Does the Catalog of California Earthquakes, with Aftershocks Included, Contain Information about Future Large Earthquakes?

John Rundle[1,2,3,4]

[1] Department of Physics, University of California, Davis, CA USA
[2] Department of Earth and Planetary Sciences, University of California, Davis, CA USA
[3] Santa Fe Institute, Santa Fe, NM USA
[4] Visiting Professor, IRIDeS, Tohoku University, Sendai, Japan

The question of whether earthquake occurrence is random in time, or perhaps chaotic with order hidden in the chaos, is of major importance to the determination of risk from these events. It was shown many years ago that if aftershocks are removed from the earthquake catalogs, what remains are apparently events that occur at random time intervals, and therefore not predictable in time. In the present work, we enlist machine learning methods using Receiver Operating Characteristic (ROC) analysis. With these methods, probabilities of large events and their associated information value can be computed. Here information value is defined using Shannon Information Entropy, shown by Claude Shannon (Shannon, 1948) to define the surprise value of a communication such as a string of computer bits. Random messages can be shown to have high entropy, surprise value, or uncertainty, whereas low entropy is associated with reduced uncertainty and high reliability. An earthquake nowcast probability associated with reduced uncertainty and greater reliability is most desirable. Examples of the latter could be the statements that there is a 90% probability of a major earthquake within 3 years, or a 5% chance of a major earthquake within 1 year. Despite the random intervals between major earthquakes, we find that it is possible to make low uncertainty, high reliability statements on current hazard by the use of machine learning methods.