Response to reviews

Response to Review Comments by Reviewer #1:

We appreciate the reviewer's critical comments and constructive suggestions. In response we have substantially refine our manuscript to clarify the themes of our paper as well as to incorporate the reviewer's comments and suggestions. We hope the revised manuscript has improved both its context and readability.

In my previously study, from the maximum peak ground acceleration (PGA), maximum peak ground velocity (PGV), maximum Modified Mercalli intensity (MMI) of each grid (grid interval is 0.1^0), I find two regions with higher maximum PGA, maximum PGV, maximum MMI. One zone extends from Hsinchu southward to Taichung, Nantou, Chiayi, and Tainan in western Taiwan and the other extends from Ilan southward to Hualian and Taitung in eastern Taiwan. And the b value distribution of each grid, it shows the low b values paths to be the same as above described. According to Wyss and Stefansson (2006) future mainshocks can be expected along zones characterized by low *b* values. For verifying these results, I originally use the Shen et al. (2007, Seism. Res. Lett. 78, they combined GPS and earthquakes data) formula to combine earthquakes data and GPS data to do "**Forecasting the Probability of Future Earthquakes of $M_w \geq 6.0$ ",** but, their formula is very complex, therefore, I try to find my new method that combining the GPS data with Benioff strain (equation 16) to forecasting the probability of future earthquakes $M_w \geq 6.0$ of each grid, and my new method result is better than Shen et al.s' formula. Moreover, I simulate the probabilities of future earthquakes with $M_w \geq 6.0$ for next 10, 20, 30, 50 years, they show the same patterns as described above. Therefore, my new method is supported by theory, experiment, and empiric. These methods are my innovations. Every journal should encourage author to innovate. Therefore, I think the reviewer should support my paper to publish on NHESS.

Reviewer thought basic flaws :

1. Language and organization of the paper :

   (a) Thank you for your comments. I had done my best to edit my manuscript, and I had asked someone (He is an English editor of something geophysical journal) to help me to edit the manuscript before submitting the paper to NHESS. Or could you introduce an editor to me to help me to edit the draft?

   (b) As described above, I find two zones, one zone extends from Hsinchu southward to Taichung, Nantou, Chiayi, and Tainan in western Taiwan and the other extends from Ilan southward to Hualian and Taitung in eastern Taiwan to present low b values distribution, and according to Wyss and Stefansson (2006) future mainshocks can be expected along zones characterized by low b values. For verifying the phenomenon, therefore, I combine the earthquakes data (Benioff strain) and maximum shear strain rate (GPS data) to estimate the probability of future earthquake with $M_w \geq 6.0$. Following, I simulate the future earthquakes to confirm the results. Because, the procedure of simulating needs a , b the two parameters, hence, I must first to calculate the two parameters before simulating. I think the data process and procedure to be right, and not suddenly jump to other step for your comments. I should focus on my new methods and results in the abstract, therefore, it is impossible, I only describe the b value in the abstract.

(c) Thank you for your comments. I think the data process is my innovation. The probability of earthquake may be a well-known equation, but, did you have saw someone to simulate the times, magnitudes of future earthquakes with magnitude $M_w \geq 6.0$? And I use other method to calculate the a , b values of Gutenberg-Richter law, not using Gutenberg-Richter law to estimate a , b values.

(d) Thank you for your comments. It is my mistake. I miss the definitions of N and T in equation 6. The definitions of N and T are respectively below : The N is the number of earthquakes with magnitude $\geq$ M , and the T is the time period used data recorded.

(e) Thank you for your comments. The definitions of M_{oL} and M_o are different. The definitions of M_{oL} and M_o are respectively below : The M_{oL} is the minimum magnitude for which the seismic catalog is complete of each grid, and it has been defined in equation 1. And the M_o is the minimum magnitude used data in equation 8.

(f) Thank you for your comments. From equation 1 to equation 6 are time-independent quantities. And we begin to define the entropy, and we introduce a certain time (time interval, or time duration by your comments) in equation 8, Taking into account the expression of probability equation 1 and the b value (equation 5), we obtain straightforward entropy :

$H(t) = \log_{10}(e * \log_{10} e) - \log_{10} b = k' - \log_{10} b$, and $k' = \log_{10}(e * \log_{10} e) \approx 0.072$ (I have described them in the manuscript). We still maintain the dependence of H on time because we expect that when we estimate b in successive time window, we have $b = b(t)$, in other words, a time dependence of b. Another simple way write the relationship between entropy and the b value is the following : $b = \frac{b_{\max}}{10^H} \approx \frac{1.2}{10^H}$ (equation 10 in my manuscript), $b_{\max} = e * \log_{10} e \approx 1.2$. Therefore, the certain time (time interval or time duration) affects the b value not a value. It is against your comments (affect a value not b value).

