also whether the sums of squares of azimuthal residuals  $\sigma(\Psi_p)$  in (6) is proportional either to  $\sigma_p$  given in the MSAS data or to a constant. Secondly, the roughness penalty integral in (4) is taken to be either of  $d\Delta d\Psi$  or  $\Delta d\Delta d\Psi$ . Finally, we also choose suitable sizes of  $M_\Delta$  and  $M_\Psi$ .

To solve these issues, we interpret the penalized sums of squares in (5) and (6) as a part of the posterior function in the Bayesian framework. The sum of squares in (5), for instance, corresponds to a Gaussian likelihood  $L$ , and the sum of the two penalty functions to a Gaussian prior distribution.

The normalizing factor for posterior distribution

$$\Lambda(w_1, w_2, \theta_M) = \int_{R^{M-1}} L(\boldsymbol{\theta}) \text{prior}(\boldsymbol{\theta}^r | w_1, w_2, \theta_M) d\boldsymbol{\theta}^r \quad (7)$$

plays an important role and is called the likelihood of the Bayesian model with respect to the hyperparameters  $w_1$  and  $w_2$ . Here the parameter vector  $\boldsymbol{\theta}$  is divided into  $(\boldsymbol{\theta}^r, \theta_M)$  such that the prior  $(\boldsymbol{\theta}^r | w_1, w_2, \theta_M)$  is proper (Ogata and Katsura 1988). The optimal hyperparameter values  $\hat{w}_1$  and  $\hat{w}_2$  are those which maximize the likelihood in (7), called type II maximum-likelihood estimates by Good (1965). Further, for the comparison of the possible options of the models summarized above, Akaike's (1980) Bayesian information criterion (ABIC) given by

$$\text{ABIC} = (-2) \max_{\rho} \{\log \Lambda(\rho)\} + 2 \dim(\rho)$$

is useful, where the hyperparameter  $\rho$  is  $(w_1, w_2, \theta_M)$  in the present case. A model with smaller ABIC is considered to be a better fit. The estimation of the function  $g$  is similarly carried out. See Ogata and Katsura (1988, 1993) and Ogata *et al.* (1991) for some detail and Appendix 1 for related technical aspects and their references.

## 4. Implementation

### 4.1. Estimation and the correction

The number of earthquakes in the world, except for Japan and its vicinity, detected and located by both MSAS and USGS from 1984 to 1988, was 1780. Since the MSAS data include quite a few outliers for the estimates of epicentral distance and azimuth, we removed them for the analysis by the criteria  $|\Delta_A - \Delta_U| \geq 2^\circ$  or  $|\Psi_A - \Psi_U| \geq 25^\circ$  in view of the empirical distributions of the differences.

In general, a larger number of parameters is preferred for a Bayesian model as far as its implementation is feasible, but a sufficient number can be judged by minimization of the ABIC based on the data set. In fact, it seems that the bias functions  $f$  and  $g$  need to be fairly complicated, because the ABIC decreases substantially as the number of subdivisions (knots) for the spline bases increases. Specifically, let the numbers  $M_\Delta$  and  $M_\Psi$  of the spline knots be restricted to the case when  $M_\Delta = M_\Psi$ . Then, for increasing numbers of knots, say  $M_\Delta = 3 \times 2^k$ , ( $k = 1, 2, \dots$ ), the series of corresponding  $\text{ABIC}_k$  values decreases and converges to a constant when the number  $M = (M_\Delta + 3)^2$  of parameters (spline coefficients) become as large as 10 000. Accordingly, we shall use 9801 parameters (i.e.  $M_\Delta = M_\Psi = 96$ ) for the estimations.

The models stated and summarized in Section 3 are applied to the data to compare their goodness of fit. The ABIC value for each model is listed in Table 1. The smallest ABIC value among the competing models suggests that the best-fitted model for the  $B$ -spline expansion of the bias  $f$  should be estimated in association with  $\sigma(\Delta) = \Delta^{1/2}$  in (5) and with  $\xi(d\Delta, d\Psi) = d\Delta d\Psi$  for the penalty  $\Phi(f)$  in (4). The best-fitted model for  $g$  should be estimated in association with  $\sigma(\Psi_p) = \sigma_p$  in (6) and with  $\xi(d\Delta, d\Psi) = d\Delta d\Psi$  for the penalty  $\Phi(g)$  in (4).

