Journal of the Meteorological Society of Japan

EARLY ONLINE RELEASE

This is a PDF of a manuscript that has been peer-reviewed and accepted for publication. As the article has not yet been formatted, copy edited or proofread, the final published version may be different from the early online release.

The DOI for this manuscript is
DOI:10.2151/jmsj.2025-025
J-STAGE Advance published date: April 24, 2025
The final manuscript after publication will replace the preliminary version at the above DOI once it is available.

Does the performance of a flood early warning system affect casualties and economic losses? Empirical analysis using open data from the 2018 Japan Floods

Hitomu KOTANI

Department of Civil and Environmental Engineering, School of Environment and Society, Institute of Science Tokyo, Tokyo, Japan

Wataru OGAWA

Department of Urban Management, Graduate School of Engineering, Kyoto University, Kyoto, Japan

and

Kakuya MATSUSHIMA

Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan

April 16, 2025

Corresponding author: Hitomu Kotani, Department of Civil and Environmental

Abstract

Flood early warning systems are crucial for mitigating flood damage; however, limitations in forecasting technology lead to false alarms and missed events in warnings. Repeated occurrences of these issues may cause people to hesitate to take appropriate action during subsequent warnings, potentially exacerbating flood damage. However, the effects of warning performance on flood damage in Japan have not been analyzed for actual flood events. This study empirically examined these effects by applying Bayesian regression analyses to open data on the 2018 Japan Floods in 127 municipalities in four prefectures (i.e., Okayama, Hiroshima, Ehime, and Fukuoka) for which data were available on the real-time flood warning map (*Kouzui Kikikuru* in Japanese) during the 2018 Japan Floods, which provides limited open data on warning performance. Based on these data, the false alarm ratio (FAR) and missed event ratio (MER) for each municipality before the 2018 Japan Floods were calculated and used as explanatory variables. The (1) fatalities, (2) injuries, (3) economic losses to general assets, and (4) economic losses to crops during the 2018 Japan Floods were used as outcome variables. The results indicate that a higher FAR was associated with an increase in fatalities, injuries, and economic losses to general assets.

Engineering, School of Environment and Society, Institute of Science Tokyo, 2-12-1, Ookayama, Meguro, Tokyo 152-8550, Japan.
E-mail: kotani.h.15c7@m.isct.ac.jp

By contrast, no prominent positive effect of MER was found for any outcome variable. Although our results are fundamental, they provide valuable insights for improving warning systems and guiding future research.

Keywords False alarms; missed events; regression analyses; disaster statistics; public response

1. Introduction

Weather forecasts and warnings offer promising solutions for reducing weather-, climate-, and water-related disaster damage (Rogers and Tsirkunov 2011; Hallegatte 2012). Scientific and technological developments have increased weather forecast skills over the past 40 years (Bauer, Thorpe & Brunet, 2015). Accurate forecasts are expected to save lives, support emergency management, mitigate impacts, and prevent economic losses due to high-impact weather conditions. With human-induced climate change leading to more extreme weather conditions, the need for early warning systems (EWS) has become increasingly crucial (World Meteorological Organization, 2022).

However, owing to the limitations of scientific knowledge, observation technology, and models, forecasts and warnings are not always accurate (Trainor et al. 2015; Bauer et al. 2015), which can lead to public complacency and undermine the effectiveness of an EWS. The performance of these systems is often measured using the false alarm ratio (FAR) and the missed event ratio (MER). False alarms refer to events that were forecasted to occur but did not (Table 1), and FAR is calculated as the number of false alarms divided by the total number of events forecasted (Trainor et al. 2015; Lim et al. 2019). Similarly, missed events and MER were calculated based on events that were not forecasted but did occur. A well-known consequence of poor warning performance is the "cry wolf effect" or "false alarm effect" (Roulston and Smith 2004; Simmons and Sutter 2009; Trainor et al. 2015; Lim et al. 2019; LeClerc and Joslyn 2015; Sawada et al. 2022). In this phenomenon, people distrust subsequent warnings and hesitate to respond because of their prior experience with false alarms. Improving forecasting and warning performance is expected to reduce the abovementioned complacency of the public, encourage protective actions, and mitigate human and property losses. Table 1

In Japan, the performance of forecasts and warnings has been improving. For example, in July 2017, the Japan Meteorological Agency (JMA) introduced a surface rainfall index and a refined basin rainfall index into criteria for issuing flood warnings (Ota 2019). Through these efforts, the success ratio (SR)[1] and probability of detection (POD)[2] of flood warnings improved from 17% and 80%, respectively, in 2012 to 41% and 95%, respectively, in 2017. Such improvements are expected to increase the trust

[1]SR is calculated as the number of hits divided by the total number of events forecasted (NOAA ; Japan Meteorological Agency e).

[2]POD is calculated as the number of hits divided by the total number of events that occurred (NOAA ; Japan Meteorological Agency e).

of local governments and residents in warnings, leading to a more accurate issuance of evacuation information by local governments and the promotion of proactive evacuation by residents (Ota 2019).

Does flood early warning system (FEWS) performance affect flood damage in Japan? We aimed to answer this question; however, this is challenging because there are almost no open data on the history of warning hits or misses in Japan, which makes it difficult to calculate FAR and MER[3]. However, exceptionally, data on the SR and POD of the "real-time flood warning map" (*Kouzui Keihou no Kikendo Bunpu* or *Kouzui Kikikuru* in Japanese) during the heavy rainfall in western Japan in 2018—the 2018 Japan Floods[4]—are presented in a technical document by the JMA (Ota 2019). The real-time flood warning map highlights the escalating risk of flood disasters in small- and medium-sized rivers owing to heavy rainfall, color-coded at five levels (Japan Meteorological Agency a). The risk level is determined using the predicted value of the basin rainfall index (up to three hours in advance), and whether the risk level is increasing due to the rapid rise in water level—characteristics of small- and medium-sized rivers—can be assessed in advance (Japan Meteorological Agency a). Based on these SR

[3]This is probably one of the main reasons why empirical studies in real-world contexts are scarce compared to theoretical studies (Sawada et al. 2022; Kotani et al. 2024).

[4]It is identified by the Global IDEntifier (GLIDE) number FL-2018-000082-JPN, available at `https://glidenumber.net/glide/public/search/search.jsp`.

and POD data, we made certain assumptions and calculated the FAR and MER of flood warnings prior to the 2018 Japan Floods. We then focused on the consequences of people's failure to take protective actions—human losses (i.e., the number of fatalities and injuries) and property losses (i.e., the number of economic losses)—during the 2018 Japan Floods in municipalities where flood warnings were issued. Using disaster statistical data on human and property damage, we empirically analyzed the relationship between pre-disaster warning performance and flood damage.

The present study's findings underscore the social value of FEWS and provide insights for designing a more effective FEWS. Revealing the effects of the performance of FEWS—FAR and MER—on flood damage could help demonstrate the social significance of improving warning performance. Additionally, identifying the performance indicators that can be improved to reduce particular types of damage can guide the development of more socially beneficial technologies and systems.

2. Literature Review

2.1 The effect of performance of EWS in the United States

Past research has empirically studied the relationship between warning performance, people's protective actions, and the resulting disaster damage, especially in the context of tornado warnings in the United States (U.S.). For example, Simmons and Sutter (2009) conducted a statistical analysis of the relationship between the FAR in tornado warnings and human casualties caused by tornadoes (Simmons and Sutter 2009). Regression analyses were conducted on over 20,000 tornadoes that occurred in the continental U.S. between 1986 and 2004, using the tornado warning FAR as the explanatory variable and the number of tornado fatalities and injuries as the outcome variables. The results showed that the number of fatalities and injuries from tornadoes was significantly higher in areas with a higher FAR.

The process by which warning performance influences protective actions, which may result in tornado damage, has also been explored. Ripberger et al. (2015) focused not only on FAR but also on MER, and examined their effects on people's perceptions of tornado warnings and trust in the agency responsible for issuing tornado warnings by conducting an online survey of residents in tornado-prone areas in the U.S. (Ripberger et al. 2015). The results indicate that residents in areas with higher actual FAR and MER perceived higher FAR and MER, respectively. The results also indicated that residents with higher perceived FAR and MER had less trust in the National Weather Service (NWS), the agency responsible for issuing tornado warnings, and respondents with less trust in the NWS were less willing to take action in response to future warnings. This suggests that residents in areas with higher actual FAR and MER may be less likely to take protective action in response to future warnings.

Trainor et al. (2015) analyzed the relationship between actual and perceived FAR and their effects on actual protective actions during tornado warnings (Trainor et al. 2015). The results of the analysis of data collected through telephone interviews with residents indicated that actual FAR had no significant effect on residents' perceived FAR, whereas actual FAR had a significant negative effect on taking protective actions (e.g., evacuation, information gathering, and property protection). This suggests that residents in areas with high actual FAR may be less likely to take protective action in response to warnings, even though they are not aware of the actual FAR.

In contrast, Lim et al. (2019) reported different findings (Lim et al. 2019). Their analysis of survey data from residents in the southeastern U.S., where most tornado fatalities occur in the country, found no significant correlation between actual and perceived FAR, and actual FAR did not significantly affect protective actions. However, residents with a higher perceived FAR were more likely to take actions such as taking shelter when a warning was issued.

Overall, while previous studies reported mixed results, they consistently analyzed how the performance of warnings—actual FAR and MER—affects protective actions and the resulting damage, considering factors such as public perception of and trust in warnings. However, these findings for tornadoes in the U.S. may not necessarily apply to floods in Japan given the differences in disaster characteristics and false alarm frequencies. For example, the FAR for tornado warnings in the U.S. was approximately 75% (Simmons and Sutter 2009), whereas the FAR for flood warnings in Japan was 59% in 2018 (Ota 2019). The effects of warning performance on protective actions may vary depending on the frequency of false alarms, hazard types, and disaster impacts.

