COMMENTARY

ethology   WILEY

Comment on "Potential short-term earthquake forecasting by farm animal monitoring" by Wikelski, Mueller, Scocco, Catorci, Desinov, Belyaev, Keim, Pohlmeier, Fechteler, and Mai

Gert Zöller[1] [ID] | **Sebastian Hainzl[2]** [ID] | **Frederik Tilmann[2,3]** [ID] | **Heiko Woith[2]** [ID] |
Torsten Dahm[2,4] [ID]

[1]Institute of Mathematics, University of Potsdam, Potsdam, Germany

[2]GFZ German Research Centre Potsdam, Potsdam, Germany

[3]Institute of Geological Sciences, Freie Universität Berlin, Potsdam, Germany

[4]Institute of Geosciences, University of Potsdam, Potsdam, Germany

Correspondence
Gert Zöller, Institute of Mathematics, University of Potsdam, Germany.
Email: zoeller@uni-potsdam.de

Editor: Wolfgang Goymann

[Corrections added on 13 November 2020, after first online publication: the reference in the abstract section has been corrected.]

Abstract

Based on an analysis of continuous monitoring of farm animal behavior in the region of the 2016 M6.6 Norcia earthquake in Italy, Wikelski et al., 2020 (Ethology, 126(9), 2020, 931) conclude that animal activity can be anticipated with subsequent seismic activity and that this finding might help to design a "short-term earthquake forecasting method." We show that this result is based on an incomplete analysis and misleading interpretations. Applying state-of-the-art methods of statistics, we demonstrate that the proposed anticipatory patterns cannot be distinguished from random patterns, and consequently, the observed anomalies in animal activity do not have any forecasting power.

KEYWORDS
animal behavior, earthquake precursor, error diagram, prediction, randomness, statistics

1 | INTRODUCTION

Earthquake forecasting is considered to be the "holy grail" in seismology. Many forecasting methods have been suggested over decades. Some of them are based on geophysical observations related to the preparatory process of an event, and others are less obviously associated with the physical properties of an earthquake. One of the most prominent examples of the latter is the forecast of large earthquakes based on anomalous animal behavior before the event (see Woith et al., 2018 and references therein). Many authors claim to have been successful in predicting single events. However, the rules defining a "successful prediction" are often ill defined.

Therefore, a proper evaluation of the predictive power of a proposed precursor should include at least the following pieces of information: (a) the number of successful predictions (earthquake with precursor), (b) the number of false alarms (precursor without earthquake), and (c) the number of failures-to-predict (earthquake without precursor). To determine these numbers, the alarm volume within time, space, and "strength" (e.g., magnitude range or ground motion range) has to be well defined by the prediction scheme. A flexible and easy tool for this purpose is the Molchan (or error) diagram (Molchan, 1990). Here, we apply this technique in order to study whether or not the anticipatory patterns between animal and seismic activity reported by Wikelski et al. (2020) (hereinafter referred to as WK2020) have significant forecasting skills. We restrict our analysis to a statistical evaluation of the forecasting power and refrain from commenting on the plausibility of such patterns or on the modeling technique used to generate the proposed precursor. In other words, we consider the time series of the proposed precursory signal without questioning its origin. For this analysis, we use the data, which have been provided online by WK2020.

2 | ANTICIPATORY PATTERNS BETWEEN ANIMAL AND SEISMIC ACTIVITY

Here, we briefly review the analysis steps in WK2020 that are relevant for our statistical analysis. More details can be found there.