**Table 1.** ABICs of Bayesian models in (5) and (6)

Model (5)	$\xi(d\Delta, d\Psi)$	$\sigma(\Delta)$	ABIC
$f(\Delta, \Psi)$	$d\Delta d\Psi$	$\Delta^{1/2}$	6 116.5
		Constant	6 604.0
	$\Delta d\Delta d\Psi$	$\Delta^{1/2}$	6 183.9
		Constant	6 736.0
Model (6)	$\xi(d\Delta, d\Psi)$	$\sigma(\Psi)$	ABIC
$g(\Delta, \Psi)$	$d\Delta d\Psi$	$\sigma_p$	12 858.7
		Constant	14 335.4
	$\Delta d\Delta d\Psi$	$\sigma_p$	12 969.2
		Constant	14 332.8

In this way, we have obtained the optimum models which minimize ABIC and also the optimum hyperparameter values  $w_1 = 4.2 \times 10^{-2}$ ,  $w_2 = 3.1 \times 10^{-6}$ ,  $w_3 = 5.0 \times 10^{-3}$  and  $w_4 = 1.8 \times 10^{-6}$ , which maximize the likelihood in (7) and the corresponding one for  $g$ . Simultaneously,  $\hat{f}$  and  $\hat{g}$  are estimated by minimizing (5) and (6), respectively. See Appendix 1 for the technical aspects.

The contour map of the  $\hat{f}$  values in Figure 2a shows how the epicentre distance of a location determined by MSAS must be shrunk in the shaded area or stretched in the white area to compensate for the bias. In the disc of radius  $100^\circ$  centred at the MSAS origin on the globe, regions of positive  $\hat{f}$  value for the MSAS location are relatively small but with highly clustered earthquakes.

Figure 2b includes a contour map of the estimated  $g$  function showing how the azimuth of a location determined by MSAS must be rotated around the origin clockwise in the white area and anticlockwise in the shaded area in order to compensate the azimuth for the bias. It appears consistent with our intuition that some portion of the boundaries between the white and shaded areas coincide with the major seismic belts of actual epicentre locations in Figure 1b. Further, although a unit degree in azimuth has different meanings from that of epicentral distance, it is seen by the comparison of Figure 2a and b that bias in azimuth is much larger than bias in epicentral distance on the average. This is also intuitively clear in view of Figure 1c.

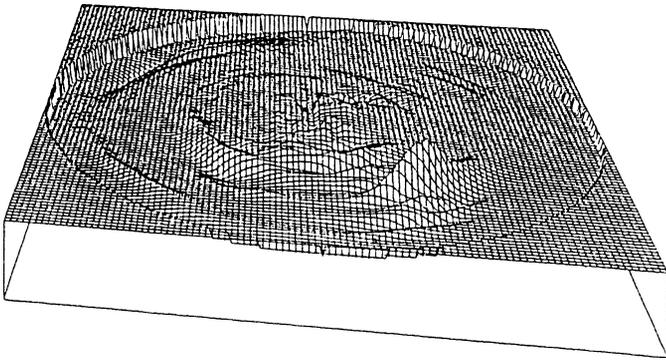
Having thus obtained optimal functions  $\hat{f}$  and  $\hat{g}$  using epicentre data of both MSAS and USGS for the time span 1984–1988 by the above procedure, the shift transformation

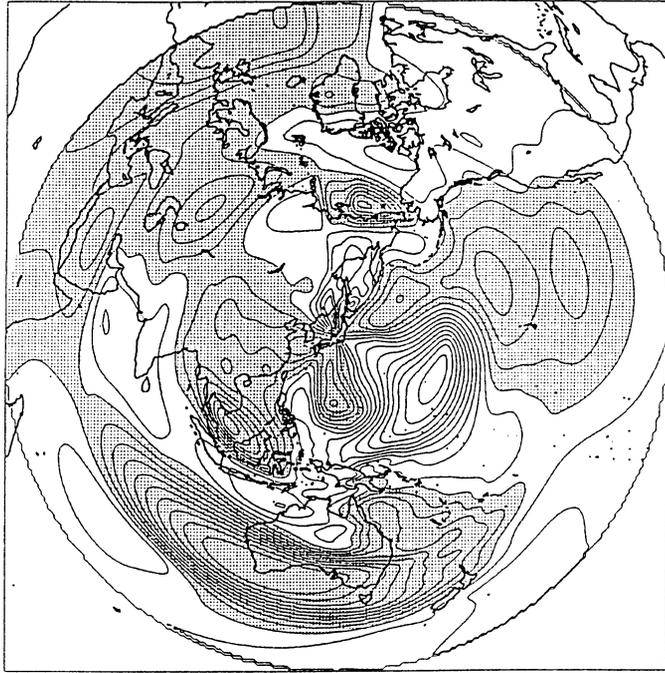
$$\varphi: (\Delta, \Psi) \mapsto (\Delta, \Psi) + (\hat{f}(\Delta, \Psi), \hat{g}(\Delta, \Psi)) \quad (8)$$

