JARQ 60 (2), 175-186 (2026) https://www.jircas.go.jp

# Has Economic Growth Mitigated Natural Disaster Damage in Vietnam? A Hybrid Approach to Panel Data Analysis

**Tetsuya MICHINAKA[1]*, Makoto EHARA[1], Tan Phuong VU[2], Thuy My Linh NGUYEN[3], Takashi OKAMOTO[4], Katsuto SHIMIZU[5], Takeshi TOMA[6], Van Tiep HA[7] and Van Tuan VU[7]**

[1] Center for Biodiversity and Climate Change, Forestry and Forest Products Research Institute (FFPRI), Tsukuba, Japan
[2] Vietnam Forest Certification Office (VFCO), Hanoi, Vietnam
[3] Vietnamese Academy of Forest Sciences (VAFS), Hanoi, Vietnam
[4] Tohoku Research Center, Forestry and Forest Products Research Institute (FFPRI), Morioka, Japan
[5] Shikoku Research Center, Forestry and Forest Products Research Institute (FFPRI), Kochi, Japan
[6] Research Planning Department, Forestry and Forest Products Research Institute (FFPRI), Tsukuba, Japan
[7] Forest Science Center of Northwestern Vietnam, Vietnamese Academy of Forest Sciences (VAFS), Son La, Vietnam

## Abstract
This study focuses on Vietnam, a nation prone to various natural disasters, including storms, typhoons, floods, flash floods, and landslides. Utilizing annual panel data from 2010 to 2020 at the provincial level, we conducted a panel data analysis of provincial natural disaster damage in monetary terms considering meteorological, hydrological, geographic, and socioeconomic factors, including elevation, terrain ruggedness index (TRI), precipitation, gross regional domestic product (GRDP) per capita, and agricultural GRDP. Our hybrid model results indicate that precipitation was a significant contributor to economic damage. Two geographical factors, elevation and TRI, also influenced it. Elevation was associated with a decrease in economic damage, while TRI was associated with an increase. Economic development, as indicated by GRDP per capita, was found to be associated with an increase in natural disaster damage at lower levels of GRDP per capita; however, this relationship shifted to one of decreasing damage after reaching a certain threshold in GRDP per capita. Provinces with higher agricultural GRDP tended to suffer greater damage, but within-province agricultural development was not statistically significant at the 5% level. This finding aligns with the understanding that crop farming and the aquaculture industry are more susceptible to natural disasters because they are more directly exposed to natural elements. This also implies that regions where agriculture comprises a significant share of the economy are less resilient to natural disasters. These findings offer valuable insights for policymakers for sustainable economic development.

**Discipline:** Social Science
**Additional key words:** agriculture, disaster damage mitigation, economic development, exposure, policy

## Introduction

Mankind has been making efforts to protect itself from the damage caused by natural disasters for thousands of years. Stories of human struggles against such disasters can be found across the world, from ancient China and India to ancient Rome and Mesoamerica. Throughout history, balancing disaster risk management with economic development has remained a challenge. Economic development can be a double-edged sword: while the pursuit of better well-being may lead to environmental disturbances that contribute to natural disasters, it also enables capacity-building and awareness, helping societies mitigate damage and escape disasters. Quantitative analysis of the relationships between economic development and disaster damage can help

*Corresponding author: michinaka_tetsuya650@ffpri.go.jp
Received 9 April 2025; accepted 8 August 2025; J-STAGE Advanced Epub 28 January 2026.
https://doi.org/10.6090/jarq.24J33

policymakers to use limited resources more effectively.

