IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10642022

ESTIMATION OF EARTHQUAKE FORECASTING PROBABILITY BASED ON SATELLITE THERMAL ANOMALIES

Zhonghu Jiao[1], Xinjian Shan[1]

[1] State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing 100029, China

ABSTRACT

An approach is proposed for estimating the probability of earthquake forecasting based on pre-seismic thermal anomalies detected through satellite remote sensing products. A spatially self-adaptive scheme identifies optimal anomaly recognition criteria for refining calculated anomalies. Bayesian probability theory is then applied to calculate the conditional probability of earthquake occurrence given the observed anomaly. This probabilistic approach accounts for uncertainties associated with anomaly-based forecasting. Analysis of global seismic regions reveals that certain parameters, such as water vapor anomaly, exhibit superior performance in terms of higher probability gains. The adoption of the Bayesian method to incorporate satellite observations of thermal anomalies into operational earthquake forecasting systems could significantly enhance forecast capacity. The probabilistic framework addresses the limitations of anomaly correlation methods and provides more quantitative estimates of earthquake likelihood.

Index Terms— Earthquake forecast, Bayes' theorem, thermal anomaly, preseismic signal, thermal infrared remote sensing

1. INTRODUCTION

The development of an operational earthquake forecasting (OEF) system hinges on advancing beyond anomaly recognition to incorporate probabilistic forecasts of large earthquakes. Pre-seismic thermal anomalies, identified through various geophysical and geochemical parameters, have been explored as potential seismic precursors [1] . Despite researchers detecting anomalous signals before strong earthquakes, these findings face criticism, primarily due to the absence of distinctive characteristics or fingerprints that reliably predict large earthquakes [2]. Statistical correlations between pre-seismic anomalies and earthquakes often result in high false-alarm ratios, hampering the practicality of operational forecasting [3, 4]. Consequently, a shift toward a probabilistic approach gains traction, acknowledging the uncertainties associated with thermal anomalies.

This transition aligns with the broader trend of adopting Bayesian probability theory in earthquake forecasting, particularly in model comparison and spatiotemporal aftershock sequence clustering. The Bayesian approach proves effective in estimating conditional probabilities of earthquake occurrence, outperforming classical statistical testing methods [5, 6]. Notably, satellite data provides a robust foundation for statistical analysis, facilitating the determination of the probability of thermal anomalies. Embracing a probabilistic framework, especially through the Bayesian method, holds promise for enhancing the short-term forecasting capacity of OEF, addressing societal needs to mitigate the economic impact and casualties resulting from strong earthquakes.

2. STUDY AREA AND DATA

2.1. Global seismically active regions

Strong earthquakes consistently occur along major fault systems. An earthquake catalog from the United States Geological Survey (USGS) comprising 4,719 events with magnitudes of ≥ 6 and focal depths of ≤ 70 km from 1980 to 2020 is utilized to delineate global seismically active regions. The globe is subdivided into a 1°×1° grid, and earthquake counts are calculated within each grid. The grid is retained if the count exceeds 1. For grids containing more than one earthquake, a 5°×5° grid surrounding it is marked, and the same procedure is repeated to identify the most earthquake-prone regions globally. The 5° buffer zone is an empirical value used to indicate possible locations of future earthquakes. These areas are considered the study's focus, based on the reasonable assumption that historical seismic activity can provide insights into future global seismicity, with similar magnitudes and focal depths within adjacent regions.

2.2. AIRS product

The AIRS/Aqua L3 daily standard physical retrieval (AIRS3STD) V7.0 product is employed to calculate pre-seismic multiparametric anomalies based on nighttime data. This product has a spatial resolution of 1° and a temporal resolution of 12 h. The five geophysical parameters in

AIRS3STD are used, consisting of skin temperature (ST), air temperature (AT), total integrated column water vapor burden (CWV), outgoing longwave radiation flux (OLR), and clear-sky OLR (COLR).