in the polar coordinates is defined to correct the further MSAS epicentre data of the world for the span 1989–1992 shown in Figure 1a. Figure 1d is the modification based on the  $\varphi$ . The out-of-focus appearance of Figure 1a has become a “focused” image in Figure 1d, which closely resembles the epicentre map in Figure 1b due to the USGS catalogue. This indicates

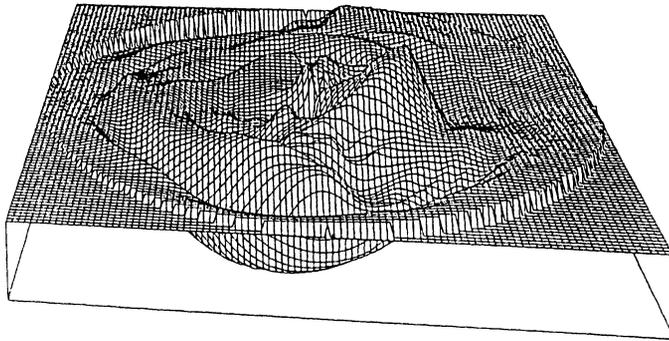


(a)





(b)



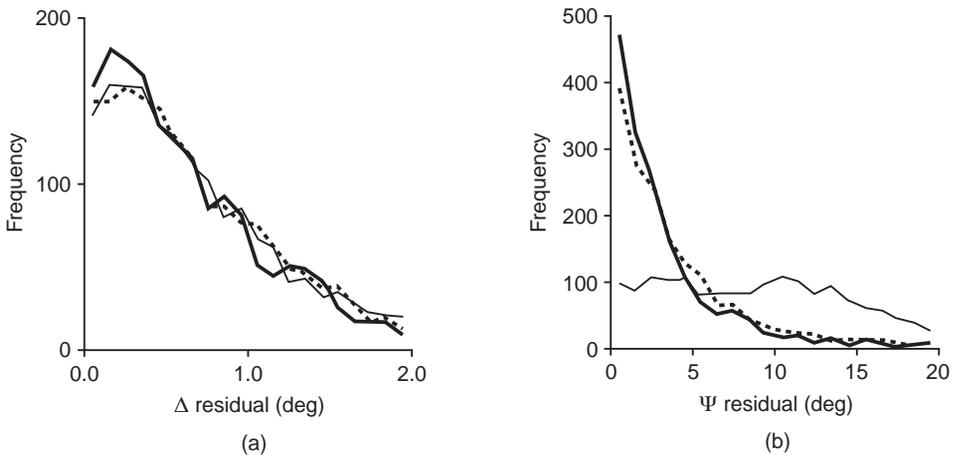
**Fig. 2.** A contour plot and a bird's-eye view of the functions (a)  $\hat{f}$  and (b)  $\hat{g}$ , showing local biases of epicentral distance and azimuth, respectively, estimated using the data from 1984–1988. The function  $\hat{f}$  varies from  $-3.5^\circ$  to  $2.5^\circ$  in the angular distance, and the contour interval is  $0.5^\circ$ . The function  $\hat{g}$  varies from  $-45.0^\circ$  to  $50.0^\circ$ , and the contour interval is  $5.0^\circ$ . Areas of negative function values are shaded.

that the performance is quite satisfactory. On the other hand, we see a number of isolated locations of epicentres in Figure 1d which are not seen in Figure 1b such as those in the central part of the Pacific Ocean and far off the west coast of California. It may be that the poor performance of the estimators for these locations originates from the fact that the MSAS locations are outliers, i.e. isolated, and have large azimuth errors.