2.2 The effect of performance of EWS in Japan

Studies of the effects of warnings and evacuation advisory performance on protective actions and disaster damage in Japan are limited. For example, Yoshii et al. (2008) and Kaziya et al. (2018) conducted questionnaire surveys and interviews with residents for whom tsunami warnings and evacuation advisories/instructions for landslides had been issued multiple times over a certain period (Yoshii et al. 2008; Kaziya et al. 2018). These studies qualitatively pointed out that one reason why residents did not evacuate when a relevant warning or evacuation advisory/instruction was subsequently issued was the perception of previous warnings or advisories/instructions as false alarms. In addition, Katada and Murasawa (2009), who conducted a questionnaire survey among residents who received a tsunami warning following the 2006 Kuril Islands earthquake, found that even a single false alarm could reduce the intention to evacuate during future earthquake-induced tsunamis (Katada and Murasawa 2009).

However, few statistical studies have been conducted. Okumura et al. (2001) defined the subjective reliance on evacuation warnings as the probability that residents will suffer damage after receiving an evacuation advisory (Okumura et al. 2001). A questionnaire survey was conducted on the level of willingness to take evacuation action (evacuating immediately, preparing for evacuation, staying at home, etc.) of residents affected by the landslide disaster of the 1999 Hiroshima torrential rainfall under hypothetical disaster information provision. The results showed that the subjective probability significantly decreased when the evacuation advisory was a false alarm but increased when the advisory was a hit or missed event. Furthermore, it was shown that residents with higher subjective probability were more willing to evacuate. Therefore, it was suggested that false alarms reduce the subjective probability and, consequently, make residents less likely to evacuate.

Oikawa and Katada (2016) conducted experiments on warning strategies and people's protective actions (Oikawa and Katada 2016). Based on the basic policy of "issuing evacuation advisories as early as possible without considering false alarms" (the guidelines for evacuation advisories issued by the Cabinet Office in 2014), they conducted an experiment to test the effects of two types of warning strategies on the decision to evacuate: (1) a low-frequency strategy prioritizing the avoidance of false alarms, and (2) a high-frequency strategy prioritizing the avoidance of missed events. The results showed that, in the short term, the high-frequency strategy increased evacuation rates, whereas the low-frequency strategy decreased them. However, in the long term, the effectiveness of both strategies was diminished, and the absence of an evacuation advisory in the high-frequency strategy significantly influenced the decision to not evacuate. The authors concluded that while high-frequency strategies might be effective in the short term, their long-term significance is limited.

However, these studies were conducted under hypothetical or experimental conditions targeting evacuation advisory, and their findings have not been empirically validated in actual disaster scenarios. To the best of our knowledge, no empirical analyses have explored the relationship between weather warning performance and actual protective actions or the resulting damage in Japan.

This study contributes to the literature by focusing on flood warnings in Japan and statistically analyzing how their performance affects actual flood damage. Building on Simmons and Sutter (2009), we performed regression analyses using warning performance as the explanatory variable and flood damage as the outcome variable. For the flood warning performance and flood damage data, we utilized the open data described in Section 3. Unlike Simmons and Sutter (2009), who considered only FAR, we included MER, drawing on the approaches of Ripberger et al. (2015) and Okumura et al. (2001). Additionally, whereas Simmons and Suter (2009) primarily focused on human casualties, which are linked to protective actions such as evacuation, we considered a broader range of damage, including economic losses to general assets and crops. These property losses can be mitigated through protective actions such as using sandbags and waterproof boards to protect land and houses from flooding, as well as moving assets (e.g., vehicles) to higher ground before flooding occurs.

3. Data

3.1 Target flood and municipalities

This study focuses on the damage caused by the 2018 Japan Floods, for which the SR and POD of a real-time flood warning map were published by Ota (2019). During the 2018 Japan Floods, river overflows and mudslides occurred simultaneously in a wide area centered in western Japan from June 28 to July 8, 2018, owing to heavy rains caused by a rainy season front and Typhoon Prapiroon (Ministry of Land, Infrastructure, Transport and Tourism 2019) (for more information on the spatiotemporal transition of rainfall and flood risk, refer to Japan Meteorological Agency (2018) (Japan Meteorological Agency 2018).). These caused more than 700 casualties (Fire and Disaster Management Agency 2019) and economic losses of approximately 1.2154 trillion yen (Ministry of Land, Infrastructure, Transport and Tourism 2018a), making it the "worst flood disaster of the Heisei Era" (The Nikkei 2018).

The unit of analysis in this study is the municipalities within the four prefectures with a large number of damaged rivers during the 2018 Japan Floods: (1) Okayama, (2) Hiroshima, (3) Ehime, and (4) Fukuoka Prefectures. The focus on these prefectures is due to the availability of SR and POD data from Ota (2019). All municipalities within these four prefectures received flood warnings during the heavy rainfall in the 2018 Japan Floods (from June 28 to July 8, 2018) (Japan Meteorological Agency e). This allows for an analysis of how people responded to the flood warnings and the extent of the resulting damage. The final sample for analysis included 127 municipalities ($n = 127$), after excluding three municipalities from the 130 municipalities in the prefectures for the reasons discussed in Section 3.3b.

3.2 Outcome variables

As the outcome variables for the regression analyses, this study focused on four types of flood damage in each municipality that could be obtained from official statistics: the numbers of (1) fatalities [persons], (2) injuries [persons], (3) economic losses to general assets[5] (general assets and business interruption losses) (hereafter, simply "economic losses (general assets)") [thousands of yen], and (4) economic losses to general assets (crops) (hereafter, "economic losses (crops)") [thousands of yen]. By analyzing these four outcome variables, the study could determine which types of damage were affected by the performance of flood warnings. Data on the numbers of (1) fatalities and (2) injuries in each municipality were derived from technical disaster damage reports compiled by the prefectures (Hiroshima Prefecture 2018; Fukuoka Prefecture 2019; Okayama Prefecture 2020; Ehime Prefecture 2023) and the Cabinet Office (Cabinet Office 2019)[6]. The data for the

[5] "Economic losses to general assets" include physical damage to buildings, household goods, business assets, and crops, as well as losses due to business interruptions (Ministry of Land, Infrastructure, Transport and Tourism 2018b).

[6] These reports compiled by the prefectures show the numbers of deaths and injuries due to direct disaster damage at the municipal level, but do not distinguish between those caused by river overflows and those caused by landslides. On the other hand, the data from the Cabinet Office disclose the number of deaths and injuries due to landslide disasters at the municipal level. In this study, the number of deaths and injuries due to landslides at the municipal level based on the Cabinet Office data was subtracted from the number of deaths and injuries due to direct disaster-related deaths at the municipal level based on the data from each prefecture, and these resulting figures were considered as the number of (1) deaths and (2) injuries due to floods in each municipality.

(3) economic losses (general assets) and (4) economic losses (crops) for each municipality were based on a statistical survey of flood damage related to the 2018 Japan Floods (Ministry of Land, Infrastructure, Transport and Tourism 2018b). The distributions of each outcome variable are shown in Fig. 1, and the descriptive statistics are presented in Appendix A. As can be seen from the figure, each variable is mostly concentrated at zero, the distribution of which is left-skewed; that is, most municipalities experienced no damage, but others experienced much greater damage. Fig. 1

3.3 Explanatory variables

a. FAR and MER

The FAR [%] and MER [%] of flood warnings before the 2018 Japan Floods for each municipality were based on Ota (2019), where the SR [%] and POD [%] of the real-time flood warning map during the 2018 Japan Floods were published. Ota (2019) compiled the level of flood warnings and damage occurrences for each river (i.e., the spatial resolution at the river level) during the 2018 Japan Floods and calculated the SR and POD for each prefecture. For example, as illustrated in Table 2, the SR and POD for each prefecture were obtained for the level of "Warning (Red)" (Level 3)[7]

[7]Ota (2019) reported only the SR and POD values and the number of rivers where damage occurred in each prefecture: 84, 69, 37, and 98 rivers were damaged in Okayama,

[8], which requires evacuation preparations and the prompt commencement of evacuation for the elderly. From these SR and POD figures, the FAR and MER for each prefecture can be calculated using Eqs. (1) and (2), respectively.

Table 2

$$\text{FAR} = 100 - \text{SR}, \tag{1}$$

$$\text{MER} = 100 - \text{POD}. \tag{2}$$

In this study, we made the following three major assumptions to derive the FAR and MER of flood warnings for each municipality before the 2018 Japan Floods from the SR and POD of each prefecture during the 2018 Japan Floods published by Ota (2019).

- **Assumption 1:** The performance of flood warnings for each municipality is consistent with the performance of the warnings corresponding to the "Warning (Red)" level in the real-time flood warning

Hiroshima, Ehime, and Fukuoka Prefectures, respectively.

[8] Longer rivers may have a higher probability of a hit (i.e., at least one instance of damage is more likely to be observed along the entire river). That is, the length of the rivers can introduce geographical bias. However, the real-time flood warning map assesses the risk of flood-related disasters in small- and medium-sized rivers, and therefore, despite some geographical bias, the impact is considered limited owing to the limited variation in river size.

map[9].

- **Assumption 2:** The performance of warnings corresponding "Warning (Red)" level of real-time flood warning map at the time of the 2018 Japan Floods is representative of warning performance before the floods (i.e., ignorance of temporal variation)[10].

[9]In Japan, five levels have been set to provide an intuitive understanding of the level of a disaster and the actions to be taken. At Alert Level 3, people are expected to check hazard maps, prepare for evacuation, and in some cases voluntarily evacuate (Japan Meteorological Agency, d). Warnings associated with Level 3 are aimed to be issued several hours before the expected event (Japan Meteorological Agency, d). Flood warnings issued for each municipality and the warnings corresponding to the "Warning (Red)" level in the real-time flood warning map fall under the same Level 3. Therefore, we assumed that they had similar performance.

[10]Many factors that affect the performance of flood forecasting are location-specific. For example, local infrastructure and conditions (e.g., "dams," "weirs," "diversion and spillways," "environmental changes due to renovation," "backwaters," and "extremely small watersheds") account for a large proportion of the factors that are assumed to contribute to the reduced performance of forecasts (according to the presentation "Current Status and Issues of Hazard Distribution (Kikikuru) from the Viewpoint of IBF [IBF no Kanten de Miru Kikendo Bunpu (Kikikuru) no Genjo to Kadai]" by Takuma Ota of the Meteorological Research Institute, JMA, at the 2023 Spring Conference of the Meteorological Society of Japan). Since these factors do not change significantly in the short term, we assumed the performance of warnings at the time of the 2018 Japan Floods to be strongly correlated with that before the floods.