3. METHODS

3.1. Pre-seismic anomaly recognition

The anomalous fluctuations of five geophysical parameters are calculated using a commonly used z-score (ZS) normalization method, represented as follows:

$$ZS(x,y,t) = \frac{v(x,y,t) - \mu(x,y)}{\delta(x,y)} \quad (1)$$

where $v(x,y,t)$ represents the current value of a given parameter (e.g., OLR) at a specific location (x,y) and time *t*. The average $\mu(x,y)$ and standard deviation $\delta(x,y)$ of the reference field are calculated at the same or similar time intervals over multiple years at the location (x,y).

However, these anomaly values alone are insufficient to associate with future earthquakes. To address this limitation, a spatially self-adaptive multiparametric anomaly identification scheme is employed to establish optimal anomaly recognition criteria for identifying anomalies in remotely sensed data [7]. To determine the optimal criteria at each 1°×1° pixel, an iterative process evaluates combinations of four criteria parameters within specified value ranges: 1) the temporal range for earthquake warning preceding the current day, 2) the absolute anomaly value threshold for identifying valid anomaly pixels, 3) the minimum percentage of anomaly pixels within a sliding spatial window, and 4) the minimum daily anomaly count within the specified temporal range. A 7°×7° sliding spatial window is utilized to extract the spatiotemporal anomaly data cube for the central pixel. A 3-fold cross-validation scheme is implemented, dividing the 2006-2020 data into training and test sets, to mitigate potential overfitting issues. The combination of parameter values that maximizes the Matthew's correlation coefficient is selected as the optimal criteria for that particular pixel.

This study is aimed to a specific case concerning forecasting practices. A fixed spatial and temporal domain is utilized—a 7°×7° spatial window to forecast earthquakes with a magnitude exceeding 6 for the upcoming 30 days. The period of observation used for forecasting can vary from 30 to 180 days [7], depending on the location and optimal parameters determined by the aforementioned procedures.

3.2. Earthquake forecasting probability

Bayesian theory is a statistical framework that offers a mathematical approach to reasoning and decision-making under uncertainty. Its fundamental principle lies in updating probabilities for a hypothesis based on new evidence or information available, combining prior knowledge with observed data to obtain a posterior probability. This iterative process revises existing predictions or theories in light of new evidence. The posterior probability for earthquake forecasting estimation based on Bayesian theory, given the recognized anomalies, can be calculated:

$$P(EQ|Ano) = \frac{P(EQ)P(Ano|EQ)}{P(Ano)} \quad (2)$$

where $P(EQ)$ is the natural probability of earthquake occurrences, also known as the a priori probability, representing our initial belief before considering the anomalies. $P(Ano|EQ)$ is the likelihood probability, representing the probability of observing the anomalies if an earthquake was to occur. $P(Ano)$ is the marginal probability of observing the anomalies, regardless of whether an earthquake occurs. Eq. (3) further expands the marginal term $P(Ano)$ into its constituent components:

$$P(Ano) = P(Ano|EQ)P(EQ) + P(Ano|\overline{EQ})P(\overline{EQ}) \quad (3)$$

where $P(\overline{EQ}) = 1 - P(EQ)$; and $P(Ano|\overline{EQ})$ is the conditional probability of anomaly occurrences without earthquakes. Therefore, estimating $P(EQ|Ano)$ requires determining three key components: $P(EQ)$, $P(Ano|EQ)$, and $P(Ano|\overline{EQ})$. This entails careful consideration of both prior knowledge and observed data within the Bayesian framework for robust earthquake forecast.

The probability gain (G_{EQ}) is defined as the ratio between $P(EQ|Ano)$ and $P(EQ)$, quantifying the increase in earthquake probability conferred by the observed anomaly [8, 9]. A robust earthquake forecast aims to simultaneously maximize success rate and minimize false alarms, ultimately maximizing G_{EQ} [10]. A G_{EQ} value exceeding 1 indicates that the observed anomaly likely carries precursory information within the seismogenic phase. Notably, higher G_{EQ} values of anomalies for geophysical parameters suggest the increased effectiveness of their P_{EQ} values in pinpointing impending large earthquakes.