## 4.2. Examination of residuals

Using the data sets for the time span 1989–1992, residuals of the USGS epicentres from the compensated MSAS epicentres by our method are investigated in comparison with the performance of the conventional method by Kobayashi *et al.* (1993). Figure 3a and b includes the distribution of residuals of epicentral distances  $r_{\Delta} = \Delta_U - \Delta_A - f(\Delta_A, \Psi_A)$  and that of azimuths  $r_{\Psi} = \Psi_U - \Psi_A - \hat{g}(\Delta_A, \Psi_A)$ , respectively, where  $(\Delta_A, \Psi_A)$  and  $(\Delta_U, \Psi_U)$  are a MSAS location and a USGS location, respectively. Figure 3a shows that the epicentral distances by the conventional modification results in about the same performance as those of the original MSAS data, but that the epicentral distances by our method are slightly improved. Figure 3b shows that the azimuths by the conventional modification are greatly improved from the original MSAS data, but that the azimuths by our method are consistently better than those by the conventional modification, particularly in closer distance to the MSAS.

Some residual plots on the global surface reveal limitations of our model. For instance, Figure 4a and b present residuals for epicentral distance  $r_{\Delta}$  and azimuth  $r_{\Psi}$ , respectively, on each USGS location in the eastern part of the Circum-Pacific seismic belt. First, sizes of the



**Fig. 3.** Distribution of residuals of (a) epicentral distance and (b) azimuth. The thick solid lines show our correction, the dotted lines are the conventional correction, and the thin solid lines are the original MSAS data.

azimuth residuals in Figure 4b suggest that the regional variability may differ from place to place, which we did not assume in the modelling of the standard deviation in (6).

Also, there are some areas where a significant bias of the residuals is seen. In particular, Figure 4b shows the dominating negative residuals (plus signs) in the eastern part of New Guinea, in addition to the dominating positive residuals (circles) in the zone from Philippine islands to western New Guinea. This means either that a MSAS location in a certain area does not always correspond to a unique real location or that the transformation of azimuth should be more drastic in some areas than that given in Figure 2b. This is a limitation of  $\hat{g}$  in the particular area since we estimated it by assuming the smoothness prior given in (6) throughout the whole region as well as assuming  $\varphi$  to be a one-to-one mapping. A geophysical reason for such phenomena can be that there are two subducting tectonic plates (i.e. Philippine Sea Plate and Pacific Plate) through which the wave paths from the area to the Matsushiro array convey the P waves anomalously fast.

Similar bias is seen along the Kamchatka Peninsula and the Chishima (Kuril) Islands arc. The shallow events are positively biased, but the deep events are negatively biased. This is another limitation of  $\hat{g}$  because we did not take the depth into consideration in our modelling.

In comparison with the azimuth residual, we cannot see any conspicuous regional bias nor variability of the residuals for epicentral distances in Figure 4a.