(g) Thank you for your comments. It is my mistakes. $\log_{10}(N) = a - b\log_{10}(PGA)$, $\log_{10}(N) = a - b\log_{10}(PGV)$, and $\log_{10}(N) = a - bI$, the three attenuation formula are all to appear in my paper (Chen et al. 2010, BSSA). The N is the number of earthquakes with ground motion acceleration $\geq$ PGA at each grid for $\log_{10}(N) = a - b\log_{10}(PGA)$, the N is the number of earthquakes with ground motion velocity $\geq$ PGV at each grid for $\log_{10}(N) = a - b\log_{10}(PGV)$, and the N is the number of earthquakes with ground motion shaking $\geq I$ (Modified Mercalli Intensity) at each grid for $\log_{10}(N) = a - bI$. I use three different data types (PGA, PGV, and MMI) to see the seismicity tendency (b-value distribution), and the results are all the same. Following, I introduce the entropy and GPS data, I want to see variational trend between entropy, b-value distribution, and maximum shear strain rate (GPS data). I have described the entropy in my manuscript, and the GPS data is from Hsu et al. (2008), therefore, I do not re-describe the data, but , I have briefly described how to calculate the strain rate for GPS data and error assessment in the section " Forecasting the Probability of Future Earthquakes of $M_w \geq 6.0$ of my manuscript.

(h) Thank you for your comments. About equation 18. The definition of Benioff strain is the square root of every earthquake energy release. Therefore, the cumulative Benioff strain of each grid is defined as : $B_{xy} = \sum_{i=1}^{N} \sqrt{E_{xy}^{(i)}}$, the parameters have been explained in the manuscript, therefore, I think the equation 18 to be very easily comprehensible.

Science

Thank you for your comments. Combining the earthquakes (Benioff strain) and GPS data (maximum shear strain rate) to estimate the probability of future earthquakes with magnitude $M_w \geq 6.0$, simulating the time, magnitude, and probability of future earthquakes with magnitude $M_w \geq 6.0$, and calculating the relation between entropy of earthquakes and shear strain rate are all my innovation. I do not know why you say it does not have any reliable scientific conclusion.

(a) Thank you for your comments. I believe everyone to know it is very difficult for predicting future earthquakes, because, there are many parameters to affect the result, the time is a factor. If, using the Poissonian distribution to make specific time-dependent predictions of future earthquakes, and the result is good. In that way, researchers could fix the time factor, then, go a step further to modify the predicted model and add other parameters. In this way, researchers maybe have the chances to obtain the true predicting model. Therefore, I think that I use the random Poissonian distribution to make specific time-dependent predictions of future earthquakes is an innovation and all right. Otherwise, including many parameters to the predicting model at one time, they probably lead to bad result, and let researchers to step back, hence, researchers would not have the chances to obtain the true predicting model.

(b) Thank you for your comments. I think the objective of NHESS journal should encourage every author has a new method and a new idea. As described above, I original use the Shen et al. (2007, Seism. Res. Lett. 78, they combined GPS and earthquakes data) formula to combine earthquakes data and GPS data to do "**Forecasting the Probability of Future Earthquakes of $M_w \geq 6.0$ "**, but, their formula is very complex, therefore, I try to find my new method that combining the GPS data with Benioff strain (equation 16) to forecasting the probability of future earthquakes $M_w \geq 6.0$ of each grid, and my new method result is better than Shen et al. formula result. Moreover, I simulate the probabilities of future earthquakes with $M_w \geq 6.0$ for next 10, 20, 30, 50 years, they show the same patterns as described above. From the result of my new method, it indicates my new method is supported by theoretical, experimental, and empirical evidence. The definition of Benioff strain is the square root of each earthquake energy release. In equation 17, the $B_{\min}$ is the minimum Benioff strain of each grid, and is equal to the minimum energy release of each earthquake of each grid, B_{xy} is the cumulative Benioff strain of each grid, the xy is the index of somewhere grid, and the B is the Benioff strain that I want to calculate minimum earthquake magnitude, for example, $M_w \geq 6.0$, and the B is the Benioff strain of magnitude M_w =6.0.

(c) Thank you for your comments. Of course, I consider carefully the questions raised by reviewer. There are 430,529 earthquakes data in Taiwan during the time period from 1900 to 2008, and I plot the Gutenberg-Richter relation (I do not show the figure in the manuscript) that I find the slope abruptly change at about M_w =2.0,.thereafter, it is almost linear until about M_w =7.0, therefore, I say the Taiwan earthquakes catalog is complete for $M_w \geq 2.0$. I also do the same procedure for $M_w \geq 5.0$ (Chen et al., 2010, BSSA), therefore, the earthquake catalog with magnitude $M_w \geq 5.0$ is complete that is no doubt, then I use the 1989 crustal earthquakes and match with the strong ground acceleration attenuation relation and ground velocity attenuation relation to calculate the peak ground acceleration (PGA) and peak ground velocity (PGV) for each earthquake data, then combine the PGA and PGV to obtain the corresponding MMI, therefore, I think the result is O.K. and high quality.

