Review of ”Geosystemics and Earthquakes” by Angelo De Santis, G. Cianchini, R. Di Giovambattista, C. Abbattista, L. Alfonsi, L. Amoruso, M. Carbone, C. Cesaroni, G. De Franceschi, Anna De Santis, D. Marchetti, F. J. Pavòn-Carrasco, L. Perrone, A. Piscini, L. Spogli and F. Santoro

This paper claims to be a review of a field the authors call *Geosystemics*. However, it lacks what I would expect from a review: An unbiased introduction to all relevant works in the particular field and a sober evaluation of their quality and importance. Instead, the manuscript is an enthusiastic description of various works from the 'earthquake-prediction' community. This community is widely critized for not following standards in investigating earthquake predictions, as I will explain below. Furthermore, this so-called review is by no means a review as it focuses solely on the publications from the investigating authors and mainly ignores critical works. Last but not least, this manuscript is certainly not a review of what the authors call *Geosystemics* but rather an advertisment for the 'earthquake-prediction' community and the lead author (14 references point to papers of the lead author). I regret but I cannot recommend this manuscript for publication.

Besides the problem that this manuscript does not constitute a review, the topic itself is very problematic. Even though the authors mention the Collaboratory for the Study of Earthquake Predictability (CSEP) with a web link, they do not consider the work of CSEP at all and ignore the standards that CSEP has set for testing earthquake forecasts or predictions and that should be followed by any publication of earthquake precursors:

1. *Present a physical model that can explain the proposed precursor anomaly.*

   To some degree, the authors present physical ideas for the precursor phenomena discussed. However, several precursory phenomena are introduced without any physical reasoning why they should work. Particularly in light of a review paper about *Geosystemics*, this lack of descriptions remains puzzling. I am not arguing that there cannot be a precursory phenomenon that is not understood but may show some predictive power, however, the presented ideas have so far failed to show their predictive power in prospective and independent tests and, therefore, their simple description does not suffice.

2. *Exactly define the anomaly and describe how it can be observed.*

   The entire paper is filled with observations and their respective interpretations. However, the authors do not define any of the anomalies in a testable way. Particularly outstanding examples of such 'anomalies' are shown in Figures 10 and 11. This type of anomalies is typical in publications of the 'earthquake-prediction' community as they suggest an anomalous signal without defining this anomaly and showing that it appears only (or at least mostly) before large earthquakes. This type of evidence always bears the problem of selection bias and it is the duty of the authors to show their readers that they did not fall for selection bias. Without a clear and quantitative description of an anomaly, all these observations are basically meaningless as they stem from observations in hindsight and prone to selection bias. It is a characteristic of such studies that the search for anomalies is conducted after large earthquakes have been observed. Usually, the parameters of the applied metric are tweaked until the anomaly is found. Thus, the authors need to fully pre-specify what the metrics for each anomaly exactly are and then apply it to future data to see if the anomaly idea holds.

3. *Explain how a precursory information can be translated into a forecast and specify such a forecast in terms of probabilities for given space/time/magnitude windows.*

   This aspect is completely lacking in the manuscript. And it is also lacking in the presented methods. Only the PI method was ever tested independently in the framework of CSEP. The RTP method has at least published their forecasts but the results were not convincing. From all the other methods I have never seen a testable model or any collaboration with CSEP. Quite the contrary, the CSEP team has tried several times to initiate collaborations but without success. I assume the lack of a testable model may have been the reason behind this.

4. *Perform a test over some time that allows to evaluate the proposed precursor and its forecasting power.*

   This is the gold standard for seismic precursor studies. Even though many precursors have been proposed a long time ago, no such tests have ever been carried out for the methods described in this manuscript. The sole execption is the PI method for which a forecast was submitted in 2006 and has been tested for the past 10 years, revealing no particular forecasting power when compared to the other forecasts in the test [*Schorlemmer et al.*, 2010; *Strader et al.*, 2017]. However, some of the authors of the PI method have published their own evaluation of the tests by changing the definitions of the metrics that were agreed upon by all participants before the tests took place [*Lee et al.*, 2011]. Not surprisingly, the PI showed higher forecasting power in their evaluation. For most of the other methods, only investigation in hindsight after large earthquake took place which do not constitute systematic testing.

5. *Report on successful prediction, missed earthquakes, and false predictions.*

   Limiting the search to periods shortly before large observed earthquakes will never be sufficient to investigate the predictive power of the precursory phenomena. The authors need to expand their search systematically and report on all hits, misses, and false alarms.

I could certainly formulate more critical points about the presented methods and why their approach to percursory phenomena is not sufficiently scientific, however, I consider it out of the scope of this article. Nevertheless, even if the authors are not responsible for non-scientific work of others, they have the duty to report on the problems, in particular the scientific problems, of the presented and introduced methods. Without it, this so-called review remains a biased advertisement for the 'earthquake-prediction' community and does not provide the means for an unbiased understanding of precursory phenomena. In light of this, the notion of *Geosystemics* seems to be very limited to the authors' works.

References

Lee, Y.-T., D. L. Turcotte, J. R. Holliday, M. K. Sachs, J. B. Rundle, C.-C. Chen, and K. F. Tiampo, Results of the Regional Earthquake Likelihood Models (RELM) test of earthquake forecasts in California, *PNAS*, *108*(40), 16,533–16,538, doi:10.1073/pnas.1113481108, 2011.

Schorlemmer, D., J. D. Zechar, M. J. Werner, E. H. Field, D. D. Jackson, T. H. Jordan, and the RELM Working Group, First Results of the Regional Earthquake Likelihood Models Experiment, *Pure and Applied Geophysics*, doi:10.1007/s00024-010-0081-5, 2010.

Strader, A., M. Schneider, and D. Schorlemmer, Prospective and retrospective evaluation of five-year earthquake forecast models for California, *Geophys. J. Int.*, *211*(1), 239–251, doi:10.1093/gji/ggx268, 2017.