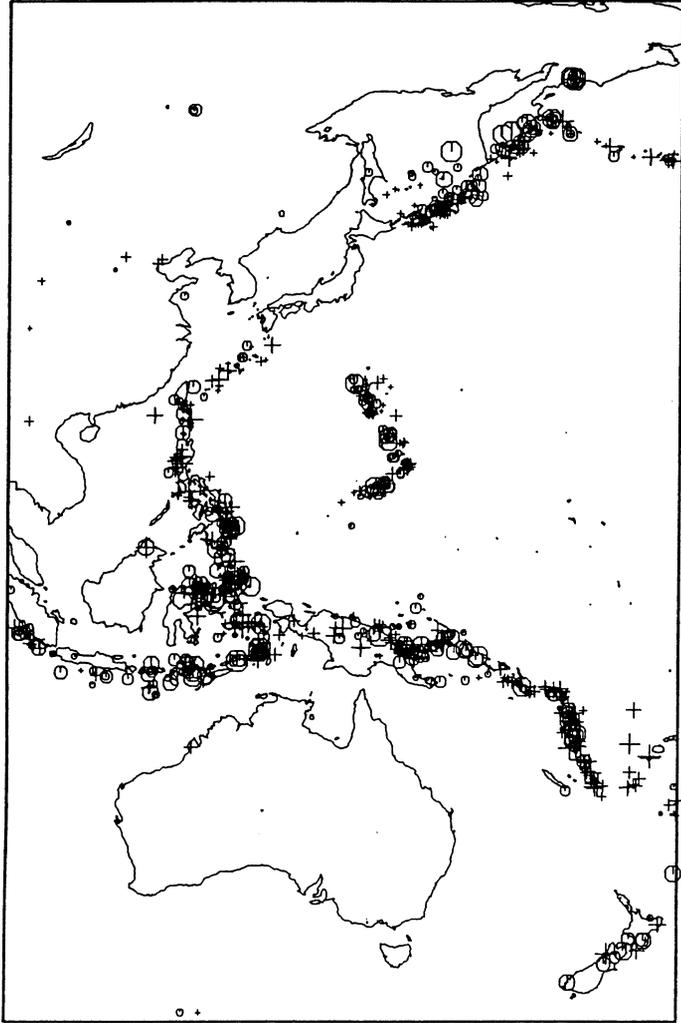
## 5. Concluding remarks

We have discussed transformations which modify the location of an earthquake determined by MSAS for a better epicentre estimation by compensating a location for a bias owing to the laterally heterogeneous structure of the Earth. The biases are systematic but complex so that a very large number of parameters are necessary to represent the transformation. An objective Bayesian method is carried out in order to choose the optimal hyperparameters for the smoothing of the transformation to regulate the parameter determination together with the selection of the models consisting of sums of squares of the variously weighted residuals and various penalties.

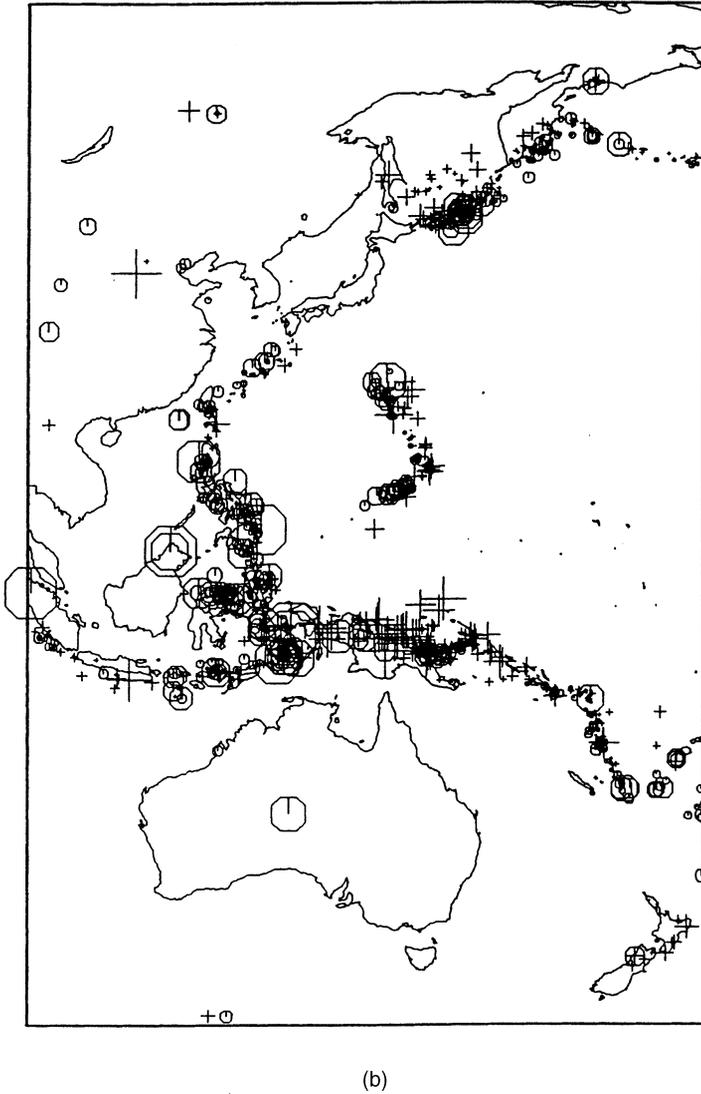
Having obtained the optimal functions  $\hat{f}$  and  $\hat{g}$  using the MSAS epicentre data relative to USGS data of the same earthquakes for the time span 1984–1988 by the above procedure, the transformation  $\varphi$  in (8) is used to correct the other part of the MSAS epicentre data for 1989–1992. The modified locations are quite similar to those from the USGS data for 1989–1992. The residual distributions suggest that our procedure improves the conventional procedure. The spatial residual analyses show the limitations of our model due to various geophysical reasons.

The separation of MSAS data at 1988–1989 is only for the demonstration of our procedure. In practice, all the 1984–1992 data removing the outliers should be used to construct the transformation for modification of the currently determined MSAS locations. An updated transformation for accumulated data at each stage can be made automatically by the present objective Bayesian procedure.