There is extensive research on the human and economic losses caused by natural disasters, but findings remain inconclusive (Skidmore & Toya 2002, Cuaresma et al. 2008, Noy & Vu 2010, Loayza et al. 2012, Klomp & Valckx 2014, González 2022). Many studies, such as González (2022), argue that developing countries are more vulnerable to the adverse effects of natural disasters. On the other hand, numerous studies have examined the impacts of economic development on disaster damage. Kahn (2005) found a decreasing trend in national average deaths per disaster and annual disaster-related deaths as real GDP per capita increased across 73 countries from 1980 to 2002. Kellenberg & Mobarak (2008) identified a nonlinear relationship between aggregate income and disaster damage, shifting from a positive to a negative correlation. Raschky (2008) explored the impact of GDP per capita on disaster losses, but the reciprocal relationships between GDP and losses were ignored, which may have introduced endogeneity bias. Schumacher & Strobl (2011), using data from 181 countries from 1980 to 2004, incorporated hazard exposure and found that the effects of economic development on disaster losses varied according to the hazard level. Padli et al. (2018) highlighted the role of economic development in mitigating disaster impacts—measured in total deaths, affected populations, and economic losses—through a panel data analysis of 79 countries. Song & Park (2019) focused on the socioeconomic impacts on human losses and economic costs, providing a damage prediction formula that includes GDP, population, and land area.

At the national and subnational levels, research presents mixed findings. Tasri et al. (2020) analyzed economic disaster losses in West Sumatra, Indonesia, and found a positive relationship between per capita income and disaster losses, arguing that economic policies have not sufficiently supported environmental sustainability. Similarly, Bahinipati & Patnaik (2020) analyzed flood-related damage in 21 flood-affected states in India and found that neither human development nor income levels significantly influenced disaster resilience, suggesting that economic growth has not effectively improved flood resilience in these states.

Investment in disaster risk reduction is crucial, but developing countries often face challenges in raising awareness and implementing effective measures (Wada et al. 2014). Fuente (2010) found that post-disaster investment often exceeds pre-disaster investment in some developing countries. Kim & Kim (2017) argued that a higher urban green ratio is associated with lower disaster damage. Limited economic resources may be a key factor slowing investment progress in developing countries.

Natural disasters are caused by the force of nature. However, human activities have a direct or indirect impact on the environment, causing natural disasters and resulting in damage, including loss of human life, injury, property and infrastructure damage, and loss of production capital.

In this study, we focus on Vietnam, which experiences a wide range of natural disasters, including storms, typhoons, floods, flash floods, landslides, and low temperatures. Using annual economic disaster damage data from 63 provinces and centrally governed cities between 2010 and 2020, we employed hybrid panel data analysis to assess the impacts of natural and human factors on total economic damage. Our study examines changes in natural disaster damage in Vietnam within the context of socioeconomic development, measured in monetary terms. This study investigated whether economic development mitigates the damage caused by natural disasters and how agricultural development contributes to this damage, considering meteorological, hydrological, and geographical factors. Furthermore, we offer national policy implications for comprehensive disaster risk management and sustainable socioeconomic development in Vietnam.

## Materials and methods

### 1. Materials

For this research, we chose Vietnam as our research site and used a provincial-level panel dataset to clarify the impact of economic development on disaster damage in Vietnam. Noy & Vu (2010), contrary to our research, examined the impact of natural disasters on economic development in Vietnam and stated the benefits of choosing Vietnam and using regional panel data in their study. The authors pointed out that Vietnam is a valuable case study because it is "a rapidly developing emerging market, with both a viable agricultural sector and a rapidly increasing manufacturing sector." This is a welcome feature for panel data analysis because it is difficult to estimate the impact if a time-variant explanatory variable experiences little change (Allison 2009). Numerous studies in this field utilize international datasets to examine countries with vastly different economic structures, policies, institutional frameworks, and macroeconomic environments. This heterogeneity can introduce biases and make it harder to obtain precise estimates. However, provincial data within the same country generally share similar legal, institutional, and economic frameworks, reducing unobserved heterogeneity. Data collection methods, definitions, and reporting standards differ across countries, which can lead to measurement errors. However, provincial data are collected under a uniform national statistical framework, resulting in greater consistency in definitions and accuracy. These characteristics help make regional datasets reliable for estimation.

In this research, we examined two categories of factors that affect economic damage: natural causes and human causes. Because disasters in Vietnam are primarily caused by meteorological and hydrological factors, natural causes are modeled using the following variables: precipitation, terrain ruggedness index (TRI), and elevation. Human causes include gross regional domestic product (GRDP) per capita, agricultural GRDP, population, land area, forest coverage, and plantation forest area. We include a linear time trend in model estimations, coded as 1, 2, 3, …, 11, to represent each consecutive year in the dataset, to examine whether there is an increasing trend of economic damage from natural disasters in Vietnam. Some studies have argued that climate change leads to an increase in the frequency and severity of natural disasters, ultimately resulting in greater damage (Coronese et al. 2019). Population and land area have also been used in previous research. Raschky (2008) notes that, due to limited space for settlement, population growth, and urban expansion, people are moving into new areas, thereby increasing the potential for damage. Raschky (2008) considers land area as a proxy for the size of a region, while Song & Park (2019) include both population and land area as indicators representing regional characteristics. Forest coverage and plantation forest area are included because we wanted to examine whether they play roles in mitigating damage.