Animal activity is detected in terms of "overall dynamic body acceleration" (ODBA) of tagged animals, which include three species (cows, dogs, and sheeps). These values are corrected for a daily cycle related to the usual daily animal activities. The final signal is the residual of these corrected values and a reference signal that is provided by a vector autoregressive data-based model (see e.g., Stock & Watson, 2001) that is supposed to account for the normal interactions of groups of animals and their reactions to ground motion. The observation period includes three intervals: T1, October to November 2016, T2, January to March 2017, and T3, March to April 2017. In the first two periods, the animals were situated in a stable, while they were roaming on a pasture in the third period. The seismic activity is represented as a time series of estimated peak ground acceleration (PGA) at the farm, which is calculated from the hypocentral distance and magnitude of each earthquake in the catalog. "Unusual" activity for both, animals and PGA, is defined by the condition that the corresponding value exceeds two standard deviations from the respective mean value. If an unusual PGA event is found, WK2020 look for an unusual animal activity event in a time window of 20 hr prior to the PGA event. If such an event is found, the time difference ("anticipation time") between both events is plotted versus the spatial distance between the location of the animal anomaly and the hypocenter of the associated earthquake ("hypocentral distance"). The anticipation of seismic activity by the animals is concluded from the observation of a "negative relationship" of anticipation time and hypocentral distance (figure 5 in WK2020) for periods T1 and T2, when animals were in the stable. In period T3, no such relationship is observed. Based on this, WK2020 suggest that only animals in a building might be sensitive to upcoming earthquakes. The interpretation of the distance–time relation is "that physical precursors of earthquakes diffuse slowly from the respective hypocenter" (caption to figure 5 in WK2020).

Finally, WK2020 propose that a monitoring system with groups of instrumented animals at different places will allow forecasts of location and time of future earthquakes by using triangulation.

2.1 | Statistical analysis

The data for our analysis are shown in the top row of Figure 1 and include the time series of "anomalous" animal activity (observed activity corrected by daily cycles and reference activity) and the seismic activity in terms of PGA, both taken directly from the supplementary material (dataset and code) of WK2020. Here, we only consider the temporal alarm volume, as the space alarm volume is simply the location of the farm, and the "strength" alarm volume is defined by WK2020 as predicted PGA values exceeding the threshold, determined as described above (different for each of the three periods considered).
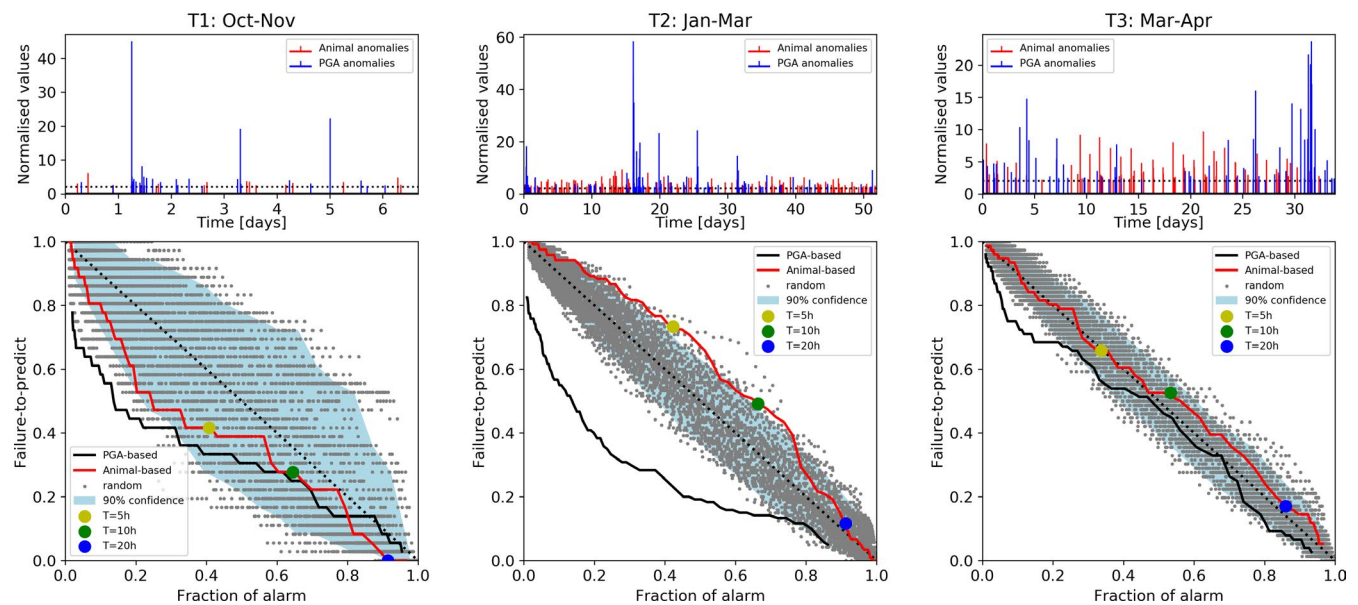