Detailed geophysical implications of compensation mappings  $\varphi$  and their residuals in



(a)



**Fig. 4.** Spatial plot of residuals of (a) epicentral distance and (b) azimuth against the corresponding location of USGS data, respectively. The circles and plus signs mean that the residual takes positive and negative values, respectively; their sizes shows the absolute values. The residual for epicentral distance ranges from  $-0.3^\circ$  to  $0.3^\circ$ , while the residual for azimuth ranges mostly from  $-2.0^\circ$  to  $2.0^\circ$ .

Japan and the world remains for a future task. In particular, relating to geophysical interpretation of seismic wave travel time anomalies, there is a difficult but challenging statistical task to extend the present compensation procedure to three-dimensional hypocentre correction including depth of an earthquake.

## Appendix 1: Computational remarks

Since both the likelihood and prior are Gaussian, the (integrated) likelihood of the Bayesian model in (7), for instance, can be provided in the algebraic form such that

$$\Lambda(w_1, w_2, \theta_M) = \frac{N}{2} \log(2\pi\hat{s}^2) - \frac{1}{2} \sum_{p=1}^N \log\{\sigma(\Delta_p)^2\} + \frac{1}{2} \log|C^r| - \frac{1}{2} \log|H|, \quad (9)$$

where  $C^r$  is the cofactor of the matrix  $C$  (with respect to  $\theta_M$ ) such that  $\theta \cdot C\theta^T = w_1\Phi_1(f) + w_2\Phi_2(g)$ ,  $\hat{s}^2 = \min_{\theta} Q(\theta; w_1, w_2)/N$  with  $Q$  in (5), and  $H = H(w_1, w_2)$  is the Hessian matrix such that  $H = -\partial^2 Q/(\partial\theta^T \partial\theta)$ .

Therefore, given the hyperparameters  $(w_1, w_2)$ , it is required to minimize the penalized least squares in (5) and (6), and also to calculate the determinants in (9). However, the required number of parameters  $\theta$ , i.e.  $M = (M_x + 3)(M_y + 3)$ , will be as large as about  $100 \times 100$  so that the Hessian matrix is as large as about  $10\,000 \times 10\,000$ . This is too large to carry out standard nonlinear optimization, unlike the Bayesian optimization procedure in Ogata and Katsura (1988, 1993) and Ogata *et al.* (1991). Nevertheless, since the present posterior is exactly Gaussian, the following two techniques are available in computing the above procedure.

First, to find minimizing parameters  $\theta$  in (5) and (6), we solve the normal equation where the symmetric Hessian matrix  $H$  is decomposed by the modified Cholesky decomposition in such a way that  $H = U^T \Lambda U$  where  $U$  is the upper-triangular matrix and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_M)$ . Second, for calculation of the determinants of  $C^r$  and  $H$  in (9), we can use  $\Lambda$  in such a way that  $\log|H| = \sum_{m=1}^M \log\lambda_m$ . Non-zero elements of both the symmetric matrices  $H$  and  $C^r$  are very sparse and located in a narrow parallel band within the distance  $3M_x + 13$  from the diagonal elements of the matrix, owing to the local bases of the cubic  $B$ -spline function.

Murata and Noro (1993) successfully compressed the square matrix of size  $M \times M$  into a rectangular matrix of size  $M \times (3M_x + 13)$  to perform the modified Cholesky decomposition efficiently. It turns out that, as the size  $M$  increases, this procedure is much faster and requires far less central processing unit memory than the method of Ogata and Katsura (1988, 1993) and Ogata *et al.* 1991 which uses the Davidon–Fletcher–Powell nonlinear minimization algorithm (see, for example, Fletcher and Powell (1963)) and also than the method of Murata (1993) using the Householder decomposition. Inoue (1986) included useful subroutines for computations of two-dimensional  $B$ -splines in (5) and their integrals for the penalty in (7).

## Acknowledgements

This research was carried out under the ISM Cooperative Research Program (93-ISM-CRP-51, 94-ISM-CRP-49 and 95-ISM-CRP-47). We thank Professor Tokuji Utsu for valuable comments which were useful clarification. We would also like to thank the case editor and the referee for helpful comments.