Panel data, also known as longitudinal data, is a set of time-series data observed on the same entities, such as countries, states, firms, or individual persons. Panel data analysis is a method for handling panel data using an econometrics approach. Panel data from 63 provinces and cities in Vietnam, spanning from 2010 to 2020, were gathered (Provincial Statistical Office (PSO), Vietnam 2013, 2016, 2020; GSO 2022). This period was chosen to avoid the need for adjustment to provincial statistics due to changes in provincial boundaries. GRDP per capita (million Dong), agricultural GRDP (billion Dong), population (persons), land area ($km^2$), forest coverage (%), plantation forest area (ha), and precipitation (mm) were sourced from the General Statistics Office (PSO 2013, 2016, 2020), Vietnam. Constant values of GRDP per capita and agricultural GRDP in 2010 were used to eliminate the impact of price changes. Accordingly, economic damage was also adjusted to 2010 values using GDP deflators sourced from the World Bank (2024). To quantify topographic heterogeneity, Riley et al. (1999) developed the TRI to measure differences in elevation from a center cell and eight directly surrounding cells. TRI provides information on local surface spatial variability and has been widely adopted (Trevisani 2023). For each province, the mean values of elevation and TRI were calculated using all cells at a 30 m resolution, based on provincial boundaries from the Database of Global Administrative Areas (GADM 2024) and elevation data from the Shuttle Radar Topography Mission (SRTM) DEM. We used R version 4.1.1 and the *raster* package to calculate elevation and TRI values (R Core Team 2024, Hijmans 2024). The basic statistics of disaster damage in monetary terms and the cause variables included in the models are presented in Table 1.

Table 2 presents the number of natural disaster events in Vietnam by type and subtype from 2010 to 2020, based on data from the Centre for Research on the Epidemiology of Disasters (CRED 2025). This summary shows that Vietnam experienced at least 78 natural disaster events between 2010 and 2020, involving four types and nine subtypes. Tropical cyclones are the most frequent disaster in Vietnam, with 37 such events over the study period, followed by riverine floods (14 events), general floods (13 events), and flash floods (7 events). In terms of types, there were 34 hydrological disasters (floods) and 41 meteorological disasters (storms). CRED (2025) also reports two major droughts over the study period. The first incident occurred from December 2015

**Table 1. Descriptive statistics of the final variables in the models**

| Variables | Unit | Observations | Min | Max | Mean | Std. dev. |
|---|---|---|---|---|---|---|
| Disaster damage | billion VND | 630 | 0 | 9,851.24 | 231.08 | 794.18 |
| GRDP per capita | million VND | 688 | 10.42 | 243.47 | 34.61 | 30.01 |
| Agri. GRDP | billion VND | 688 | 954.00 | 27,703.01 | 8,192.39 | 5,300.27 |
| Precipitation | mm | 658 | 512 | 5,805 | 1,884 | 595 |
| TRI | - | 693 | 0.89 | 11.51 | 4.51 | 3.16 |
| Elevation | m | 693 | 1.67 | 999.89 | 251.6 | 284 |

Notes: Data are at the provincial level. All VND data are values in the year 2010.

