A METHOD FOR ESTIMATING FAULT PROPERTIES OF A FAULT POPULATION

Zhuoheng Chen and Quanquan Fu
University of Petroleum, Changping, 102200, Beijing, P. R. China
(Received: 31 October 1995; Accepted: 31 January 1996)

Abstract: Reservoir which has been subject to brittle deformation contains faults with displacements from a few millimeter to tens and hundreds meters, but only part of these faults may be mapped by seismic survey. Despite the complexity of the faults and fault system, two major factors, the resolution of detection tool and the grid density of line search, influence the detectability of a fault. This paper discusses the impact of seismic grid pattern on the unmapped faults and proposes the use of Horvitz-Thampson estimator for estimation of number and sizes of fault unmapped by seismic survey in a region under investigation. Computer generated fault populations in space are then used to illustrate the application of this probability model.

Key Words: Unmapped faults, detectability, Horvitz-Thampson estimator

INTRODUCTION

Fault as a geological phenomena plays an important role in petroleum accumulation and reservoir compartmentation. Reservoir which has been subject to brittle deformation contains faults with displacements from a few millimeter to tens and hundreds meters. However only part of these faults may be mapped due to seismic grid line density and seismic resolution in a petroleum exploration. Given data quality, the denser the seismic grid is, the larger the number of faults can be mapped in a reservoir horizon. General knowledge of rock mechanics and analog to outcrop observations suggest that small faults are also present in reservoir rock and the frequency of fault occurrence increases with decreases in fault displacements. These small faults in conjunction with fractures contribute in a large portion to reservoir heterogeneity and exert influence on the subsurface fluid flow. It is therefore important to investigate the characteristics and spatial distribution of the small faults and fractures as early as possible in a reservoir study.

The distribution characteristics of a fault population have been intensely studied in recent years. The power law or fractal model has been suggested and tested by several researchers (Heffer, 1989; Marret et al. 1992, Walsh etc., 1991 and 1994). However problems still remain in systematically estimating both the number and size distribution characteristics of faults missed by seismic grid line search of a region. This article proposes a moment estimator based on a well-known finite population survey theory, the Horvitz-Thampson (H-T) estimator, to estimate the properties of the finite population of faults. The application of the method to the prediction of the number and sizes of unmapped faults is then illustrated through computer simulated fault patterns in space.

THE HORVITZ-THAMPSON (H-T) ESTIMATOR

Suppose that there are N faults in an area under investigation, representing a finite population U of faults. For a given seismic grid line density #(A,B), n faults are mapped by the seismic survey (n \( \leq N \)), and their associated magnitudes \( Y = \{y_1, y_2, \ldots, y_n\} \), representing one property of the fault population. The magnitude can be fault length, throw or displacement, and other measures. If \( g(y_i) \) is a single valued function of \( y_i \), an unbiased estimate of the property \( G = \sum_{i=1}^{n} g(y_i) \), from the parent population U, is then

\[
G = \frac{n \sum_{i=1}^{n} g(y_i)}{\sum_{i=1}^{n} \pi_i(n)}
\]

where \( \pi_i(n) \) is the detection probability of identifying item \( y_i \) and is a function of seismic grid and fault properties. For example, we may be interested in the total quality of magnitude of the fault population. Let \( g(y_i) = y_i \), T, the total magnitude can then be estimated by

\[
T = \sum_{i=1}^{n} y_i / \pi_i(n)
\]

The number of faults in the population, N, can be estimated by

\[
N = \sum_{i=1}^{n} 1 / \pi_i(n)
\]

If we subdivide the n faults into m size classes according to their magnitudes, and estimate the number of faults in each class, the estimated numbers of faults in individual size classes represent an estimate of the empirical magnitude distribution of the fault population.

The variance of the estimate is given by the following expression:
\[ T = \sum_{i=1}^{n} h^2(y_i) \left[ (1 - \pi_i) / \pi_i^2(n) \right] \]  

(4)

If \( h(y_i) = y_i \), equation (4) gives the variance of the estimate of total magnitude, whereas \( h(y_i) = 1 \), equation (4) provides the variance for the estimated total number of faults.

ESTIMATION OF DETECTION PROBABILITY

The key issue in the application of the H-T estimator to the problem of estimation both the number and size distribution of undetected faults is the determination of the detection probability \( \pi_i(n) \). In a study of the application of H-T estimator to predicting the number and size distribution of prospects missed by seismic grid line search in petroleum exploration, Kaufman (1994) has shown a method of determining the detection probability that a specific prospect will be included in the n mapped prospects. The estimation problem in his study is very similar to the one in our study. As a fault trace in a reservoir horizon is a zone of deformation, the width of the zone changes from maximum at the point with largest displacement around the middle part of the fault and to zero at the two tip points. The projection of fault geometry on a reservoir horizon can therefore be treated as an ellipse with a large major to minor axes ratio. Figure 1 is an ideal geometric model of a normal fault (from Needham et al.). The method discussed by Kaufman for estimating the detecting probability of prospect can be applied to our problem directly with only a minor modification on the model geometry.

To compute the probability of detecting an anomaly of a fault by continuous seismic grid line search, we assume that an anomaly with boundary curve \( C \) lying in the x-y plane is detected if at least one grid line crosses its boundary curve. The model for the computation of the detecting probability from Kaufman (1994) is adopted in this study, and the following model is based on Kaufman’s probability model (1994):

A. A rectangular lattice with lattice sides of \( A \) and \( B \) is superposed on the x-y plane.
B. An anomaly is detected if at least one lattice (grid) line crosses its boundary curve.
C. The location of each anomaly is uniformly distributed on the plane, and orientations of the fault are uniform on \( (0,2\pi) \).
D. Anomaly positions are mutually independent.