- **Assumption 3:** The performance of flood warnings issued for each municipality does not differ significantly within the same prefecture (i.e., ignorance of spatial variation within the same prefecture) [11].

Based on these assumptions, the FAR and MER of flood warnings issued in each municipality before the 2018 Japan Floods are assumed to be the same as those corresponding to the "Warning (Red)" level for each prefecture in the real-time flood warning map, as reported in Ota (2019). Thus, the FAR and MER values for each prefecture in Table 2 were used in the analysis as the FAR and MER for the municipalities within each prefecture.

b. Basin rainfall index criterion

Selecting appropriate confounding variables for which to control is crucial for reliable causal inference. Variables that influence both the cause and outcome should be included as explanatory variables in the model to minimize omitted variable bias (VanderWeele 2019). As the primary objective of the regression analysis in this study was to estimate the effects of the FAR and MER of flood warnings on the damage (outcome variables), it was important to control for confounding factors that influence both warning performance and flood damage.

[11] We assumed that the variation in local infrastructure and conditions, mentioned in footnote 10, is relatively small within a prefecture compared with between the prefectures.

This study took the basin rainfall index criterion (*Ryuiki Uryō Shisū Kijun* in Japanese) [.] as a primary confounding factor. The basin rainfall index criterion or the combination of the surface rainfall index[12] (Japan Meteorological Agency b). and basin rainfall index has been established for each municipality as the issuance criterion for flood warnings (Japan Meteorological Agency b). The basin rainfall index measures how rainfall in a river's upper reaches increases the risk of flooding in downstream target areas. It is calculated using a tank model and kinetic equations to quantify the volume of rainwater that flows into rivers over time via the ground surface and underground, and then flows down along the river, by dividing the river basin into a grid (mesh) of 1 km squares for approximately 20,000 rivers nationwide (Japan Meteorological Agency b). Lower criteria of the basin rainfall index may result in more frequent warnings, potentially increasing the number of false alarms. Therefore, the basin rainfall index criterion was considered to be correlated with the warning performance (FAR and MER). In addition, the basin rainfall index criterion reflects, to some extent, the conditions of levees and other infrastructure (Japan Meteorological Agency c). For example, areas with advanced infrastructure tend to have a higher basin rainfall index criterion. Flooding is less likely to occur in these areas,

[12] The surface rainfall index quantifies the amount of rain accumulated on the ground surface, considering factors such as ground cover, geology, and topographical gradient

resulting in reduced flood damage. In other words, the basin rainfall index criterion is also considered to be correlated with flood damage. Thus, the basin rainfall index criterion can influence both the performance of flood warnings (FAR and MER) and the extent of flood damage (outcome variables).

The basin rainfall index criteria for all the municipalities used in this analysis were obtained from the JMA's list of criteria for issuing warnings (Japan Meteorological Agency b). When a municipality had multiple basins and more than one criterion, the median value of the criteria was used. Due to the absence of basin rainfall index criteria, three municipalities—(1) Kamijima-cho, Ehime Prefecture; (2) Ikata-cho, Ehime Prefecture; and (3) Oto-machi, Fukuoka Prefecture—were excluded from the analysis. Descriptive statistics for the basin rainfall index criteria are provided in Appendix A.

c. Other variables

In addition to the basin rainfall index criteria, the following five variables were included as explanatory variables: (1) flooded area (residential land and others) [m^2], (2) flooded area (farmland) [m^2], (3) population [persons], (4) percentage of population over 65 years old [%], (5) sex ratio[13] [.] for each

[13] The sex ratio is the number of males per 100 females.

municipality. Covariate control recommends that variables that influence the outcome (i.e., flood damage) should also be included as explanatory variables in the regression analyses (VanderWeele 2019). Previous studies have indicated that the scale of hazards and local population density have significant positive effects on the number of fatalities and injuries (Simmons and Sutter 2009). Additionally, age and gender have been found to significantly influence the protective actions taken when a warning is issued (Trainor et al. 2015; Lim et al. 2019). Based on these findings, the aforementioned five variables were selected[14].

Data for these variables were sourced from public records. Specifically, (1) flooded area (residential land and others) [m^2] and (2) flooded area (farmland) [m^2] in each municipality were obtained from the disaster statistics (i.e., Flood Damage Statistics Survey in 2018) (Ministry of Land, Infrastructure, Transport and Tourism 2018b); (3) population [persons], (4) percentage of population over 65 years old [%], and (5) sex ratio [.] in each municipality were taken from the 2015 Census (Ministry of Internal Affairs

[14]Explanatory variables that only affect the outcome variables reduce the standard error of the estimated parameter (Yasui 2020). As we included the main confounding variable (i.e., the basin rainfall index criterion), the influence of other explanatory variables on FAR or MER is expected to be minimal. Therefore, although we can include as many variables as possible that could only affect the outcome, it would not substantially affect the means of the posterior distributions.

and Communications 2017). Descriptive statistics for these variables are provided in Appendix A. The maximum correlation between the explanatory variables including FAR, MER, and the basin rainfall index criterion was approximately 0.45 in absolute value, which is well below the 0.80–0.95 threshold typically associated with multicollinearity (Munro 2005; Matsuura 2022), suggesting that multicollinearity is not a concern in this analysis.

4. Regression Models

This study employed two types of regression models tailored to the nature of the outcome variables, which were either discrete or continuous data with non-negative values: For the discrete variable—(1) fatalities and (2) injuries—we used zero-inflated negative binomial (ZINB) models as described in Section 4.1; for the continuous variables—(3) economic losses (general assets) and (4) economic losses (crops)—we used the hurdle lognormal (HL) model as detailed in Section 4.2 [15] [16].

[15] As we constructed regression models for each outcome variable, the results for one outcome variable do not affect those for any other outcome variable.

[16] The dataset in this study is nested, with each municipality (the unit of analysis) belonging to a specific prefecture. This nested structure may introduce group differences owing to prefecture-level factors (e.g., variations in disaster-management systems across prefectures) that are not captured by the municipal-level explanatory variables alone (Snijders and Bosker 2011; Matsuura 2022). The dummy-variable approach is rec-

4.1 Zero-inflated negative binomial models

The variables representing fatalities and injuries contain many zeros and exhibit overdispersion, as described in Section 3.2, thus making the ZINB model appropriate (Liu et al. 2019; Feng 2021; Young et al. 2022). The ZINB model assumes a two-step data generation process. In the first process, a sample has a probability $1 - q$ of being 0 ($y = 0$), and in the second process, a sample has a probability q of following a negative binomial distribution. This two-step process effectively handles data with an excess of zeros. In addition, a negative binomial distribution is appropriate for overdispersed count data because it accounts for heterogeneity in the mean parameter of the Poisson distribution (Cameron and Trivedi 2005; Simmons and Sutter 2009). In this case study, the probability q represents whether a flood hazard occurs in a municipality (the first process), and next, the likelihood of deaths or injuries is captured (the possibility of no deaths or injuries is also considered) when the hazard occurs (the second process). The probability mass function for the outcome variable y is as follows:

ommended when the number of groups ($N < 10$) is small (Snijders and Bosker 2011). However, the prefecture dummies (Okayama Prefecture set as the reference level) were strongly correlated with FAR and MER (0.62 to 0.96 in absolute value), suggesting serious multicollinearity in our small sample size. Therefore, we focused on models with a non-nested structure.

$$\mathrm{ZINB}(y|q,\mu,\theta) = \begin{cases} 1 - q + q \cdot \mathrm{NB}(0|\mu,\theta) & \text{if } y = 0, \\ \\ q \cdot \mathrm{NB}(y|\mu,\theta) & \text{if } y > 1. \end{cases} \tag{3}$$

$\mathrm{NB}(y|\mu,\theta)$ is a negative binomial distribution with mean μ and variance $\mu + \mu^2/\theta$, and θ (> 0) is the dispersion parameter. The negative binomial probability mass function is given by

$$\mathrm{NB}(y|\mu,\theta) = \frac{\Gamma(\theta+y)}{\Gamma(\theta)\Gamma(y+1)} \left(\frac{\theta}{\theta+\mu}\right)^{\theta} \left(\frac{\mu}{\theta+\mu}\right)^{y}, \tag{4}$$

where Γ is the gamma function. As θ approaches infinity, the NB reduces to the Poisson distribution (therefore, small values of θ indicate overdispersion). In this study, the probability q of hazard occurrence was simplified to follow a Bernoulli process, while the mean μ of $\mathrm{NB}(y|\mu,\theta)$, which is primarily related to the amount of damage, was regressed on the explanatory variables.

The mean μ_i is formulated as follows:

$$\begin{aligned} \ln \mu_i = &\ln x_{Population,i} + \beta_0 + \beta_1 x_{FAR,i} \\ &+ \beta_2 x_{BasinRainfall,i} + \beta_3 x_{FloodedResidential,i} \\ &+ \beta_4 x_{FloodedFarmland,i} + \beta_5 x_{Elderly,i} + \beta_6 x_{Sex,i}, \end{aligned} \tag{5}$$

where $i \in \{1, \ldots, n\}$ denotes a municipality i. $x_{Population,i}$ is the population, $x_{FAR,i}$ the FAR, $x_{BasinRainfall,i}$ the basin rainfall index criterion, $x_{FloodedResidential,i}$ the flooded area (residential land and others), $x_{FloodedFarmland,i}$

the flooded area (farmland), $x_{Elderly,i}$ the percentage of population over 65 years old, and $x_{Sex,i}$ the sex ratio for Municipality i. When examining the effect of the MER, we replace $x_{FAR,i}$ with $x_{MER,i}$. The parameters β_k ($k = 0, \ldots, 6$) are the intercept and coefficients of the explanatory variables, respectively. These parameters, along with q and θ, are to be estimated. The main focus is on the estimation of β_1, the coefficient of FAR or MER. A positive β_1 indicates that a municipality with a higher FAR (or MER) has more fatalities or injuries. The first term $\ln x_{Population,i}$ on the right side of Eq. (5) is an offset term that allows the model to account for the number of fatalities or injuries relative to the population of each municipality (Christensen et al. 2010).