4. RESULTS AND DISCUSSION

4.1. Global probability of earthquake forecasting

Figure 1 illustrates the global earthquake forecasting probability (P_{EQ}) based on OLR anomaly, including three components integral to the Bayesian formula for probability calculation. $P(ANO|EQ)$ consistently remains below 0.25, while $P(Ano|\overline{EQ})$ can exceed 0.6. This suggests that anomalous fluctuations are more prevalent in regions, as depicted by the yellow to red shading in Figure 1b. It can serve as an indicator for quantitatively evaluating the statistical correlation between pre-seismic anomalies and impending earthquakes: $P(ANO|EQ)$ should surpass $P(Ano|\overline{EQ})$ to enhance forecasting capacity.

The earthquake forecasting probability, $P(EQ|ANO)$, on December 31, 2020, is depicted in Figure 1c. Most regions exhibit $P(EQ|ANO)$ values below 0.2, and certain seismically active areas, such as Western China and Southeast Asia, however still manifest high probabilities. Figure 1d displays the *a priori* probability of earthquakes, with the circum-Pacific seismic belt showing significantly greater values, correlating with a higher likelihood of anomalous phenomena. Simultaneously, it is demonstrated that recognized anomalies convey pertinent information about earthquake preparation only when the probability gain (G_{EQ}) exceeds 1, as shown in Figure 1e, indicating a heightened chance of imminent earthquakes when anomalies occur.

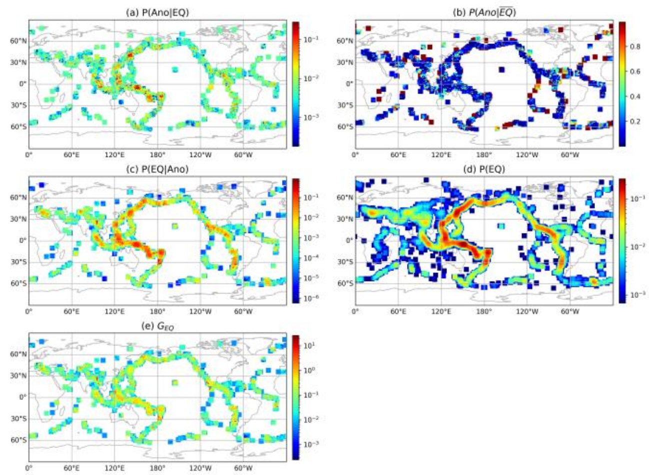

Figure 1 Component analysis of earthquake forecast probability for the OLR anomaly on December 31, 2020. (a) Probability of OLR anomaly occurrence before earthquakes; (b) Probability of OLR anomaly occurrence without earthquakes; (c) Posterior probability of earthquake occurrence given the OLR anomaly; (d) Natural probability of earthquake occurrence; (e) Probability gain from the OLR anomaly.

To further illuminate the spatial variability of calculated probabilities in Figure 1, we focus on the Mainland China and its vicinity (Figure 2). As expected, $P(Ano|EQ)$ consistently exceeds $P(Ano|\overline{EQ})$ in Figure 2a and b, corroborating our previous findings of a higher false alarm ratio in the study area [11]. However, $P(EQ|Ano)$ exhibits regional hotspots within the North-South and the Xinjiang Pamir to Tianshan North-South seismic belts (Figure 2c). Especially, the Taiwan region emerges as a hotspot with $P(EQ)$ values surpassing 0.15 (Figure 2d). Figure 2e reveals several regions within the high-probability areas of Figure 2c where G_{EQ} exceeds 1. This signifies that OLR anomalies in these regions effectively enhance the forecast capacity for imminent strong earthquakes.

Overall, these findings demonstrate a significant benefit in utilizing OLR anomalies for earthquake forecast. While phenomenological occurrence alone does not constitute definitive proof of precursory behavior, the observed spatial correlation between P_{EQ} and seismically active regions, coupled with the high G_{EQ} values in specific areas, provides compelling evidence supporting the feasibility and potential of thermal anomaly for global earthquake forecast.