**Table 2. Number of disaster events by type from 2010 to 2020**

| Year | Climatological | Biological | Hydrological | | | Meteorological | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| | No. 1 | No. 2 | No. 3 | No. 4 | No. 5 | No. 6 | No. 7 | No. 8 | No. 9 | |
| 2010 | 0 | 0 | 1 | 0 | 3 | 0 | 3 | 0 | 0 | 7 |
| 2011 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 1 | 0 | 5 |
| 2012 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 0 | 4 |
| 2013 | 0 | 0 | 0 | 0 | 6 | 0 | 4 | 0 | 0 | 10 |
| 2014 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 3 |
| 2015 | 1 | 0 | 1 | 1 | 0 | 0 | 2 | 0 | 0 | 5 |
| 2016 | 1 | 1 | 0 | 3 | 1 | 1 | 3 | 0 | 0 | 10 |
| 2017 | 1 | 0 | 1 | 4 | 0 | 0 | 4 | 0 | 0 | 10 |
| 2018 | 0 | 0 | 2 | 1 | 0 | 0 | 4 | 0 | 0 | 7 |
| 2019 | 1 | 0 | 1 | 4 | 0 | 0 | 2 | 0 | 0 | 8 |
| 2020 | 1 | 0 | 1 | 0 | 0 | 0 | 8 | 1 | 1 | 12 |
| Total | 5 | 1 | 7 | 13 | 14 | 1 | 37 | 2 | 1 | 81 |

Notes: No. 1: Drought; No. 2: Viral disease; No. 3: Flash flood; No. 4: Flood (General); No. 5: Riverine flood; No. 6: Storm (General); No. 7: Tropical cyclone; No. 8: Lightning/Thunderstorms; No. 9: Extra-tropical storm

to February 2017, resulting in an estimated USD 6.75 billion in total damage. Based on the 2016 average inter-bank exchange rate and the national GDP in 2016 (GSO 2018), this amounts to approximately 3% of GDP. The second drought, which lasted from July 2019 to February 2020, affected approximately 675,000 people (CRED 2025). Some natural disaster events, such as earthquakes, were not recorded because of fatality standards. Figure 1 shows the average annual disaster damage by province and city from 2010 to 2020. The North Central Coast and South Central Coast regions suffered more damage than other regions. Figure 2 shows the changes in annual disaster damage from 2010 to 2020 at the provincial level. Damages in 2013, 2017, and 2020 were significant. Within five weeks starting from 30 September 2013, three typhoons—Wutip, Nari, and Haiyan—struck the central region of Vietnam, causing severe damage (Bocchini 2014, CRED 2025). Typhoon Damrey, which occurred in November 2017, was reportedly the strongest storm to make landfall in the South Central Coast region in the past 20 years (International Federation of Red Cross and Red Cresent Societies 2025). Damrey affected 4.3 million people in nine provinces, killed 107 people, destroyed 3,400 houses, and damaged 141,100 houses (CFE-DM 2025). In 2017, floods, landslides, and the drought that began in 2015 caused severe damage. In 2020, CRED (2025) reported the largest number of disaster events in 11 years (Table 2). Floods, flash floods, landslides, drought, and storms hit Vietnam, causing serious damage (CFE-DM 2025).

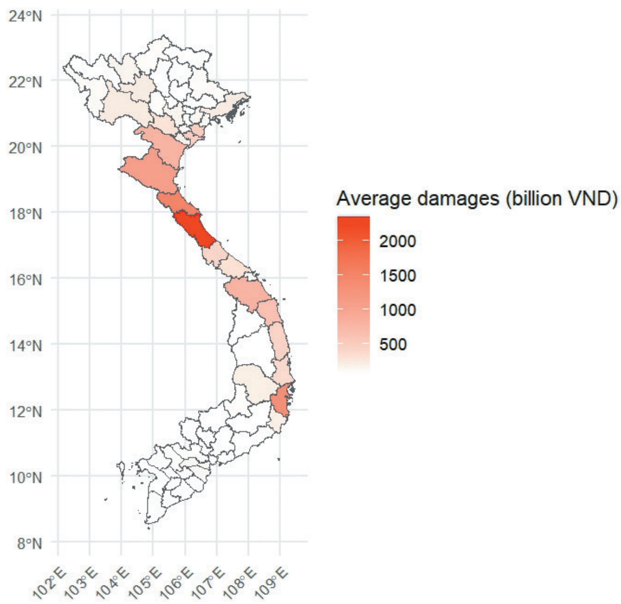


**Fig. 1. Provincial average annual disaster damage from 2010 to 2020**

Source: Created by the authors using statistics obtained from the GSO and shapefiles from GADM. Disaster damage is presented in 2010 constant values. In 2010, 1 USD = 18,613 VND (World Bank 2024). In other words, 1,000 billion VND = 53.726 million USD.

