

Interactive comment on “Combining earthquakes and GPS data to estimate the probability of future earthquakes with magnitude $M_w \geq 6.0$ ” by K.-P. Chen et al.

Anonymous Referee #1

Received and published: 3 January 2014

The paper claims to combine earthquake and GPS data to attempt a prediction of future $M > 6.0$ earthquakes in Taiwan. However, the paper fails completely to deliver any reliable result, as it contains a repetition of some older well-known elementary formulas, with extremely poor data processing, and the arbitrary introduction of a probability definition for future earthquakes (which the authors themselves call “naïve”), without any theoretical, experimental or even empirical support. The data handling and presentation is very poor, while the text is almost incomprehensible in many sections. It is clear that the paper cannot be published in any form, even with major revision, therefore I suggest that it is rejected. Some of the most important paper basic flaws are

C2183

mentioned in the following

Language and organization of the paper

The paper is written in very poor English. In many places I could hardly figure what the authors wanted to say. I suggest that an English speaking or a colleague who is proficient in English reviews any future paper versions that the authors may want to re-consider. The paper organization is another very poorly handled issue. The Abstract is a typical example: The authors mention b values in the first sentence, jump to Benioff strain and shear strains in the second sentence, return to b values in the 3rd sentence and suddenly discuss future earthquake probabilities in the 4th sentence, without any coherence and reasoning in the text. Several sentences of the Abstract are repeated in the Introduction section, without any references or justification, while elementary information and well-known relations are given in Section 2 (Probability of earthquakes and estimates of a and b values in the Gutenberg–Richter law). Even worse, while most equations in this Section 2 are well-known variants derived from the G-R relation, the authors miss to provide explanations for several equations (e.g. no explanation is provided for N and T in the a definition, eq. 6), while symbols change without any explanation e.g. M_{0L} in equation (1) changes to M_0 in equation (8). The time-independent quantities of equations (1) to (6) (well-known G-R results) become time-dependent in the entropy definition (equation 8), where the authors introduce “. . . a certain time t . . .” (page 5734, line 6) without any explanation. In fact t is not time but the simple time duration of the employed data sample, which affects the a value of the G-R relation and not b , however this subtlety is clearly beyond the authors analysis.

Section 3 (Data processing and interpretation) is another very poorly written section, where the authors start with a G-R description of predicted (from the observed catalogue) PGA/PGV/MMI values using an undefined attenuation formula and then proceed to seismicity (a and b values), GPS strain and entropy estimations over a grid, without any explanation of the actual data, their completeness, spatial window employed, strain rate error assessment, etc. Section 4 is even more poorly written, especially the

C2184

explanation of eq. (18), which is simply incomprehensible to the average reader. The paper contains additional weaknesses at the remaining section, but I will mostly focus on the fundamental scientific errors of these sections.

Science

This is the main reason that I recommend the paper rejection. The paper simply exhibits 3 major flaws, which render it practically useless for any reliable scientific conclusion:

a) The authors use the seismicity in various forms (a and b values of the G-R relation, entropy measures, Benioff-strain) from their catalogue to compute essentially the average seismicity level. This seismicity is Poissonian, e.g. random and memory-less, as is evident in their formulation (e.g. a scales with $\log T$ in eq. 6). This is further recognized by the authors in Section 5, where the elementary Poissonian probability relation (eq. 23) is presented. In the same section the authors make the unbelievable claim (immediately after eq. 23) that they can use the average estimates of a memory-less Poissonian distribution to make specific time-dependent predictions of future earthquakes, which they present in Figs. 7 and 8. I do not have the time and patience to educate the authors but this is equivalent to adding the mean interevent time of $M > 6.0$ events in a region to the date of the last similar ($M > 6.0$) event, in order to "predict" the expected occurrence time of the next $M > 6.0$ event, something fundamentally not possible for random-Poissonian processes. Time-independent seismicity measures simply cannot be used to perform specific time-dependent earthquake forecasts.

b) The fundamental equation (eq. 16) that they employ to "combine" seismicity information (expressed through Benioff-strain) and maximum shear strains (expressed through $A(x)$), is simply not supported by theoretical, experimental or any other even empirical evidence. Moreover their definition of "average" B in equation (17) is simply incomprehensible and not supported by similar evidence. In other words, the whole combinatory "prediction" concept is not only "naïve", as the authors themselves recognize in the last

C2185

line of page 5737 but most probably wrong. It is interesting to mention that eq.(16) is also mathematically wrong, as the provided probability definition is not properly scaled (e.g. in order to give 1 for all spatial and time scales, after integration)

c) The data processing is simply obscure or even wrong in many places and the results or terms are often misleading. For example in Section 3 and in the first line of page 5735 the authors make the unbelievable claim that the Taiwan catalogue is complete for $M > 2.0$ since 1897, when it is clear that this cannot be the case (I doubt that it is the case even today for small events in the sea regions of Taiwan!). No completeness analysis is presented for the catalogue, however, the authors use it without any such consideration (!) and compute synthetic PGA/PGV/MMI for incomplete data, which they process as if they are complete. They mention that for this computation an attenuation law was used, without providing these laws and their uncertainties. Even worse, within the same section they make computations, e.g. for average magnitudes using eq. (12), using a lower minimum magnitude $M_{0L} = 1.0$ without any explanation for this choice !!! No spatial and time completeness study is provided for the study area, clearly leading to erroneous results as such a low M completeness is simply not available almost anywhere in the world, even nowadays that modern digital networks exist. For this computation (as well as for all grid computations) they provide the grid step (0.1°) but not the spatial window employed. In the next section (Section 6) they compute strain rates without providing any insight on the original GPS data, their density, accuracy, errors, etc. For the derived strain rates they do not perform any error propagation analysis (which strongly and non-linearly affects strain rates) but vaguely mention a mean strain-rate error, without any explanation about its derivation. They make claims about spatial correlations but the only correlation they show is Fig.5, which exhibits a huge scatter and a very poor correlation, and for which they do not provide even the standard linear regression error estimates (standard error, linear correlation coefficient, t-test for linearity, etc.). In some cases (e.g. last lines of page 5740) they claim that a correlation exist between different quantities (e.g. strain rate and b values) without providing any figure with a simple correlation to support their statement.

C2186

d) The previous inconsistencies, flaws, etc. are also observed in many other aspects of the manuscript. For example the authors recognize that they should have declustered the catalogue (lines 10-12 in page 5741) but state that did not to it because this would result in too few data for some sub-areas of Taiwan. In other parts they suggest that the application of alternative approaches was not performed because "...their formulas is very complex, and the result is not good..." (line 20, page 5737). Mishandling data or methods because the results are not easy or convenient, certainly does not belong to a proper research paper.

In order to summarize, I think the paper is fundamentally flawed, contains a lot of arbitrary statements and hypothesis combined with well-known elementary seismicity equations, mishandles the data and their interpretation. In my opinion, the authors need to relook at their approach and adopt a much more solid and scientific approach to data processing, if they want similar work to be published in high-quality journals.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 1, 5729, 2013.