FIGURE 1 Time series of normalized animal anomalies and target PGA events (top) and Molchan diagrams (bottom) for the three different time periods (see titles). In the top panels, the black dashed horizontal line refers to the threshold of two standard deviations which is used to define times of anomalous animal behavior (red vertical lines) and PGA target values (blue vertical lines). In the bottom row, the red lines refer to the result for the observed animal anomalies using different alarm times T, where the results for specific values of T = 5, 10, and 20 hr are highlighted by colored points. For comparison, gray dots represent results for randomized occurrences of the animal anomalies, where the blue-shaded region defines their 90% confidence interval. The black line shows the result of forecasting based on the PGA anomalies alone (see text) [Colour figure can be viewed at wileyonlinelibrary.com]

It is important to note that WK2020 propose a short-term earthquake forecasting based on the anticipation of seismic activity by the animals. However, they only consider the case that unusual seismic activity is preceded by unusual animal activity in the preceding 20 hr. The cases of animal activity without ensuing seismic activity in this time window and of seismic activity without preceding animal activity are ignored. Even if the relationship between anticipation time and hypocentral distance for two associated events shows robustly some specific behavior, a forecasting scheme will only be reasonable, if also false alarms and failures-to-predict in terms of anticipation are studied in addition to successful predictions. More precisely, the distance–time relation is only beneficial in terms of a forecast, if the animal anomaly is indeed followed by a seismic event. For this reason, the key question will be whether or not there is a significant temporal association of animal anomalies and PGA anomalies. This association must be characterized by a high number of successful predictions and low numbers of false alarms and failures-to-predict.

The question raised above can be addressed by using Molchan (or error) diagrams (Han et al., 2017; Molchan, 1990; Molchan & Kagan, 1992): This is a two-dimensional plot with fraction of total alarm volume (here: fraction of cumulative anticipation time) on the horizontal axis and fraction of earthquakes without precursor (failure-to-predict) on the vertical axis. A rule for declaring an "alarm" with fixed length (e.g., 20 hr as in WK2020) results in a single point in the Molchan diagram. The variation of the alarm duration leads to a line connecting two end-member cases: first, the point (0,1) corresponding to the optimist's strategy (alarm time zero, all events are missed), and second, the pessimist's strategy at point (1,0) (permanent alarm, all events are "forecasted"). The diagonal line represents alarms based on random guesses. A scheme with high forecasting skill should be close to the origin: low fraction of alarm time and low number of missed events. The bottom row of Figure 1 shows the Molchan diagrams for the three periods T1, T2, and T3. The filled circles denote the anticipation times 5, 10, and 20 hr. For comparison, we repeated the same procedure for random perturbations of the time series of animal anomalies. The gray dots show the results for 100 randomizations evaluated for 100 different anticipation times. The results indicate for all periods that the association of animal and seismic activity is close to the diagonal line and many random signals perform better than the true signal of animal activity. Put differently, a complete analysis considering not only successes, but also false alarms and failures-to-predict demonstrates that the anticipation patterns reported in WK2020 can be explained as random patterns. The animal signal has thus no forecasting power. Specifically, we note that the 20-hr alarm window suggested by WK2020 is close to the right edge of the diagrams, implying the alarm of impending anomalous ground motion is nearly always on. Whereas this unsurprisingly "predicts" nearly all incidences of anomalous PGA events, such a nearly always-on alarm state would be useless in practice.

We further note that earthquakes are not uniformly distributed in time, that is, as a homogeneous Poisson process, but have a tendency to cluster due to the fact that some earthquakes trigger other earthquakes; commonly these triggered earthquakes are described as aftershocks. Whereas models of various sophistication have been developed to describe the statistical properties of this clustering (Cattania et al., 2018; Ogata, 1988), we consider here a naive forecasting scheme that simply raises an alarm when the PGA threshold is exceeded, that is, the PGA time series is used to predict its own future. The alarm is kept active for a set anticipation time. Again, a Molchan diagram can be constructed by adjusting the anticipation time. The results of this naive forecasting experiment are shown as a black line in Figure 1. The PGA-based forecast outperforms the animal-based forecast for (nearly) all configurations and time periods. For the T2 period, the naive forecast is actually significantly better than random. However, this is not because the algorithm shows any particular skill but is simply an artifact of the time clustering of earthquakes. Any proposed precursory phenomena must significantly outperform forecasts based on earthquake clustering, otherwise they are likely simply proxies for current seismic activity, even if they show some forecasting ability. The very different behavior of the naive forecast for the three time periods also demonstrates the variability of seismicity, highlighting the fact that much longer time series are needed for serious tests of forecasting schemes.