4.2 Hurdle lognormal model

The economic losses (general assets) and economic losses (crops) are non-negative continuous data with many zeros, as shown in Section 3.2; thus, we used HL models, which are well-suited to these data characteristics (Cameron and Trivedi 2005; Hamada et al. 2019). The HL models also assume a two-step data generation process. In the first process, a sample has a probability $1-q$ of being 0 ($y = 0$), and in the second process, a sample has a probability of q of following a lognormal distribution. This two-step process can represent data containing many zeros. In our case study, the probability of q represents whether a flood hazard occurs in a municipality (the first process), and the economic losses then always arise ($y > 0$) when the hazard occurs (the second process). The probability density function for the outcome variable y is as follows:

$$\mathrm{HL}(y|q,\mu,\sigma) = \begin{cases} 1-q & \text{if } y = 0, \\ q \cdot \mathrm{Lognormal}(y|\mu,\sigma) & \text{if } y > 0. \end{cases} \tag{6}$$

$\mathrm{Lognormal}(y|\mu,\sigma)$ represents the probability density function for the log-normal distribution given by

$$\mathrm{Lognormal}(y|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}y}\exp\left(-\frac{(\log y-\mu)^2}{2\sigma^2}\right), \tag{7}$$

where $\ln y$ follows a normal distribution with mean μ and standard deviation σ. As in Section 4.1, the mean μ of $\mathrm{Lognormal}(y|\mu,\sigma)$ was regressed on the explanatory variables.

The mean μ_i is formulated as follows:

$$\begin{aligned} \ln\mu_i =& \beta_0 + \beta_1 x_{FAR,i} + \beta_2 x_{BasinRainfall,i} + \beta_3 x_{FloodedResidential,i} \\ &+ \beta_4 x_{FloodedFarmland,i} + \beta_5 x_{Elderly,i} + \beta_6 x_{Sex,i} + \beta_7 x_{Population,i}. \end{aligned} \tag{8}$$

The parameters β_k ($k = 0,\ldots,7$), q, and σ are estimated.

4.3 Bayesian estimation

a. Overview of estimation

We employed a Bayesian approach to estimate the models. This method treats parameters as random variables. Drawing on Bayes' theorem, the prior probability distribution of unknown parameters, that is, the prior distribution, is updated, given the data obtained, to a posterior distribution (Gelman et al. 2013; Lee and Wagenmakers 2013; Levy and Mislevy 2017; Matsuura 2022). That is, $p(\boldsymbol{\eta}|\boldsymbol{D}) = p(\boldsymbol{D}|\boldsymbol{\eta})p(\boldsymbol{\eta})/p(\boldsymbol{D}) \propto p(\boldsymbol{D}|\boldsymbol{\eta})p(\boldsymbol{\eta})$, where $\boldsymbol{\eta}$ is an unknown parameter vector, $\boldsymbol{D}$ is data, $p(\boldsymbol{\eta})$ is a prior distribution of the parameters, $p(\boldsymbol{D}|\boldsymbol{\eta})$ is a likelihood, and $p(\boldsymbol{\eta}|\boldsymbol{D})$ is a posterior distribution. In most instances, the posterior distribution, which expresses the uncertainty of the parameters, is obtained by simulation using so-called Markov chain Monte Carlo (MCMC) methods. Sampling-based Bayesian methods depend less on asymptotic theory, and therefore have the potential to produce more reliable results, even with small samples, than those obtained by the maximum likelihood method (Song and Lee 2012; Van De Schoot et al. 2017). Our data are from only four prefectures; thus, the sample is not large, which justifies the use of the Bayesian method. Furthermore, the Bayesian method is more flexible with complex datasets and modeling (Hamada et al. 2019; Kruschke 2021). As our analysis incorporates zero-inflated and hurdle processes (as shown in Sections 4.1 and 4.2), the Bayesian approach is considered suitable.

b. Prior distributions

In the estimation, we used noninformative and weakly informative priors as follows:

$$\beta_k \sim \mathrm{Normal}\,(0, 10) \tag{9}$$

$$q \sim \mathrm{Uniform}\,(0, 1) \tag{10}$$

$$\theta \sim \mathrm{Gamma}\,(1, 1) \tag{11}$$

$$\sigma \sim \mathrm{Normal}^{+}(0, 5) \tag{12}$$

where $\mathrm{Uniform}\,(0, 1)$ is a continuous uniform distribution on the interval $[0, 1]$. $\mathrm{Gamma}\,(1, 1)$ is a gamma distribution whose density function is $\mathrm{Gamma}\,(\theta|a = 1, b = 1) = b^a \theta^{a-1} \exp\,(-b\theta)/\Gamma(a)$ with mean a/b and standard deviation $\sqrt{a}/b$. $\mathrm{Normal}^{+}(0, 5)$ is a normal distribution with a mean of 0 and a standard deviation of 5, truncated to positive values. Eq. (11) was only applied to ZINB models, and Eq. (12) was applicable only to HL models.

c. Computations

We conducted a Bayesian estimation using the Stan program (Carpenter et al. 2017) using RStan (Stan Development Team). We ran the MCMC with $16,000$ iterations, following a burn-in of 1000 iterations for each of the four chains, and every fifth iteration was saved for each chain. We drew $12,000$ $(= (16,000 - 1000) \times 4 \div 5)$ samples for each parameter.

Before running the simulation, we transformed the data to ease the convergence (Matsuura 2022) as follows: the FAR, MER, percentage of population over 65 years, and sex ratio were divided by 100. The flooded area (residential land and others), flooded area (farmland), and basin rainfall index criteria were standardized. The population was standardized only for HL models.

The MCMC chains were checked in terms of convergence and resolution. Specifically, model convergence was assessed using the Gelman-Rubin statistic (Gelman and Rubin 1992). In the following estimation, all parameters reached statistical values lower than the recommended value of 1.1. Posterior samples should be less autocorrelated and the effective sample size (ESS)[17] should be sufficient to obtain stable parameter estimates, particularly for the stable limits of credible intervals (Kruschke 2014, 2021). The ESS of each parameter exceeded the recommended value of $10,000$.

[17]The ESS is the effective number of steps in the MCMC chain after the clumpiness of autocorrelation is factored out.

5. Results

The estimation results for the posterior distributions of the FAR and MER parameters for each outcome variable—the (1) fatalities, (2) injuries, (3) economic losses (general assets), and (4) economic losses (crops)—are presented in Sections 5.1 through 5.4, respectively. Detailed results for the posterior distributions, including other parameters, are provided in the Supplementary Materials.

5.1 Fatalities

Figure 2a displays the posterior distribution of the parameter β_1 for the FAR; Fig. 2b shows the same for the MER. Each posterior distribution is depicted with the posterior mean in a circle and the 90% highest density interval (HDI)[18] on a line.

A positive trend was observed for FAR, where the 90% HDI did not overlap with 0, and the probability that the parameter was positive was extremely high ($\Pr(\beta_1 > 0) = 0.997$). This suggests that municipalities with higher FAR experienced more fatalities.

In contrast, the posterior distribution for MER was centered around

[18] The 90% HDI summarizes the distribution by specifying an interval that spans most of the distribution, say 90%, such that every point inside the interval has a higher credibility than any point outside it (Kruschke 2014).

0. It implies that there is no strong evidence to suggest that MER has a substantial effect on the number of fatalities. Fig. 2

5.2 *Injuries*

A positive trend in FAR was also observed for injuries (Fig. 3a). The 90% HDI did not overlap with 0, and the probability that the parameter was positive was extremely high ($\Pr(\beta_1 > 0) = 0.999$). This suggests that municipalities with higher FAR experienced more injuries.

For the MER parameter, a negative trend was observed, where the 90% HDI did not overlap with 0, and the probability that the parameter was positive was extremely low ($\Pr(\beta_1 > 0) = 0.018$) (Fig. 3b). This result suggests that a higher MER may be associated with fewer injuries. Fig. 3

5.3 *Economic losses (general assets)*

For economic losses (general assets), a positive trend was observed for the FAR parameter (Fig. 4a). The 90% HDI did not overlap with 0, and the probability that the parameter was positive was extremely high ($\Pr(\beta_1 > 0) = 1.000$). A positive parameter means that municipalities with higher FAR suffered greater economic losses (general assets).

For the MER parameter, the posterior distribution showed a negative trend, but the 90% HDI overlapped with 0 (Fig. 4b). This result suggests that there is no strong evidence for a positive effect of MER on economic losses (general assets). Fig. 4

5.4 Economic losses (crops)

Although positive trends were observed for both FAR and MER parameters regarding economic losses (crops), these effects were not as pronounced as those observed for the other outcome variables (Fig. 5). The 90% HDIs for both FAR and MER overlapped with 0, and the posterior means were close to 0, indicating that neither FAR nor MER had a strong or clear effect on economic losses (crops). Of the variables examined, the effect of FAR on general losses (crops) appeared to be the weakest. Fig. 5

6. Discussion and Conclusions

Frequent false alarms or missed events may erode public trust in warnings and their issuers, potentially leading to a decreased likelihood of protective action in response to future warnings, thereby increasing disaster damage. In this study, we used limited open data on FAR and MER in Japan to analyze their effects on human and property damage at the municipal level during the 2018 Japan Floods, employing Bayesian statistical models. We discuss which types of damage are associated with FAR and MER (Section 6.1) and suggest measures for improving the effectiveness of

FEWS (Section 6.2).

6.1 *Effect of FAR and MER*

The results in Section 5 suggest that we cannot deny the possibility that higher FAR increases several types of flood damage. Specifically, Figs. 2a, 3a, and 4a suggest that FAR may be associated with higher (1) fatalities, (2) injuries, and (3) economic losses (general assets), as indicated by the 90% HDI of the posterior distribution, which does not overlap with 0.

The finding that FAR is associated with the number of fatalities and injuries aligns with that of Simmon and Sutter (2009), who studied tornado warnings in the U.S. It is also consistent with previous studies (Ripberger et al. 2015; Trainor et al. 2015) that found that a higher FAR hampers protective actions in the future and during actual tornado warnings in the U.S. This suggests that among the measures of performance of flood warnings, the FAR is particularly strongly associated with life-saving behavior (e.g., evacuation).