## References

- Akaike, H. (1980) Likelihood and Bayes Procedure. In Bernard *et al.* (eds), *Bayesian Statistics*, pp. 143–166. Valencia: University Press.
- Bolt, B.A. (1976) *Nuclear Explosions and Earthquakes*. San Francisco, CA: W.H. Freeman.
- Brillinger, D.R. (1985) A maximum likelihood approach to frequency wave number analysis. *IEEE Trans. Acoust. Speech Signal Processing*, **33**, 1076–1085.
- Carder, D.S., Tocher, D., Bufe, C., Stewart, S.W., Eisler, J. and Berg, E. (1967) Seismic wave arrivals from Longshot, 0° to 27°. *Bull. Seismol. Soc. Amer.*, **57**, 573–590.
- Fletcher, R. and Powell, M.J.D. (1963) A rapidly convergent descent method for minimization. *Comput. J.*, **6**, 163–168.
- Good, I.J. (1965) *The Estimation of Probabilities*. Cambridge, MA: MIT Press.
- Good, I.J. and Gaskins, R.A. (1971) Nonparametric roughness penalties for probability densities. *Biometrika*, **58**, 255–277.
- Herrin, H. and Taggart, J. (1968) Source bias in epicenter determination. *Bull. Seismol. Soc. Amer.*, **58**, 1791–1796.
- Inoue, H. (1986) A least square smooth fitting for irregularly spaced data: finite element approach using the cubic *B*-spline basis. *Geophysics*, **51**, 2051–2066.
- Jeffreys, H. and Bullen, K.E. (1940) *Seismological Tables*. London: British Association for Advancement of Sciences, Gray–Milne Trust.
- Kobayashi, A., Mikami, N. and Ishikawa, Y. (1993) Epicentral correction for the hypocenter of Matsushiro Seismic Array System (in Japanese with English captions for figures). *Tech. Rep. Matsushiro Seismol. Obs. Japan Meteorol. Agency*, **12**, 1–9.
- Matsushiro Seismological Observatory (1984–1992) *Seismol. Bull. Matsushiro Seismolog. Obs., Series 3*, 1–8. Nagano City: Japan Meteorological Agency.
- Mikumo, T. (1965) Determination of phase velocity and direction of wave approach from station arrays. *Bull. Disaster Prevention Res. Inst.*, **15**, 31–43.
- Murata, Y. (1993) Estimation of optimum average surficial density from gravity data: an objective Bayesian Approach. *J. Geophys. Res.*, **98**, 12 097–12 109.
- Murata, Y. and Noro, H. (1993) Optimization of estimation procedure of bouguer density by use of the ABIC-minimization method (in Japanese with English summary and captions for figures). *Butsuri-Tansa*, **46**, 120–127.
- Ogata, Y., Imoto M. and Katsura, K. (1991) 3-D spatial variation of *b*-values of magnitude-frequency distribution beneath the Kanto District, Japan. *Geophys. J. Int.*, **104**, 135–146.
- Ogata, Y. and Katsura, K. (1988) Likelihood analysis of spatial inhomogeneity for marked point patterns. *Ann. Inst. Statist. Math.*, **40**, 29–39.
- Ogata, Y. and Katsura, K. (1993) Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues. *Geophys. J. Int.*, **113**, 727–738.
- Osada, H., Kasiwabara, S., Nagai, A., Takayama, H., Wakui, S., Morisita, I. and Tanaka, Y. (1984)

- Matsushiro Seismic Array System (I)—Summary and hypocenter determination of near earthquakes (in Japanese with English figure captions). *Tech. Rep. Matsushiro Seismol. Obs. Japan Meteorol. Agency*, **5**, 13–31.
- Otsuka, M. (1966) Azimuth and slowness anomalies of seismic waves measured on the Central California Seismographic Array. Part I. Observations. *Bull. Seismol. Soc. Amer.*, **56**, 223–239.
- Ringdal, F. (ed.) (1990) *Bull. Seismol. Soc. Amer.: Regional Seismic Arrays and Nuclear Test Ban Verification*, **80**. El Cerrito, CA: Seismological Society of America.
- Takayama, H., Wakui, S. and Nagare, S. (1985) Matsushiro Seismic Array System (II)—Hypocenter determination of very-near earthquakes (in Japanese with English figure captions). *Tech. Rep. Matsushiro Seismol. Obs. Japan Meteorol. Agency*, **6**, 67–72.
- Vere-Jones, D. and Smith, E.G.C. (1981) Statistics in seismology. *Commun. Statistics—Theory Methods A*, **10**, 1559–1585.

Received October 1995 and revised September 1997