Finally, we consider the main argument of WK2020, that is, the apparent relation between anticipation time and hypocentral distance for cases where a seismic anomaly was preceded by an animal anomaly. The observation of a negative trend of the anticipation time with distance to the future earthquake has been interpreted as an indication of a diffusion process. However, the potential effect of the space–time clustering of earthquakes has been neglected in the calculation of significance level of the observed trends. Here, we reproduce the analysis by selecting all animal anomalies which preceded an anomalous PGA value within 20 hr and for all such pairs calculated the anticipation time and the hypocentral distance to the corresponding earthquake. Following the procedure in WK2020, we then fit a line to the data points. We finally repeat this calculation for 1,000 randomizations, where the same number of animal anomalies is randomly distributed in time, but the original PGA time series is used. For all three time periods, the observed values for slope and intercept lie within the cloud of slope–intercept pairs obtained from the randomized time series (Figure 2). For T1 and T3, the expectation value of the randomized relationship is near a zero slope (i.e., absence of a distance–anticipation time correlation), but the negative slopes in the observed dataset are not significant, as already noted by WK2020 themselves. At T2, the negative slope for the observed data has passed the significance test for non-zero slope applied by WK2020 but the expectation value of the randomized realizations is also negative due to the particular space–time distribution of earthquakes in this period. Thus, the observed trend seems to be solely related to the space–time clustering properties of the earthquakes and not related to the animal behavior.
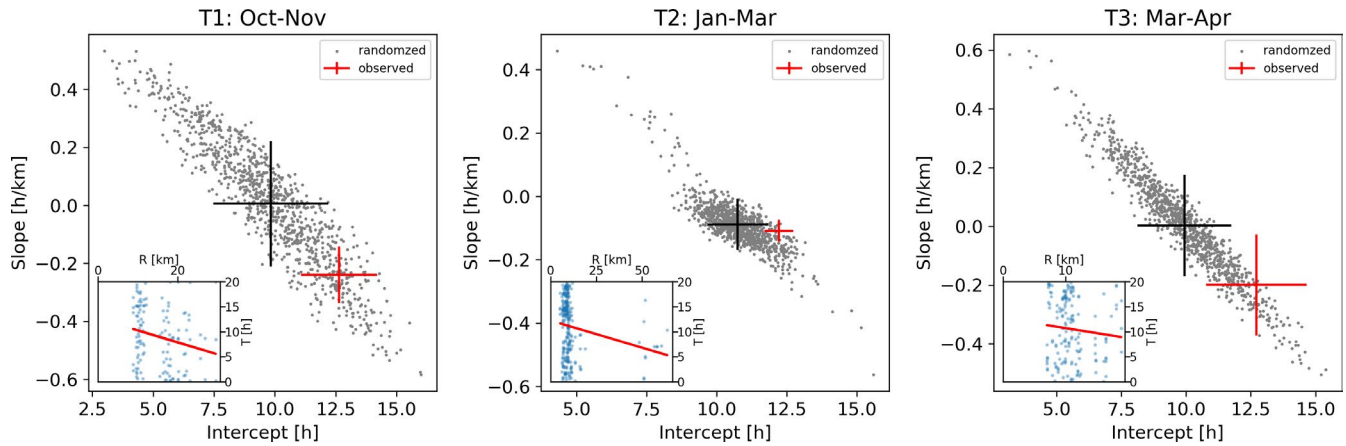
FIGURE 2 Results of the linear regression in terms of slope and intercept between the precursory times T of anomalous animal behavior and the hypocentral distance R to the respective earthquakes, where the three panels refer to the three analyzed time periods (see titles). Insets show scatter plots for the observed T-R pairs, with the red line representing the least-squares fit. In the main panels, the red cross shows the corresponding value of the slope and intercept of the fitted line, where the error bar represents plus/minus one standard deviation of the corresponding uncertainty. For comparison, the gray dots show the results of fits obtained for 1,000 randomizations of the times of the observed animal anomalies, where the black error bar represents their mean value plus/minus one standard deviation [Colour figure can be viewed at wileyonlinelibrary.com]