Several reasons could explain why the FAR did not have as strong an effect on the other variable (i.e., economic losses (crops)). One possible reason is the "risk perception paradox," where higher risk perception does not necessarily lead to disaster preparedness actions (Wachinger et al. 2013). A systematic review by Wachinger et al. (2013) attributed this paradox to confusion or ignorance about the appropriate actions to take and a lack of capacity and resources to help oneself. While some of these factors were accounted for in this study (e.g., population over 65 years of age and sex ratio), there may be unmeasured effects that influence the outcomes. During the 2018 Japan Floods, even if people trusted the warnings, they might not have had the ability or knowledge to act.

Other possible reasons could be the characteristics of flood warnings. Flood warnings are issued when serious flooding is expected to occur, but they do not explicitly instruct people on the actions they should take, unlike evacuation orders (Yamori, 2016). Consequently, flood warnings might not have been strongly associated with intentions related to protective actions and might not have had significant effects on flood damage.

Conversely, MER did not show a positive association with the casualties or economic losses (Figs. 2b–5b). A possible reason is the influence of past disaster experiences in addition to the reasons mentioned above. Wachinger et al. (2013) cite past disaster experience, in addition to trust in warnings, as one factor that influences heightened risk perception. Municipalities with more missed events may have suffered significant damage in the past, and as a result, it can be inferred that residents had a higher risk perception, and some residents took action when a warning was issued. Okumura et al. (2001) also showed that when a missed event occurred, unlike in the case of a false alarm, people increased their subjective reliance on evacuation warnings and were more willing to take evacuation actions. The fact that the posterior distribution of the MER parameter showed a negative trend for some outcome variables (Model 1 for Figs. 3b and 4b) is consistent with their findings. Therefore, we conclude that we obtained the result that higher MER does not necessarily increase flood damage.

6.2 Implication for effective FEWS

Our findings suggest that issuing frequent warnings, which may result in a large number of false alarms, can have negative consequences, as concluded by Oikawa and Katada (2016) based on their experiments. One possible mechanism is that frequent false alarms decrease people's trust in warnings, resulting in their reluctance to take protective action (e.g., evacuation) in response to subsequent warnings. Therefore, a strategy issuing frequent warnings must consider the adverse effects of false alarms on protection actions and reduce such adverse effects. For example, LeClerc and Joslyn (2015) suggested that providing information on probabilistic forecasts, in addition to information on deterministic forecasts, may increase trust in and responsiveness to weather information. In the context of floods in Japan, offering probabilistic data may encourage residents to take protective action. Examples of providing probabilistic information on floods and other hazards can be found in Millet et al. (2020) and Watanabe et al. (2022) (Millet et al. 2020; Watanabe et al. 2022).

Our findings also suggest that the development of technologies and systems that contribute to reducing the FAR may be particularly effective in reducing flood damage. Tanaka et al. (2008) and Ota (2019) discussed the changes in the numbers of false alarms and missed events following the introduction of new flood warning criteria in May 2008 and July 2017, respectively (Tanaka et al. 2008; Ota 2019). Both studies demonstrated that the new criteria based on the basin rainfall index and surface rainfall index significantly reduced the number of false alarms, while largely maintaining the number of missed events. In other words, the FAR reduction was achieved without increasing the MER. Such improvements in warning criteria are considered effective in reducing flood damage, especially casualties, and similar improvements in technologies and systems will be required in the future[19].

[19]Needless to say, we do not deny the practical or potential importance of reducing the MER without increasing the FAR; however, our results imply that reducing the FAR without increasing the MER should be a priority.

6.3 Limitations and future directions

This study has several limitations. The first and most significant limitation is the reliance on three major assumptions in calculating the FAR and MER for each municipality, as discussed in Section 3.3a. These assumptions were made because of the limited availability of open data on FAR and MER in Japan. Future work would benefit from more granular and widely available data on false alarms and missed events at the municipal and monthly levels, eliminating the need for such assumptions. Once more detailed data become available, panel data analysis and other methods can provide deeper insights into the effects of warning performance.

The second limitation is the use of the basin rainfall index criterion as a confounding factor. This variable is reasonable as the main factor, as discussed in Section 3.3b.; however, as is often the case with cross-sectional regression, we acknowledge that we may have missed some variables that affect both warning performance and flood damage, leading to omitted variable bias. The methods discussed in the first limitation can help reduce this bias.

The third limitation is the study's focus on the direct relationship between warning performance (FAR and MER) and flood damage without explicitly analyzing the intervening processes. As discussed in Section 2, the effects of FAR or MER on damage are likely to involve public perceptions of and trust in warnings and issuers. Understanding these processes is important for developing better risk communication strategies that lead to protective actions, given that improving the performance of weather forecasts in a short time and at low cost is not feasible. Another possibility that has not been discussed extensively is the intervening influence of other stakeholders, such as local governments. For example, municipalities experiencing frequent false alarms (high FAR) might anticipate public reluctance to act and increase efforts to encourage evacuation (e.g., call for evacuation), potentially increasing individuals' protective actions and mitigating damage despite a higher FAR. Future studies should explore these processes in greater detail.

The fourth limitation is the exclusive focus on flood warnings, as they were issued for all municipalities during the 2018 Japan Floods. Analyzing higher-level weather warnings (e.g., emergency warnings (*Tokubetsu Keihou* in Japanese)) and directives for action (e.g., evacuation orders) could help clarify which types of information are most effective in mitigating damage and should be prioritized for improvement.

Despite these limitations, this study is the first to empirically examine the effects of FAR and MER on flood damage in Japan, where open data on flood warning performance are scarce. These findings provide useful information for warning providers and developers of weather forecasting and warning systems, highlighting the potential disaster mitigation effects of warning performance and the future direction of effective warning strategies and system development. The study also underscores the importance of making weather forecasting and warning data more openly available in Japan, which could stimulate further research into weather forecasting and warnings.

Supplement The supplementary material includes the estimation results (i.e., the summary of the posterior distributions of all the parameters for each model).

Data Availability Statement

The dataset and codes for the analyses are available at `https://doi.org/xxxxxxx`. [The doi number is issued after the acceptance of the article.]

Acknowledgements

The authors would like to thank Masamitsu Onishi for valuable discussions. This study was partially supported by the Japan Society for the Promotion of Science (KAKENHI Grant No. 22K18822) and JST (Moonshot R&D Program Grant No. JPMJMS2281). The authors declare that

they have no known competing financial interests or personal relationships that could appear to influence the work reported in this study.

A. Sample characteristics

The descriptive statistics for the outcome variables are presented in Table 3, while the statistics of the data for the explanatory variables (excluding FAR and MER) are shown in Table 4.

Table 3

Table 4

References

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525(7567)**, 47–55.

Cabinet Office, 2019: Damage caused by the 2018 Japan Floods [heisei 30 nen 7 gatsu gouu niyoru higai jokyo tou ni tsuite]. `https://www.bousai.go.jp/updates/h30typhoon7/index.html`. Accessed: 2023-02-13 (in Japanese).

Cameron, A. C., and P. K. Trivedi, 2005: *Microeconometrics: Methods and Applications*. Cambridge University Press.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell, 2017: Stan: A probabilistic programming language. *Journal of Statistical Software*, **76**.

Christensen, R., W. Johnson, A. Branscum, and T. E. Hanson, 2010: *Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians*. CRC press.

Ehime Prefecture, 2023: Casualties and damage to residential properties as a result of 2018 Japan Floods [heisei 30 nen 7 gatsu gouu niyoru jintekihigai, juukahigai nitsuite]. `https://www.pref.ehime.jp/h12200/h3007-gouu-saigai-oshirase-.html`. Accessed: 2023-02-13 (in Japanese).

Feng, C. X., 2021: A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *Journal of Statistical Distributions and Applications*, **8(1)**, 8.

Fire and Disaster Management Agency, 2019: Damage caused by the 2018 Japan Floods and Typhoon No. 12, and the response of fire-fighting and other agencies (60th report) [heisei 30 nen 7 gatsu gouu oyobi taifu 12 gou ni yoru higai jokyo oyobi shobo kikan tou no taiou jokyo (dai 60 pou)]. `https://www.fdma.go.jp/disaster/info/items/190820nanagatugouu60h.pdf`. Accessed: 2024-01-15 (in Japanese).

Fukuoka Prefecture, 2019: Disaster annual report in 2018 [heisei 30 nen saigai nenpo]. https://www.pref.fukuoka.lg.jp/contents/saigainenpou-30.html. Accessed: 2023-02-13 (in Japanese).

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, 2013: *Bayesian Data Analysis*. CRC Press.

Gelman, A., and D. B. Rubin, 1992: Inference from iterative simulation using multiple sequences. *Statistical Science*, **7(4)**, 457–472.

Hallegatte, S., 2012: A cost effective solution to reduce disaster losses in developing countries: hydro-meteorological services, early warning, and evacuation. *World Bank Policy Research Working Paper*, **(6058)**.

Hamada, H., A. Ishida, and H. Shimizu, 2019: *Bayesian Statistical Modeling for the Social Sciences [Shakaikagaku no tameno Beizu Toukei Moderingu]*. Asakura Publishing. (in Japanese).

Hiroshima Prefecture, 2018: Damage caused by the 2018 Japan Floods [heisei 30 nen 7 gatsu gouu saigai niyoru higai tou ni tsuite]. https://www.pref.hiroshima.lg.jp/site/bousaisaigaijouhou/list3673-13867.html. Accessed: 2023-02-13 (in Japanese).

Japan Meteorological Agency, 2018: The 2018 Japan Floods [heisei 30 nen

7 gatsu gou]. https://www.data.jma.go.jp/stats/data/bosai/report/2018/20180713/20180713.html. (in Japanese).

Japan Meteorological Agency, a: Real-time risk map: Flood (hazard distribution of flood warning) [kouzui kikukuru (kouzui keihou no kikendo bunpu)]. https://www.jma.go.jp/jma/kishou/know/bosai/riskmap_flood.html. Accessed: 2024-02-01 (in Japanese).

Japan Meteorological Agency, b: List of criteria for issuing warnings and advisories [keiho chuiho happyou kijun ichiran]. https://www.jma.go.jp/jma/kishou/know/kijun/index.html. Accessed: 2023-07-26 (in Japanese).