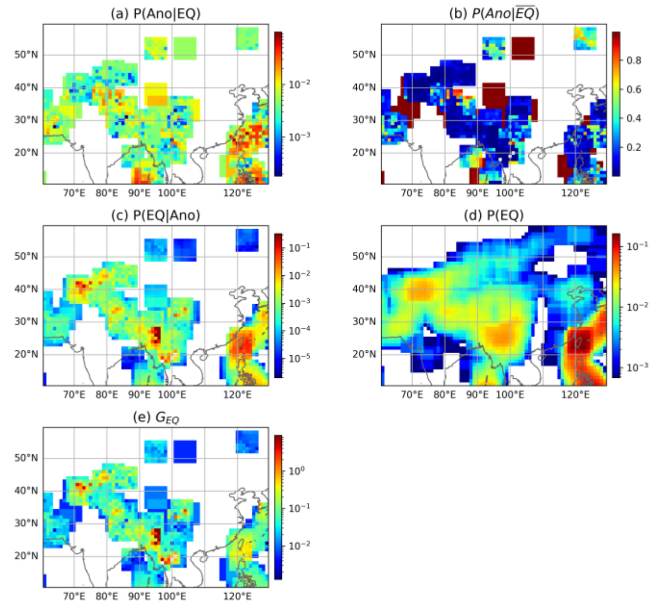

Figure 2 Component analysis of earthquake forecast probability for the OLR anomaly on December 31, 2020, focused on Mainland China and adjacent regions (similar to Figure 1).

4.2. Probability gain of each parameter

A valid earthquake forecast hypothesis identifies spatiotemporal sub-volumes with statistically elevated probability of strong earthquakes, demonstrably exceeding baseline levels. A valid hypothesis, in the context of earthquake forecast, exhibits G_{EQ} values consistently exceeding 1. Figure 3 presents the histograms of each parameter for the entirety of the seismically active regions. The statistical distributions exhibit remarkable similarity across all parameters, suggesting that each can potentially generate valid forecasts in some regions despite observed spatial variability. In particular, the CWV anomaly has the best performance, with a record of 304 valid pixels, a maximum G_{EQ} of 27.01, and a mean G_{EQ} of 2.95. This could be explained by the hypothesis that water vapor condensation plays a crucial role in generating atmospheric thermal anomalies [12], with anomalous CWV fluctuations potentially reflecting this intermediary process. In contrast, the AT anomaly exhibits the lowest G_{EQ} of 2.29, with such values indicating a modest risk of large earthquakes [13]. These observations collectively demonstrate the potential of thermal anomalies in earthquake forecast, irrespective of the specific parameter. Importantly, the results highlight that while every parameter can contribute precursory information about upcoming earthquakes, no single parameter emerges as definitively superior. This underscores the potential value of a multiparametric approach for earthquake forecast, combining non-seismological (e.g., land and atmospheric parameters used in this study) and seismological parameters [14].

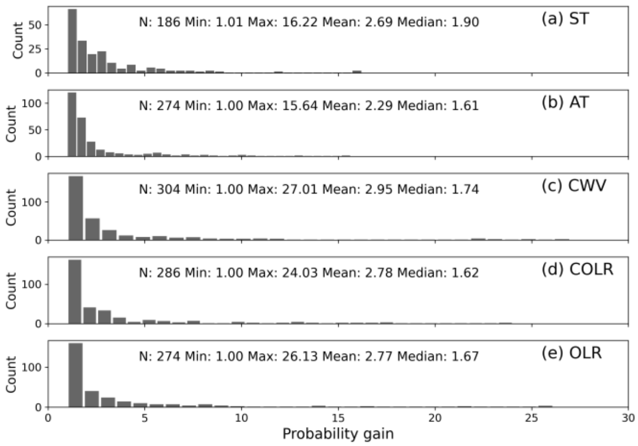

Figure 3 Histogram of probability gains of anomalies derived from (a) ST, (b) AT, (c) CWV, (d) COLR, and (e) OLR on December 31, 2020.

5. CONCLUSIONS

This study highlights the potential of a probabilistic framework to improve OEF using satellite observations of thermal anomalies. Bayesian theory facilitates the estimation of conditional probabilities for imminent large earthquakes based on detected precursor signals. The analysis reveals variability in the reliability of precursors, with water vapor anomalies demonstrating the highest probability gains. Despite its limitations, the probabilistic approach contributes to addressing false alarms and uncertainties inherent in anomaly correlation methods.