## 2. Methods

(1) Hybrid approach of panel data analysis

A hybrid approach of panel data analysis was adopted in this research. As stated in the Introduction, natural disasters can impact economic development, but our objective was to clarify the impact of economic

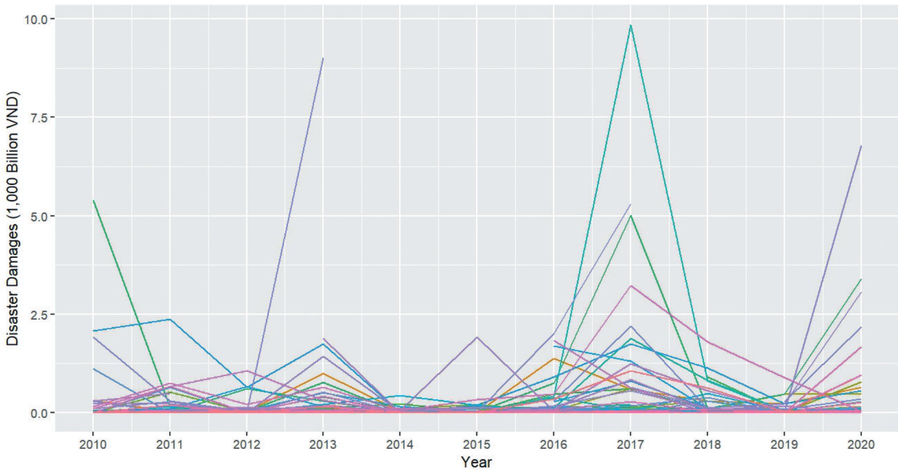


**Fig. 2. Changes in annual provincial disaster damage from 2010 to 2020**
Source: Created by the authors using statistics obtained from GSO. Disaster damage is presented in 2010 constant values.

development on disaster damage. This fact indicates a reciprocal relationship between these two factors. There are at least four cases. First, natural disasters in year *t* may cause loss of life, personal injury, damage to property, infrastructure, and productive capital and negatively impact economic development in year *t* and subsequent years. Second, due to disaster recovery efforts, natural disasters can stimulate demand and have a positive impact on economic development. Third, economic activities in year *t* cause natural disasters and cause damage in year *t* and later. Lastly, economic development can enhance the capacity to reduce disaster damage through pre-disaster investment and mitigate future disaster damage. The first two cases illustrate the impact of disasters on economic development, while the latter two highlight that of economic development on disaster damage. This research concentrates on the comprehensive impact of the latter two cases by setting an economic development variable, proxied by GRDP per capita in year *t*–1, as a predetermined variable in model estimation. This treatment avoids simultaneous relationships between economic development and disaster damage. The empirical model is as follows (Equation 1, a log-log function):

$$lddm_{it} = \alpha + \gamma T + \beta_1 lpgrdp_{i,t-1} + \beta_2 lagrdp_{i,t-1} + \beta_{others} OTHERS_{it} + \mu_i + e_{it} \quad (1)$$

where *l* stands for log; $ddm_{it}$ for disaster damage in monetary terms in year *t* for province or city *i*; *T* for the time trend; *pgrdp* for GRDP per capita; *agrdp* for agricultural GRDP; *OTHERS* for other variables, including standardized TRI, standardized elevation, log (precipitation), log (land area), log (population), forest coverage (percentage points), log (plantation forest area), and year dummy variables. $\mu_i$ is an individual-specific effect, $\alpha$ is the intercept, $\gamma$ and $\beta$s are parameters, and $e_{it}$ is an error term. Because $ddm_{it}$ can be zero, we use $lddm_{it}$, calculated as log($ddm_{it}$ + 0.5) (Yamamura 1999). Dummy variables for the year were included to treat time-fixed effects. We also added a time trend to see whether there is a trend in disaster damage.