This simple model can be easily modified to incorporate nonuniformly distributed anomaly positions and orientations of fault when additional geological information are available. A geometric model is shown in Figure 2. Draw a line \( L \) through any fixed point \( P \) interior to the anomaly that divides its projective area on the x-y plane into two regions. The line \( L \) is to be regarded as fixed in the anomaly. Suppose that \( L \) is at angle \( \theta \) with respect to the x-axis. Enclose the anomaly’s boundary in a box with sides parallel to x- and y-axes and with all four sides tangent to the boundary curve. The caliper length at angle \( \theta \) in the x-direction is the distance \( L_x(\theta) \) between lines parallel to the y-axis touching the boundary curve of the anomaly; \( L_y(\theta) \) is the anomaly’s caliper length at angle to the x-axis. Inscribe the caliper length rectangle and compute the ratio of the area of the rectangle within a grid rectangle in which the center of the caliper length rectangle must lie to avoid crossing a grid line. The probability that the anomaly is not detected by a grid line is \( \Pr(\text{anomaly not detected} \mid \#=(A,B)) = 1 - \frac{(A-L_y(\theta))(B-L_x(\theta))}{AB} \) if \( A \leq L_y(\theta) \) and \( B \geq L_x(\theta) \) and is zero otherwise. Consequently, given grid dimensions \#=(A,B),

\[
\Pr(\text{anomaly with boundary curve } C \text{ detected} \mid \# \text{ and } L \text{ at angle } \theta) = \begin{cases} 
1 - \frac{(A-L_y(\theta))(B-L_x(\theta))}{AB} & \text{if } A \leq L_y(\theta) \text{ and } B \geq L_x(\theta) \\
1 & \text{otherwise}
\end{cases}
\]  

(5)

Knowing the seismic grid line density \#=(A,B) along with an observation of mapped faults in the region of interest, the number and the size distribution of faults missed by the seismic grid line search can then be estimated by the H-T estimator.

APPLICATION EXAMPLES

We use computer simulated fault patterns to illustrate the application of the proposed method to the estimation of fault population properties. As suggested by many researchers, a power law relationship
exists between fault trace length \( L \) and the number of fault in a fault population. The relation is modeled as a Pareto distribution in our example with the following form:

\[
f(x) = \alpha \beta x^{-(\beta + 1)}
\]

(6)

where \( \alpha \) is a constant, \( \beta \) is the shape parameter of the Pareto model, \( x \) is the fault trace length. Based on the randomly generated probability values, we calculate the fault trace length using the following relation:

\[
x = \alpha (1 - F)^{-1/\beta}
\]

(7)

where \( F \) is the probability distribution function.

A fault trace length population is first simulated in our example using equation (7). The parameters used in the simulation are: \( \alpha = 30, \beta = 1.5 \), total number of faults \( N = 50 \), and a ratio of the major to minor axes of the fault ellipse is set at 150. We assume that the orientations and the positions of the faults are randomly distributed. Figure 3 shows one realization of the fault population. The total length of the fault population by this realization is 4177.1. If we chose a seismic grid density of 100x100 unit superposed on the \( x \)-\( y \) plane in which faults are located to observe the fault population, 32 of the 50 faults are mapped, and the other 18 faults are missing (Figure 4). The method discussed in previous section is used to estimate the detection probability that a fault with given length \( x_i \) being mapped. An estimate of total number of faults in the population by the H-T estimator is 50.8 with standard deviation of 6.5, and the
estimated total length of faults is 4208 with standard deviation of 24.8. Figure 5 shows the observations, estimated experimental distribution and the population distribution of the faults. Figure 6 shows the extrapolation of a straight-line fitted to the observations of fault that represent a completely sampled distribution segment for inferring the number and sizes of faults in the population.

The method can also provide other estimates rather than magnitude, as long as the estimated property has a single valued function of the magnitudes. For example, if the objective of the study is to estimate the total amount of extension due to normal faulting, using the fault trace length to calculate detection probabilities, replacing \( g(y_i) \) with the measurements on fault width, the total extension can be then estimated using equation (2).

Complex seismic grid line pattern rather than the one in figure 4 can be also handled by this method. To calculate the detection probabilities in complex circumstances, it is referred to Kaufman (1994) for methodology details.

**CONCLUDING REMARKS**

Although the power law relationship between fault length and fault number has been suggested by several researchers, different opinion exists (Wen, 1995). The H-T estimator is a nonparametric model, i.e., non pre-assumption on characteristics of the parent population is required. It is therefore applicable in the case that fault magnitude does not follow power law relation.

The detectability of a fault by seismic depends not only on the grid density, but also on the resolution of seismic survey. The faults present in nature have a wide spectrum of magnitudes from several millimeter to hundred kilometers, but only those with size above seismic resolution in the subsurface may be mapped. The prediction of small faults below seismic resolution remains a problem in proposed method. However the estimation of detection probability may provide a quantitative guide line for other method such as graphic fitting by Pareto mode in predicting the unmapped faults distribution.

**ACKNOWLEDGEMENT**

This research is partly supported by a grant from the Foundation of Distinguished Young Teacher of the China National Education Commission.

**REFERENCES**