The shift from anomaly recognition to probabilistic forecasting aligns with contemporary perspectives on earthquake predictability. Through the combination of thermal infrared data and Bayesian statistics, this methodology can enhance existing forecasting models by providing improved situational awareness of seismic hazard likelihood. Further validation of reliable precursors in diverse tectonic environments will aid in refining this probabilistic forecasting approach. Overall, the framework illustrates a pathway for enhancing OEF through the probabilistic integration of satellite observations and statistical modeling. What's more, it is not limited to the studied parameters from AIRS3STD product but can also include other geophysical and geochemical parameters, such as greenhouse gases, total electron content, and electric and magnetic fields.

6. REFERENCES

[1] Z.-H. Jiao, J. Zhao, and X. Shan, "Pre-seismic anomalies from optical satellite observations: a review," *Natural Hazards and Earth System Sciences,* vol. 18, no. 4, pp. 1013-1036, 2018.

[2] H. Chen, S. Qin, L. Xue *et al.*, "Universal precursor seismicity pattern before locked-segment rupture and evolutionary rule for landmark earthquakes," *Geoscience Frontiers,* vol. 13, no. 3, pp. 101314, 2022.

[3] C.-C. Fu, L.-C. Lee, D. Ouzounov, and J.-C. Jan, "Earth's Outgoing Longwave Radiation Variability Prior to M ⩾6.0 Earthquakes in the Taiwan Area During 2009 – 2019," *Frontiers in Earth Science,* vol. 8, pp. 364, 2020.

[4] E. Pavlidou, M. van der Meijde, H. van der Werff, and C. Hecker, "Time Series Analysis of Land Surface Temperatures in 20 Earthquake Cases Worldwide," *Remote Sensing,* vol. 11, no. 1, pp. 61, 2018.

[5] W. Marzocchi, J. D. Zechar, and T. H. Jordan, "Bayesian Forecast Evaluation and Ensemble Earthquake Forecasting," *Bulletin of the Seismological Society of America,* vol. 102, no. 6, pp. 2574-2584, 2012.

[6] A. Darzi, B. Halldorsson, B. Hrafnkelsson, and K. S. Vogfjörð, "Short-term Bayesian ETAS spatiotemporal forecasting of the Ölfus 2008 earthquake sequence in Iceland," *Tectonophysics,* vol. 839, pp. 229522, 2022.

[7] Z. Jiao, Y. Hao, and X. Shan, "A Spatially Self-Adaptive Multiparametric Anomaly Identification Scheme Based on Global Strong Earthquakes," *Remote Sensing,* vol. 15, no. 15, pp. 3803, 2023.

[8] C. Fidani, "West Pacific Earthquake Forecasting Using NOAA Electron Bursts With Independent L-Shells and Ground-Based Magnetic Correlations," *Frontiers in Earth Science,* vol. 9, pp. 673105, 2021.

[9] J. D. Zechar, and T. H. Jordan, "Testing alarm-based earthquake predictions," *Geophysical Journal International,* vol. 172, no. 2, pp. 715-724, 2008.

[10] R. Console, "Testing earthquake forecast hypotheses," *Tectonophysics,* vol. 338, no. 3, pp. 261-268, 2001.

[11] Z.-H. Jiao, and X. Shan, "Statistical framework for the evaluation of earthquake forecasting: a case study based on satellite surface temperature anomalies," *Journal of Asian Earth Sciences,* vol. 211, pp. 104710, 2021.

[12] S. Pulinets, and D. Ouzounov, "Lithosphere–Atmosphere–Ionosphere Coupling (LAIC) model – An unified concept for earthquake precursors validation," *Journal of Asian Earth Sciences,* vol. 41, no. 4–5, pp. 371-382, 2011.

[13] K. Katsumata, and M. Nakatani, "Testing the seismic quiescence hypothesis through retrospective trials of alarm-based earthquake prediction in the Kurile–Japan subduction zone," *Earth, Planets and Space,* vol. 73, no. 1, pp. 100, 2021.

[14] N. Genzano, C. Filizzola, K. Hattori *et al.*, "Statistical Correlation Analysis Between Thermal Infrared Anomalies Observed From MTSATs and Large Earthquakes Occurred in Japan (2005–2015)," *Journal of Geophysical Research: Solid Earth,* vol. 126, no. 2, pp. e2020JB020108, 2021.