Japan Meteorological Agency, c: Basin rainfall index [ryuiki uryo shishu]. https://www.jma.go.jp/jma/kishou/know/bosai/ryuikishisu.html. Accessed on 2024-01-15 (in Japanese).

Japan Meteorological Agency, d: Correspondence between meteorological disaster prevention information and alert levels [bousaikisyouzyouhou to keikaireberu no taiou nituite]. https://www.jma.go.jp/jma/kishou/know/bosai/alertlevel.html. Accessed: 2024-01-18 (in Japanese).

Japan Meteorological Agency, e: Verification of the accuracy of disaster prevention weather information in cases of heavy rainfall and improvement of announcement criteria [oame jirei tou niokeru bousai kisho joho no seido kensho to happyou kijun no kaizen]. https://www.jma.go.jp/jma/kishou/know/jirei/index.html. Accessed: 2024-01-15 (in Japanese).

Katada, T., and N. Murasawa, 2009: Present conditions and issues of correspondence of administrations and inhabitants for distant tsunami. *Journal of Disaster Information Studies*, **7**, 94–103. (in Japanese).

Kaziya, A., K. Akaishi, T. Yokota, F. K. Kusano, N. Sekiya, and Y. Takahashi, 2018: Reduction of evacuation rate after Izu Oshima Sediment Disaster in 2013 and examination of its cause and measures based on questionnaire survey. *Journal of Disaster Information Studies*, **16(1)**, 37–47. (in Japanese).

Kotani, H., W. Ogawa, K. Matsushima, J. Susaki, and Y. Sawada, 2024: Modeling the dynamics of social preparedness with regional interactions to explore efficient flood early warning systems. *Weather, Climate, and Society*, **16(4)**, 771–788.

Kruschke, J., 2014: *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.

Kruschke, J. K., 2021: Bayesian analysis reporting guidelines. *Nature Human Behaviour*, **5(10)**, 1282–1291.

LeClerc, J., and S. Joslyn, 2015: The cry wolf effect and weather-related decision making. *Risk Analysis*, **35(3)**, 385–395.

Lee, M. D., and E.-J. Wagenmakers, 2013: *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Levy, R., and R. J. Mislevy, 2017: *Bayesian Psychometric Modeling*. Chapman and Hall/CRC.

Lim, J. R., B. F. Liu, and M. Egnoto, 2019: Cry wolf effect? evaluating the impact of false alarms on public responses to tornado alerts in the southeastern united states. *Weather, Climate, and Society*, **11(3)**, 549–563.

Liu, L., Y.-C. T. Shih, R. L. Strawderman, D. Zhang, B. A. Johnson, and H. Chai, 2019: Statistical analysis of zero-inflated nonnegative continuous data. *Statistical Science*, **34(2)**, 253–279.

Matsuura, K., 2022: *Bayesian Statistical Modeling with Stan, R, and Python*, Volume 526. Springer.

Millet, B., A. P. Carter, K. Broad, A. Cairo, S. D. Evans, and S. J. Majumdar, 2020: Hurricane risk communication: Visualization and behavioral science concepts. *Weather, Climate, and Society*, **12(2)**, 193–211.

Ministry of Internal Affairs and Communications, 2017: Census in 2015 [heisei 27 nen kokusei chousa]. https://www.e-stat.go.jp/stat-search/files?stat_infid=000031594311. Accessed: 2023-02-15 (in Japanese).

Ministry of Land, Infrastructure, Transport and Tourism, 2018a: Flood damage statistical survey/flood damage statistical survey/flood damage by extreme weather conditions (table-10), 2018 [suigai toukei chousa/heisei 30 nen suigai toukei chousa/ijo kisho betsu suigai higai (hyou-10)]. https://www.e-stat.go.jp/stat-search/files?stat_infid=000032223721. Accessed: 2024-01-15 (in Japanese).

Ministry of Land, Infrastructure, Transport and Tourism, 2018b: Flood damage statistical survey/flood damage statistical survey/general property and other flood damage by major extreme weather events by municipality (table-5), 2018. [suigai toukei chousa/heisei 30 nen suigai toukei chousa/shikuchouson betsu shuyo ijo kisho betsu ippan shisan tou suigai higai (hyou-5)]. https://www.e-stat.go.jp/stat-search/files?stat_infid=000032223688. Accessed: 2023-02-13 (in Japanese).

Ministry of Land, Infrastructure, Transport and Tourism, 2019: Overview of the 2018 Japan Floods [heisei 30 nen 7 gatsu gou no gaiyo]. https://www.mlit.go.jp/river/shinngikai_blog/chisui_kentoukai/dai03kai/dai03kai_siryou6.pdf. (in Japanese).

Munro, B. H., 2005: *Statistical Methods for Health Care Research*, Volume 1. Lippincott Williams & Wilkins.

NOAAForecast verification glossary. https://www.swpc.noaa.gov/sites/default/files/images/u30/Forecast%20Verification%20Glossary.pdf. Accessed: 2025-01-27.

Oikawa, Y., and T. Katada, 2016: Effects of repetitive false evacuation advisory on residents' behavior. *Journal of Disaster Information Studies*, **14**, 93–104. (in Japanese).

Okayama Prefecture, 2020: Records of the 2018 Japan Floods [heisei 30 nen 7 gatsu gouu saigai kirokushi]. https://www.pref.okayama.jp/page/653529.html. Accessed: 2023-02-13 (in Japanese).

Okumura, M., M. Tsukai, and T. Shimoaraiso, 2001: Reliance on disaster warning and response. *Infrastructure Planning Review*, **18**, 311–316. (in Japanese).

Ota, T., 2019: Chapter 2: Accuracy verification of "index and criteria" used for heavy rainfall and flood warnings [dai 2 sho ooame kouzui keihou ni mochiiteiru "shisu to kijun" no seido kensho]. https://www.jma.go.jp/jma/kishou/books/yohkens/24/chapter2.pdf. Accessed: 2024-07-15 (in Japanese).

Ripberger, J. T., C. L. Silva, H. C. Jenkins-Smith, D. E. Carlson, M. James, and K. G. Herron, 2015: False alarms and missed events: The impact and origins of perceived inaccuracy in tornado warning systems. *Risk Analysis*, **35(1)**, 44–56.

Rogers, D., and V. Tsirkunov, 2011: Costs and benefits of early warning systems. *Global Assessment Report on Disaster Risk Reduction*.

Roulston, M. S., and L. A. Smith, 2004: The boy who cried wolf revisited: The impact of false alarm intolerance on cost–loss scenarios. *Weather and Forecasting*, **19(2)**, 391–397.

Sawada, Y., R. Kanai, and H. Kotani, 2022: Impact of cry wolf effects on social preparedness and the efficiency of flood early warning systems. *Hydrology and Earth System Sciences*, **26(16)**, 4265–4278.

Simmons, K. M., and D. Sutter, 2009: False alarms, tornado warnings, and tornado casualties. *Weather, Climate, and Society*, **1(1)**, 38–53.

Snijders, T. A., and R. Bosker, 2011: *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. SAGE.

Song, X.-Y., and S.-Y. Lee, 2012: A tutorial on the Bayesian approach for analyzing structural equation models. *Journal of Mathematical Psychology*, **56(3)**, 135–148.

Stan Development Team RStan: the R interface to Stan. `https://mc-stan.org/`. R package version 2.26.24.

Tanaka, N., T. Ota, and Y. Makihara, 2008: Flood warning/advisory improvement based on JMA runoff index. *Sokkou Jihou*, **75(2)**, 35–69. (in Japanese).

The Nikkei, 2018: One month after the 2018 Japan Floods: the worst flooding in the Heisei era, scars on the Japanese islands [nishinihon gouu 1 kagetsu heisei saiaku no suigai, rettou ni kizuato]. `https://www.nikkei.com/article/DGXMZO33875110W8A800C1CC1000/`. Accessed: 2024-07-29 (in Japanese).

Trainor, J. E., D. Nagele, B. Philips, and B. Scott, 2015: Tornadoes, social science, and the false alarm effect. *Weather, Climate, and Society*, **7(4)**, 333–352.

Van De Schoot, R., S. D. Winter, O. Ryan, M. Zondervan-Zwijnenburg, and S. Depaoli, 2017: A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods*, **22(2)**, 217.

VanderWeele, T. J., 2019: Principles of confounder selection. *European Journal of Epidemiology*, **34**, 211–219.

Wachinger, G., O. Renn, C. Begg, and C. Kuhlicke, 2013: The risk perception paradox—implications for governance and communication of natural hazards. *Risk Analysis*, **33(6)**, 1049–1065.

Watanabe, M., S. Ito, W. Ma, and D. Yamazaki, 2022: Importance of two-way communication with users for effective utilization of ensemble flood forecast information. *Journal of Japan Society of Hydrology and Water Resources*, **35(2)**, 104–121. (in Japanese).

World Meteorological Organization, 2022: , Early warnings for all the UN global early warning initiative for the implementation of climate adaptation executive action plan 2023–2027. Technical report, World Meteorological Organization.

Yamori, K., 2016: Disaster information from the viewpoint of speech act theory: Constative, performative, and declarative utterances. *Journal of Disaster Information Studies*, **14**, 1–10. (in Japanese).

Yasui, S., 2020: *Introduction to Causal Inference/Econometrics [Kouka kensyou nyuumon]*. Gizyutsuhyouronsya, inc. (in Japanese).

Yoshii, H., I. Nakamura, H. Nakamori, and Y. Jibiki, 2008: The information dissemination and behaviors of the inhabitants in the 2 earthquakes in Hokkaido 2006-2007. *Disaster-Information Management*, **14**, 1–55. (in Japanese).

Young, D. S., E. S. Roemmele, and P. Yeh, 2022: Zero-inflated modeling part I: Traditional zero-inflated count regression models, their applications, and computational tools. *Wiley Interdisciplinary Reviews: Computational Statistics*, **14(1)**, e1541.