(d) Thank you for your comments. It is my mistake. The acceleration and velocity attenuation laws are form Liu and Tsai (2005, BSSA).

(e) Thank you for your comments. The equation (12) $\overline{M} = \sum_{i=1}^{N} \frac{n_i M_i}{n_i}$ is to calculate the average magnitude for the linear part of Gutenberg-Richter relation for each grid , for example, 5 magnitude M_w =2.0, 3 magnitude M_w =3.0, 2 magnitude M_w =4.0, 1 magnitude M_w =5.0, before equation 12, the average magnitude is $\frac{2.0+5.0}{2} = 3.5$, but, I consider the average magnitude should be $\frac{5*2.0+3*3.0+2*4.0+1*5.0}{5+3+2+1} = 2.9$ and it is reasonable. M_{oL} is the minimum magnitude of linear part of Gutenberg-Richter relation of each grid for which the seismic catalog is complete, I have briefly described it in my manuscript, and it may be different for each grid (as example above, the M_{oL} is 2.0).

(f) Thank you for your comments. As described above, I have done the test for magnitude $M_w \geq 2.0$ and $M_w \geq 5.0$, and they are all complete, therefore, it is complete without doubt for time from 1900 to 2008. My study area is Taiwan island (E120^0-122^0, N21.9^0-25.3^0), and I used data to include Taiwan area (119^0-123^0, N21^0-26^0), therefore, the spatial is complete without doubt. My paper (Chen et al., 2010, BSSA) the grid interval is also $0.1^0 x 0.1^0$ for the Taiwan island (E120^0-122^0, N21.9^0-25.3^0), and the results and this study results are all good, hence, I think the spatial is complete without doubt for $0.1^0 x 0.1^0$ grid space.

(g) Thank you for your comments. As described above, in this study, I use the GPS data to be from Hsu et al. (2009, Tectonophysics), hence, I do not re-describe the data, I only describe how to calculate the maximum shear strain rate. During data processing calculation of the maximum shear strain rate may cause uncertainty. The source of this uncertainty comes from observed data. It is transferred to the velocity field and ultimately the shear strain rate; i.e., uncertainty is cumulative. In this study, the shear strain average uncertainty is 0.113 ± 0.0146, I have described it in my manuscript.

(h) Thank you for your comments. Yes. The relation between entropy of earthquakes and shear strain rate (Figure 5) exhibits a big scatter, the relation is Y(entropy)=(0.27 ± 0.077)X(shear strain rate)+(0.01 ± 0.02), I only show the average value in my manuscript: Y(entropy)=(0.27)X(shear strain rate)+0.01. I want use the relation to do related research at lack of seismic data or GPS data.

(i) Thank you for your comments. I think from Figure 4 has supported my statement, for example, from Figure 4(b), there are two zones to present low *b* values distribution, one zone extends from Hsinchu southward to Taichung, Nantou, Chiayi, and Tainan in western Taiwan and the other extends from Ilan southward to Hualian and Taitung in eastern Taiwan, and from Figure 4(c), the paths of high shear strain rate distributions are the same as Figure 4(b).

(j) Thank you for your comments. For increasing the data volume and confidence level for some seldom seismic active areas , therefore, I do not remove aftershocks data and foreshocks data. Probabilities are appropriate for forecasting earthquakes. Shen et al. (2007) described earthquake probability at a given spot x and magnitude M as:

$$P(x, M) = A(x)F(M)$$

where $A(x)$ is proportional to the maximum shear strain rate field, and

$$F(M) = 10^{-1.5\beta(M - M_{\min})} \exp\left[10^{1.5(M_{\min} - M_c)} - 10^{1.5(M - M_c)}\right]$$

Which is the tapered Gutenberg-Richter magnitude distribution. M_c is the corner magnitude, and β is the exponential falloff rate for the seismic moment distribution, which is complex than my formula $\overline{B} = \frac{B_{xy} - B_{\min}}{B - B_{\min}}$, and my new method result is exactly better than Shen et al. (2007) result.

(k) Thank you for your comments. There are my new methods and new ideas in this study. As you say, the NHESS is a high-quality journal, therefore, I think the primary goal of NHESS should to encourage every author to have new methods and new ideas.

Finally, we hope that above response has adequately addressed the issues raised by the reviewer. Once again we deeply appreciate the reviewer's thoughtful comments and suggestions. They have been very helpful to us in the revision of our paper.