Readers may have noticed that TRI, elevation, and land area are time-invariant variables. In panel data analysis, there are two basic, traditional approaches: the fixed effects (FE) and random effects (RE) models. In econometrics, FE models assume that the effects of time-invariant variables, observed or unobserved, do not change within a province or city. By controlling for these time-invariant variables and allowing correlations between time-varying variables and individual-specific intercepts for different provinces and cities, FE models can consistently clarify the impacts of explanatory factors. FE models are useful because many social, cultural, and other factors cannot be observed. The FE model controls time-invariant variables but does not estimate their impact. RE models assume that the individual-specific effect, or intercept, is a random variable uncorrelated with the explanatory variables. FE models only use within-individual variations, while RE models incorporate both within- and between-individual variations, thereby providing lower standard errors. Because RE models treat time-invariant variables as random, and the error structure has a specific form of correlation that violates the assumption of ordinary least squares (OLS), their effects can be estimated using generalized least squares regression (GLS). GLS corrects this by applying a quasi-demeaning transformation that adjusts for within-individual correlations (Baltagi 2021). RE models can estimate the effects of time-invariant variables, which is useful when researchers want to understand their impact. An RE model is appropriate if it does not violate the assumption of correlation between the intercept and explanatory variables.

To use the advantages of FE to control time-invariant variables and RE to estimate the effects of time-invariant variables, a hybrid approach has been proposed (Mundlak 1978, Hausman 1978). Allison (2009) further develops and explains a hybrid approach by combining FE and RE models for panel data analysis. His basic ideas can be seen in the following equation:

$$y_{it} = \alpha + \beta \boldsymbol{dx}_{it} + \gamma \boldsymbol{mx}_i + \delta \boldsymbol{z}_i + \mu_i + e_{it}, \text{for } i = 1, \ldots, n;\ t = 1, \ldots, T, \quad (2)$$

where $\boldsymbol{mx}_i = mean(\boldsymbol{x}_{it}) = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}_{it}$ and $\boldsymbol{dx}_{it} = \boldsymbol{x}_{it} - \boldsymbol{mx}_i$. Here, $\alpha$ is an overall intercept, *dx* are the deviations from individual-specific means of *x* over the years, *mx* are their individual-specific means of *x* over the years (which are time-invariant), and $\beta$, $\gamma$, and $\delta$ are coefficients. *x* and *z* stand for time-varying and time-invariant variables, respectively. Instead of using OLS estimation as in FE, a GLS estimation in a hybrid model is employed, as in RE. Allison (2009) argues that this approach provides better estimates of time-invariant variables than those in RE models by controlling them (i.e., by time-varying predictors). This approach distinguishes the effects of deviations from the mean (*dx* in Equation 2) and the effects of the mean itself (*mx* in Equation 2). Deviations from the mean (within-variations) capture short-term effects, while the mean of the variable (between-variations) captures long-term effects. We developed a hybrid model for this study, as well as a reference RE model. We conducted several tests to evaluate the models and their estimation results. Because we aim to investigate the impacts of socioeconomic development while incorporating geographic factors—such as TRI, elevation, and land area (i.e., the size of a region)—we do not use an FE model because FE models cannot accommodate time-invariant variables.

If the difference between the effects of deviations from individual means (*dx*) and the effects of variable means (*mx*) is not statistically significant in the hybrid model, separating *dx* and *mx* is unnecessary. In this case, the original variable, *x*, should be used in the model instead of *dx* and *mx* in Equation 2. If separation is unnecessary for all time-varying variables, the hybrid model is not a better choice than an RE model.

(2) Model selection and residual diagnosis

There are variations in characteristics across different provinces and cities, which makes panel data analysis necessary and meaningful. It is important to deal with these variations. For a panel dataset, if one ignores the fact that they are from different individuals (provinces and cities) and different time points (years), a pooled OLS model can be estimated. Therefore, whether an RE model is preferred to the pooled OLS model should be tested by using the Breusch-Pagan Lagrange multiplier test for individual effects. When individual effects are significant, an RE model is preferred over the pooled model. Other tests, including those for time effects, serial correlation, and heteroskedasticity, were also implemented.

