Overestimates of Earthquake Prediction Efficiency in a “Post-Prediction” State

Shozo Matsumura

National Research Institute for Earth Science and Disaster Prevention, Tsukuba 305, Japan

A basic procedure in earthquake prediction research is to find out possible relationship between precursory phenomena and earthquake occurrences. After the occurrence of strong earthquakes, we usually arrange observed data on a common temporal axis, and carefully examine if there has been some anomalous phenomena prior to the events. The result obtained from such procedure is so-called “post-prediction” state. It should be noted that “post-prediction” often states more than the truth, because, in the above procedure of finding possible precursors, it is necessary to introduce several free parameters, such as threshold levels, precursor time (duration or preceding time of precursor) and others, and they are liable to be intentionally selected to make the final result optimum. After all, such phenomenological handling will result in exaggeration of prediction efficiency compared with “true-prediction.”

The aim of this study is to clarify this situation, and to give a numerical estimation of it by taking a simplified model, where only precursor time $T_p$ is treated as a free parameter.

A model pattern of precursor and earthquake sequence is given in Fig. 1, where five precursors, or candidates of precursors, $p_1, \cdots, p_5$, precede occurrences of five earthquakes, $e_1, \cdots, e_5$. It seems that whoever sees this pattern will have an impression of good correlation between appearance of $p_i$ and occurrence of $e_i$. One way to judge whether or not the impression is true, is to execute a statistical test for a hypothesis that the pattern happens by accident. For this purpose, a probability $P$ that at least $n$ earthquakes among $e_1, \cdots, e_m$ accidentally fall within specified periods is used to be calculated, and compared with a level of statistical significance. For the example of Fig. 1, those periods are specified as shown in Fig. 2, where four among five earthquakes fall in the heavy-lined periods. In this case, the total duration of specified periods is $5T_p$, so $P = 5C_4p^4q + 5C_5p^5$, where $p = 5T_p/T_0 (= 0.25)$, and $q = 1 - p$. Now, the calculated value of $P$, 0.016, looks small enough to conclude that the hypothesis should be rejected. However, in this case, the value of $T_p$ has been intentionally selected to make $P$ minimum. In other words, $P$ is biased from a reasonable value. Then, how much value of $P$ should be recognized to be reasonable, or what value should be adopted for the significance test? In order to solve this problem, we have carried out a computer simulation as follows.

Two extreme cases are investigated, and compared. One is case-A, where the free parameter is intentionally selected in order to get an optimum solution. The other, case-B, corresponds to non-intentional selection. First, five $e_i$ (earthquakes) and five $p_i$
An example pattern presenting five earthquakes ($e_1, \cdots, e_5$), and five candidates of precursors ($p_1, \cdots, p_5$).

A free parameter $T_p$ (duration or preceding time of precursor) is brought into the pattern of Fig. 1. The value of $T_p$ is intentionally selected to get an optimum solution.

Accumulated appearance probability functions of $P$ for possible patterns composed of five earthquakes and five precursors obtained for both cases. The solid line corresponds to the case of intentional selection for $T_p$ (case-A). The dotted line corresponds to non-intentional selection (case-B). The broken line is proportional to $P^2$ given as a reference.

(candidates of precursor) are independently generated on a time axis of unit length. At each trial, $T_p$ is selected to make $P$ minimum in case-A, while $T_p$ is selected at random on a unit time length in case-B, and the values of $P$ are calculated for both cases. Finally after 100,000 trials, appearance probabilities of $P$ are obtained, and compared. The results are shown in Fig. 3, where accumulated probability functions $F(P)$ are drawn in a solid line for case-A, and in a dotted line for case-B. The latter approximately equals to a function of $P^2$. This can be intuitively understood because the appearance probability of $P$ must be proportional to itself, so that the integrated and normalized function becomes $P^2$. Here, let us re-evaluate the significance of the above-introduced
Fig. 4. An example of earthquake prediction reported by Keilis-Borok et al. (1988). Small arrows indicate occurrences of strong earthquakes ($M > 6.4$). Shaded and unshaded windows correspond to different thresholds for a special parameter. For a case of unshaded windows, all of the strong earthquakes occurred within the warning periods. This means that the prediction resulted in a perfect success.

Example. In case-A, that is, when $T_p$ has been intentionally selected, the calculated value $P=0.016$ corresponds to $F(P)=6.0\%$, which means that the hypothesis cannot be rejected under the test with 5\% level of significance. After all, the first impression should be overthrown. On the contrary, supposing that the parameter $T_p$ is a priori determined, that leads to an opposite conclusion because of very small value of $F(P)=0.09\%$ in case-B. In an actual case, the true value of $F(P)$ is considered to exist between both extreme cases as shown in Fig. 3. The value of $P$ itself lies between this extent as shown by the broken line; however, the object to be tested is not $P$, but $F(P)$.

This idea can be applied to an actual example. Keilis-Borok et al. (1988) have reported successful predictions as shown in Fig. 4, where 9 strong earthquakes ($M > 6.4$) in the Southern California region, and 5 strong earthquakes in the Northern California region could be recognized to occur within the warning periods. Values of $P$ are formally calculated to be 0.0028 (the specified period ratio $p$ is 0.52) for the Southern California data, and 0.0034 ($P=0.32$) for the Northern California. The similar procedures are carried out by taking nine earthquakes model for the Southern California, and five earthquakes model for the Northern California. Although the number of precursor appearance is not definitely specified in their paper, it has been confirmed that the final result is scarcely dependent on it. The true values of $F(P)$ are estimated to lie between 1.8\% (case-A)-0.02\% (case-B) for the Southern California case, and between 1.5\% (case-A)-0.01\% (case-B) for the Northern California. Thus, Keilis-Borok’s result could be guaranteed to be significantly successful even if it may be derived from a “post-prediction” state. However, this is not always the case. Generally speaking, result obtained from a “post-prediction” state should be carefully assessed because of its relatively low level in significance.

The author is grateful to Dr. Y. Okada for many valuable suggestions and comments.

REFERENCES


Vol. 41, No. 1, 1993