List of Figures

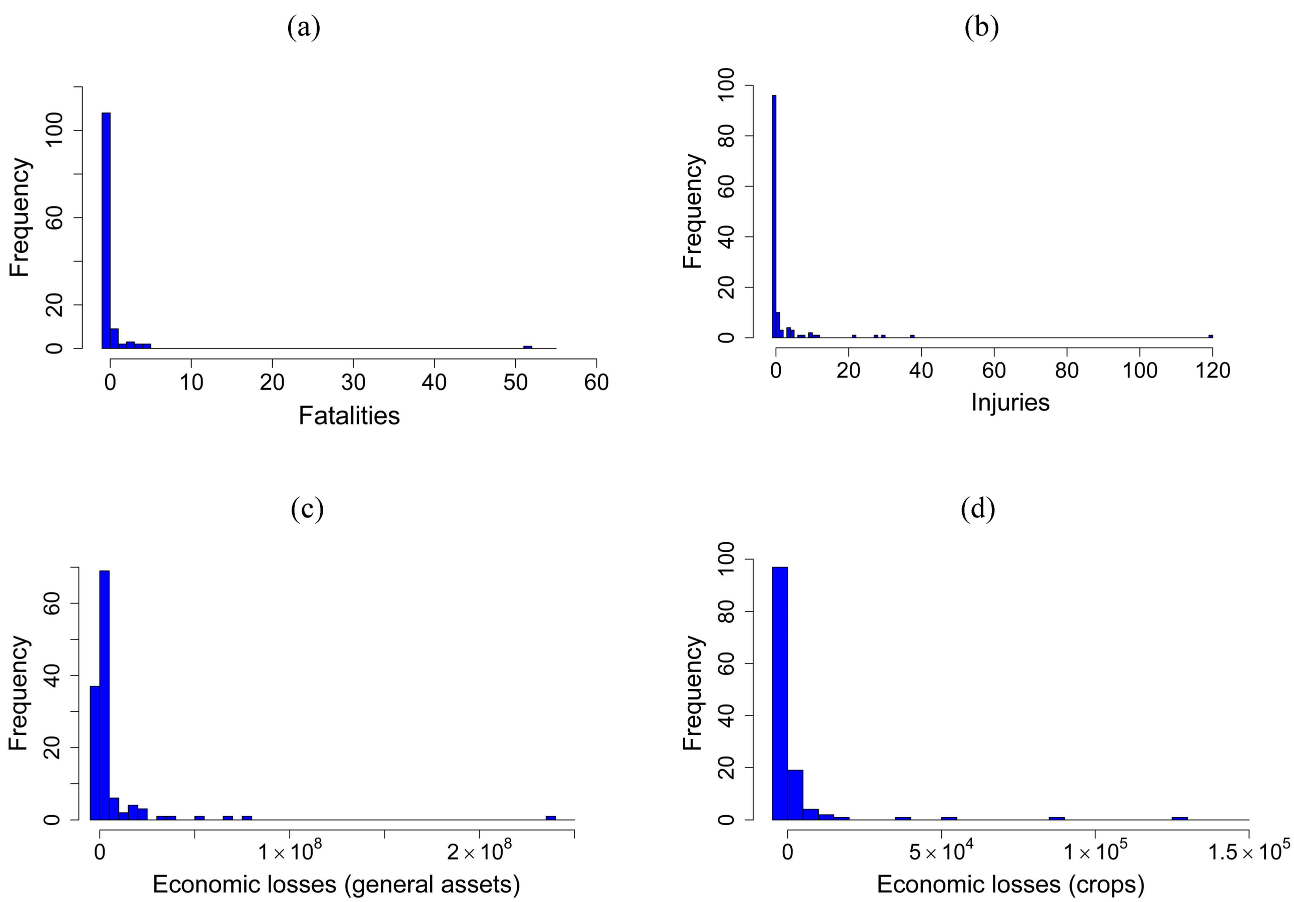


Fig. 1. Histograms of (a) fatalities, (b) injuries, (c) economic losses (general assets), and (d) economic losses (crops).

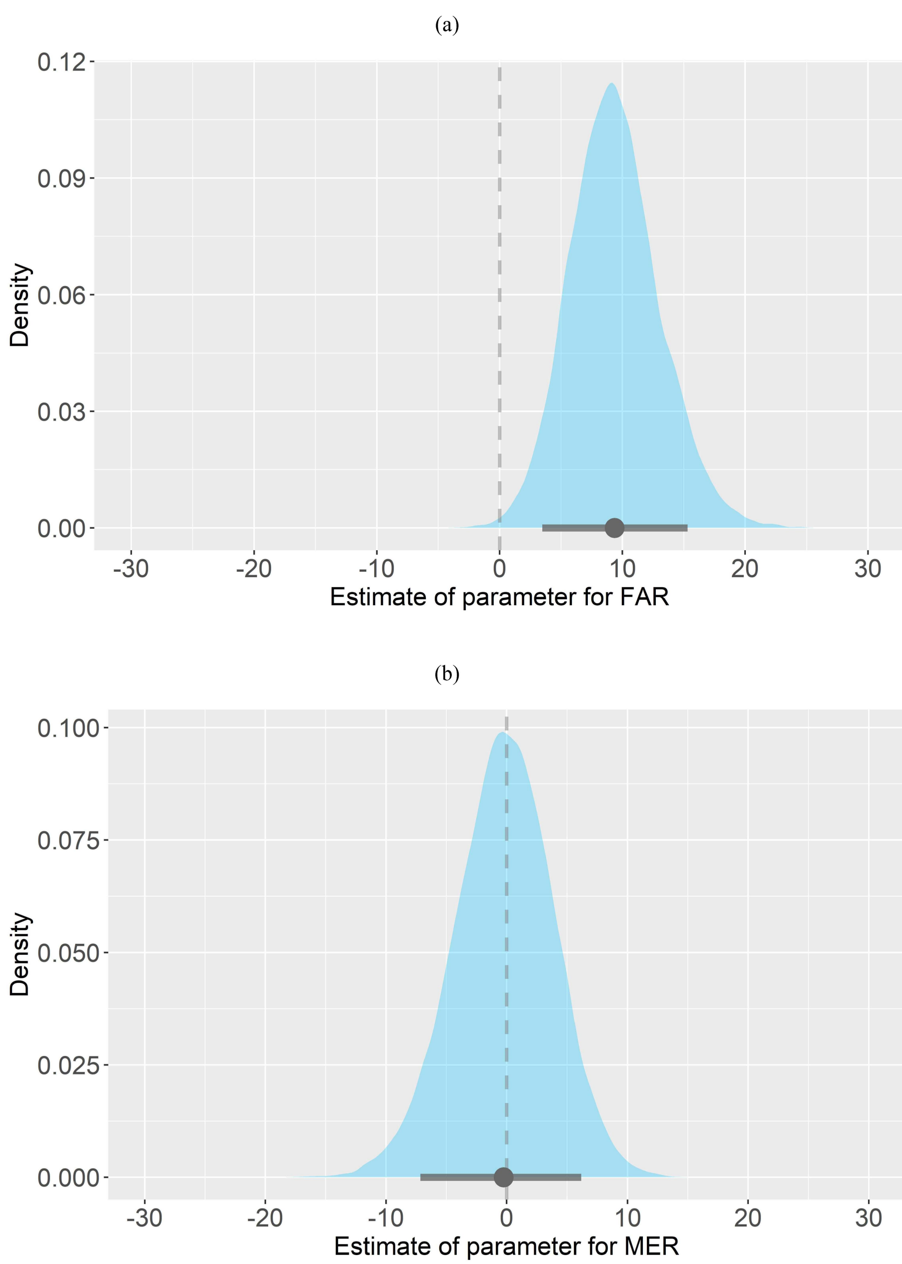


Fig. 2. Estimation results for fatalities: (a) Posterior distribution (mean and 90% HDI) of FAR parameter, and (b) that of MER.