Residuals from the best model were also examined. Residuals in a regression model represent the portion that the explanatory variables cannot explain, also referred to as errors or disturbances. A pattern detected in the residuals suggests that the model estimations are biased. First, scatter plots were prepared (omitted), and correlation coefficients between residuals and explanatory variables were examined. No statistically significant correlations implied that the endogeneity issue was well treated. An autocorrelation test on residuals was implemented to check whether there is an autocorrelation in residuals. Testing for heteroskedasticity was also needed to determine whether the assumption was violated. Autocorrelation and heteroskedasticity can cause the standard errors of the coefficients to be underestimated. However, using robust standard errors (heteroskedasticity- and autocorrelation-consistent (HAC) standard errors) and related p-values helps to ensure that our statistical inferences are reliable.

We used the Akaike information criterion (AIC) to determine whether to include a variable in a model. Because RE GLS regression cannot provide the log-likelihood and therefore AIC, we used maximum likelihood (ML) estimation for RE models to select variables. ML regression showed estimation results similar to those of GLS regression. For the final models, we still used GLS regression because the estimates required robustness. When a variable did not improve model fit based on AIC criteria, especially if it was far from 5% level of statistical significance, it was deleted from the models. A log-log function form was adopted. A quadratic term for GRDP per capita was attempted, which was not a causal factor itself. However, together with GRDP per capita, the nonlinear relationship was captured. Finally, Stata software (StataCorp 2023) was used in our panel data analysis.

We implemented a Ramsey regression equation specification error test (RESET) to determine whether there was omitted variable bias or an incorrect functional form in the nonlinear functions for the linear regression model (Ramsey 1969). In the test, we added the squared and cubic fitted values to the model and tested their significance.

## Results

### 1. Test results

Table 3 shows the results of various tests. The first test, comparing a pooled model with an RE model, revealed a significant individual effect. Therefore, the pooled model should not be used because it ignores the individual effect. The second test was implemented for time effects. The test results showed that the year dummy variables were jointly significant at the 1% level in both the RE and hybrid models; therefore, they were retained. Even though cross-sectional dependence could not be tested due to missing values, adding the year dummy variables lessened their effects if present. The third test, a Wooldridge test, showed that autocorrelation was not an issue in our panel data. Eleven years is short for an annual dataset. Usually, autocorrelation can be ignored in short- and wide-panel datasets. We used HAC standard error estimation to treat autocorrelation and heteroskedasticity as shown in the fourth test, a modified Wald test, which demonstrated that the problem of heteroskedasticity existed.

### 2. Model estimation results

Model results for the RE and hybrid models are shown in Table 4. Time trend, land area, population, forest coverage, and plantation forest area were confirmed to not contribute to the model fit and were deleted from the final RE and hybrid models. Actually, these variables were not statistically significant at the 5% level. Year dummy variables remained in the models, because adding these variables together lowered AIC, and the F and Wald tests showed F-values or Chi-squares with p-values less than 0.01, rejecting the null hypothesis that the coefficients for the years were jointly equal to zero at the 1% significant level. However, the estimates for the year dummy variables are not shown in Table 4. Finally, a quadratic GRDP per capita term remained in the models because it improved the fit (i.e., a lower AIC) and was statistically significant at the 1% level.

Finally, precipitation, GRDP per capita, agricultural GRDP, elevation, and TRI were included in the models. For precipitation, the Wald test showed that the deviation from its mean and the mean had no statistically significant difference (Chi-square = 0.88, $p$-value = 0.349). Therefore, precipitation was used in the model instead of separate variables representing the deviation from the mean (*dprecipitation*) and the mean (*mprecipitation*) in the final hybrid model. GRDP per capita was expressed as *lpgrdplag* ($\log(pgrdp_{t-1})$), the logarithm of GRDP per capita in lag 1. Similarly, agricultural GRDP is expressed as *lagrdplag* ($\log(agrdp_{t-1})$). We started by estimating the versions of an RE model, which includes both *lpgrdplag* and *lpgrdplag*$^2$ and yields a lower AIC (2,820.05) than the model with *lpgrdplag* alone (2,836.31). Therefore, we included both *lpgrdplag* and *lpgrdplag*$^2$ in our RE model. We also used *lpgrdplag* and *lpgrdplag*$^2$ in the hybrid model due to the non-significant difference at the 5% level between the effects of the deviation variable and the mean variable for the first-order term (Chi-square = 2.64, $p$-value = 0.104) and a coefficient for the first order deviation term that is non-significant at the 5% level ($\beta = -1.711$, $p$-value = 0.302). When testing the difference between the deviation from the mean for agricultural GRDP (*dagrdplag*) and the mean of agricultural GRDP (*magrdplag*), we obtain a Chi-square of 3.45 and a $p$-value of 0.063, implying that it was significant at the 10% level but not at the 5% level. To examine the impact of within-individuals and between-individuals effects, we separated the agricultural GRDP into two variables: deviations (*dagrdplag*) and means (*magrdplag*) in the hybrid model.