3 | DISCUSSION AND CONCLUSION

The conclusion of WK2020 that anomalous behavior of farm animals can be used for "potential short-term earthquake forecasting" is not justified by the results of their analysis. Their main argument lies in a relation between anticipation time and hypocentral distance for cases where a seismic anomaly was preceded by an animal anomaly. The other cases— animal anomaly without seismic anomaly and vice versa— had not been analyzed. In this comment, we perform a more complete analysis by constructing Molchan diagrams, which are well-established for the evaluation of earthquake forecasting schemes. The results clearly indicate that the anticipatory patterns are governed by purely random behavior. An earthquake forecasting scheme in terms of a yes/no statement on earthquake occurrence in a given time window based on animal anomalies will be as good as forecasts based on random guesses.

One might argue that the negative relation between anticipation time and hypocentral distance could have a certain value in terms of physical processes preceding an earthquake. However, the plots shown in figure 5 in WK2020 are hardly convincing to argue for the reality of such a relationship. Due to the apparent strong clustering of the earthquakes, the analysis is effectively based only on a small number of independent data, that is, the number of clusters. Again, a comparison with randomized sequences shows that the animal signal and random signals cannot be distinguished.

Summarizing, WK2020 presents an excellent data record for testing the ability of farm animals to predict earthquakes. However, their positive conclusion is based on a misleading analysis of this re- cord. Using standard methods for evaluating earthquake prediction

schemes shows that there is zero evidence that animals can predict earthquakes.

ORCID

Gert Zöller https://orcid.org/0000-0002-2384-2724
Sebastian Hainzl https://orcid.org/0000-0002-2875-0933
Frederik Tilmann https://orcid.org/0000-0002-7439-8782
Heiko Woith https://orcid.org/0000-0001-7662-7780
Torsten Dahm https://orcid.org/0000-0001-6432-7422

REFERENCES

Cattania, C., Werner, M. J., Marzocchi, W., Hainzl, S., Rhoades, D., Gerstenberger, M., Liukis, M., Savran, W., Christophersen, A., Helmstetter, A., Jimenez, A., Steacy, S., & Jordan, T. H. (2018). The forecasting skill of physics-based seismicity models during the 2010–2012 Canterbury, New Zealand, earthquake sequence. *Seismological Research Letters*, 89, 1238–1250. https://doi.org/10.1785/0220180033
Han, P., Kattori, H., Zhuang, J., Chen, C.-H., Liu, J.-Y., & Shuji, Y. (2017). Evaluation of ULF seismo-magnetic phenomena in Kakioka, Japan by using Molchan's error diagram. *Geophysical Journal International*, 208, 482–490. https://doi.org/10.1093/gji/ggw404

Molchan, G. M. (1990). Strategies in strong earthquake prediction. *Physics of the Earth and Planetary Interiors*, *61*, 84–98. https://doi.org/10.1016/0031-9201(90)90097-H

Molchan, G. M., & Kagan, Y. Y. (1992). Earthquake prediction and its optimization. *Journal of Geophysical Research*, *97*, 4823–4838. https://doi.org/10.1029/91JB03095

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of American Statistical Association*, *83*, 9–27. https://doi.org/10.1080/01621459.1988.10478560

Stock, J. H., & Watson, M. W. (2001). Vector Autoregressions. *Journal of Economic Perspectives*, *15*, 101–115. https://doi.org/10.1257/jep.15.4.101

Wikelski, M., Mueller, U., Scocco, P., Catorci, A., Desinov, L. V., Belyaev, M. Y., Keim, D., Pohlmeier, W., Fechteler, G., & Mai, P. M. (2020). Potential short-term earthquake forecasting by farm-animal monitoring. *Ethology*, *126*(9), 931–941. https://doi.org/10.1111/eth.13078

Woith, H., Petersen, G. M., Hainzl, S., & Dahm, T. (2018). Can animals predict earthquakes? *Bulletin of the Seismological Society of America*, *108*(3A), 1031–1045. https://doi.org/10.1785/0120170313