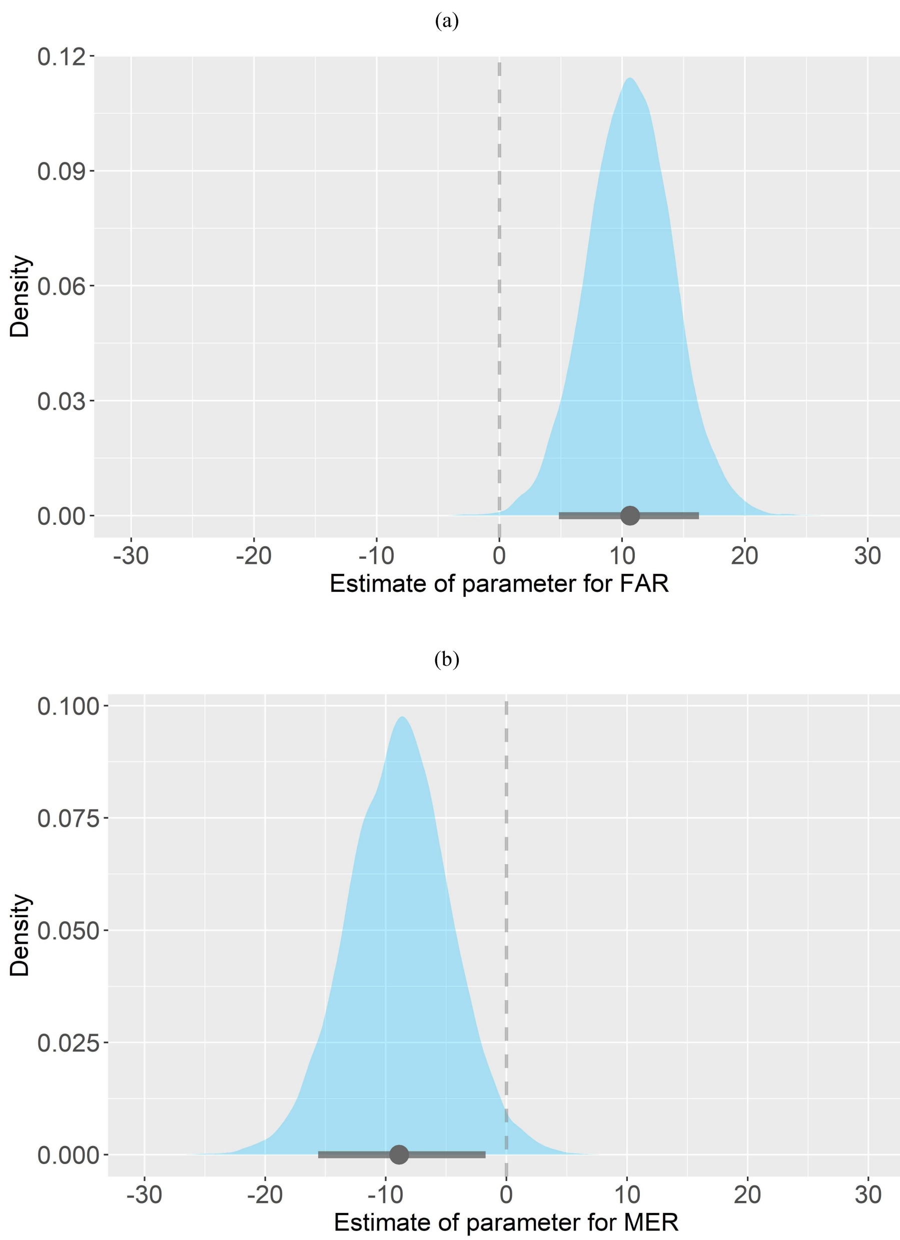
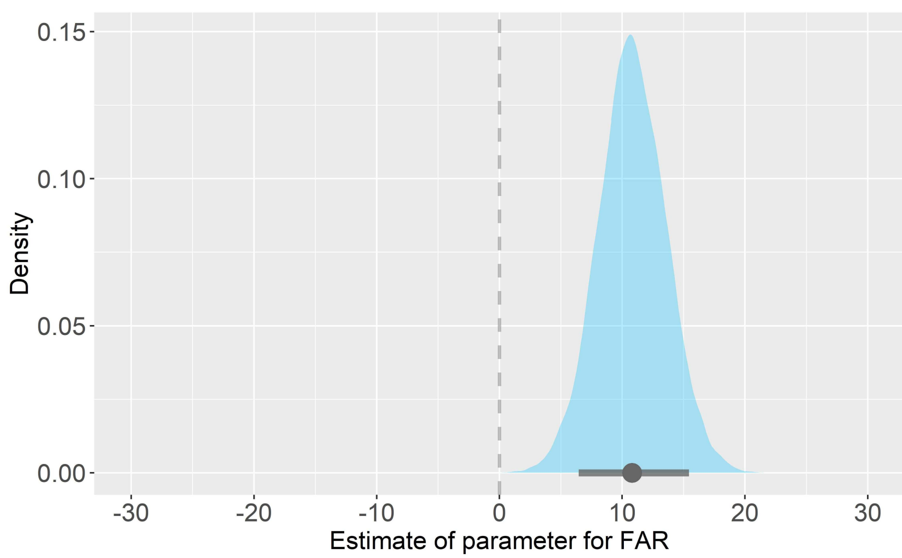


Fig. 3. Estimation results for injuries: (a) Posterior distribution (mean and 90% HDI) of FAR parameter, and (b) that of MER.

(a)



(b)

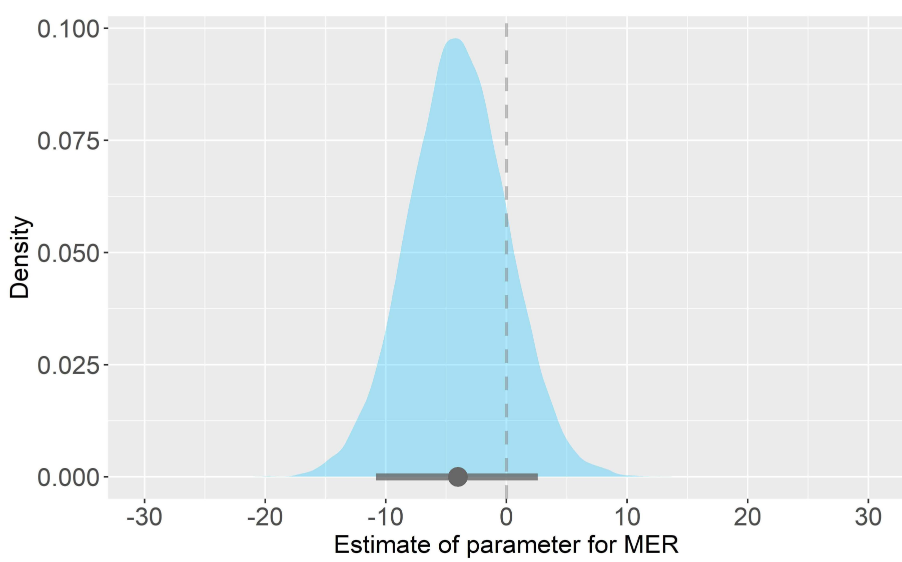


Fig. 4. Estimation results for economic losses (general assets): (a) Posterior distribution (mean and 90% HDI) of FAR parameter, and (b) that of MER.

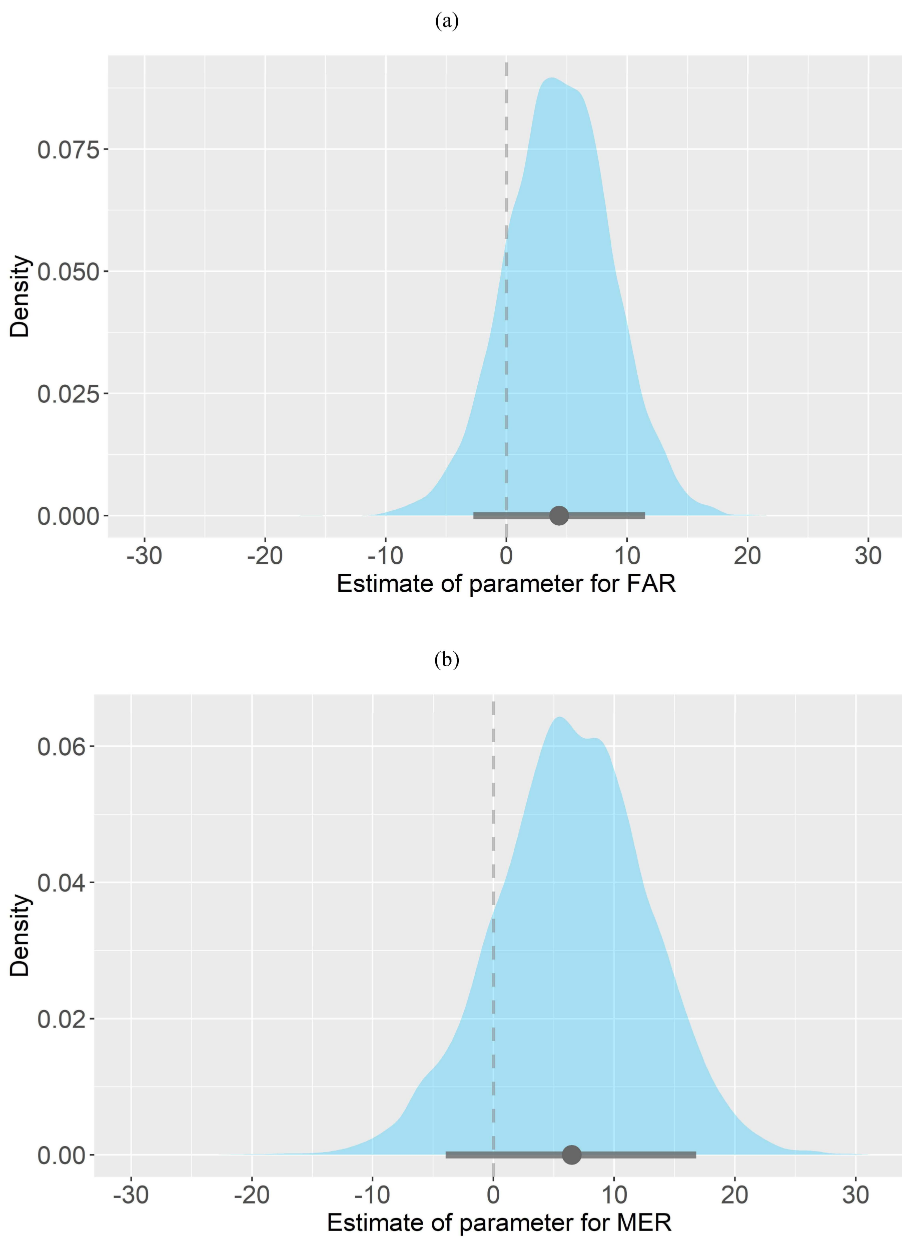


Fig. 5. Estimation results for economic losses (crops): (a) Posterior distribution (mean and 90% HDI) of FAR parameter, and (b) that of MER.

List of Tables

Table 1. Warning performance typology

		Hazard observed	
		Yes	No
Hazard forecasted	Yes	Hit	False alarm
	No	Missed event	All clear

Table 2. SR and POD according to Ota (2019); FAR and MER used for this study

		Okayama	Hiroshima	Ehime	Fukuoka
Ota (2019)	SR [%]	23	21	13	40
	POD [%]	74	93	78	87
This study	FAR [%]	77	79	87	60
	MER [%]	26	7	22	13

Table 3. Descriptive statistics of outcome variables

	Mean	Variance	Minimum	Maximum
Fatalities [persons]	0.72	21.76	0	52
Injuries [persons]	2.70	140.35	0	120
Economic losses (general assets) [thousand yen]	5.91×10^6	5.74×10^{14}	0	239737892
Economic losses (crops) [thousand yen]	3.06×10^4	2.17×10^{10}	0	1288800

Table 4. Descriptive statistics of explanatory variables

	Mean	Variance	Minimum	Maximum
Basin rainfall index criterion [.]	1.28×10	5.10×10	3.7	49.1
Flooded area (residential land and others) [m^2]	5.04×10^5	4.42×10^{12}	0	21084039
Flooded area (farmland) [m^2]	4.95×10^5	5.58×10^{12}	0	22850940
Population [persons]	8.84×10^4	4.29×10^{10}	866	1538681
Percentage of population over 65 years old [%]	3.22×10	4.08×10	16	49
Sex ratio [.]	9.05×10	1.45×10	82	106