**Table 3. Test results**

| Test items | Methods | Results | Remarks |
|---|---|---|---|
| (1) Comparing models: pooled vs RE | Breusch-Pagan Lagrange Multiplier Test | $\chi^2 = 15.05$, $p = 0.000$ | RE is preferred because unobserved heterogeneity exists. |
| (2) Testing for time effects | Wald or F test | $p < 0.001$ for RE and hybrid models | Year dummy variables are jointly significant at 1%. |
| (3) Testing for serial correlation | Wooldridge test | F (1, 60) = 2.428, $p = 0.125$ | Autocorrelation is not confirmed. |
| (4) Testing for heteroskedasticity | Modified Wald test | $\chi^2 = 3{,}358$, $p = 0.000$ | Heteroskedasticity exists. Robust estimations are needed. |

**Table 4. Results of model estimates (response variable: *lddm*)**

| Variables | RE model | Hybrid model |
|---|---|---|
| *lprecipitation* | 2.491$^{***}$ | 2.462$^{***}$ |
| | (0.517) | (0.527) |
| *lpgrdplag* | 8.254$^{**}$ | 7.721$^{*}$ |
| | (3.139) | (3.223) |
| (*lpgrdplag*)$^2$ | −1.305$^{**}$ | −1.238$^{**}$ |
| | (0.389) | (0.399) |
| *lagrdplag* | 1.407$^{***}$ | - |
| | (0.343) | |
| *zelevation* | −0.808$^{*}$ | −0.803$^{*}$ |
| | (0.332) | (0.327) |
| *zTRI* | 2.207$^{***}$ | 2.163$^{***}$ |
| | (0.404) | (0.405) |
| *dagrdplag* | - | 4.192$^{+}$ |
| | | (2.283) |
| *magrdplag* | - | 1.339$^{***}$ |
| | | (0.343) |

Notes: Signs for significance levels: $^{***}$: 0.1%; $^{**}$: 1%; $^{*}$: 5%; $^{+}$: 10%. Values in brackets are robust (HAC) standard errors. *lprecipitation*, *lpgrdplag*, and *lagrdplag* are in logarithmic form. Elevation and TRI are standardized values, and *magrdplag* = mean (*lagrdplag*); *dagrdplag* = *lagrdplag* – *magrdplag*.

*lprecipitation* was statistically significant at the 0.1% level for RE and hybrid models. Precipitation had a positive coefficient and was highly elastic. *lpgrdplag* and *lpgrdplag*$^2$ were both significant at the 1% or 5% level for both models. Positive *lpgrdplag* and negative *lpgrdplag*$^2$ showed that these functions have inverted-U shape curves. *lagrdplag* was significant in the RE model. In the hybrid model, *lagrdplag* was decomposed into *dagrdplag* and *magrdplag*. *dagrdplag* was only significant at the 10% level, but not at the 5% level ($p$-value = 0.066). The 95% confidence interval of *dagrdplag* (−0.282 to 8.666) was asymmetric around zero, indicating that the coefficient was more likely to be positive. However, given the wide range, its precise magnitude was difficult to determine. Using the 5% level criteria, the results showed that the agricultural GRDP growth within a province did not cause an increase in damage. *magrdplag* was significant at the 0.1% level. An elastic, positive, and significant *magrdplag* indicated that provinces with higher agricultural GRDP faced higher disaster damage.

These two models had similar estimations, signs, magnitudes, and significance levels, as expected for the effects of elevation and TRI. Results for elevation showed that elevation was associated with an adverse change in disaster damage. However, TRI was associated with a positive change in disaster damage.

**Table 5. Turning points in GRDP per capita and $R^2$ in models**

| Index | RE model | Hybrid model |
|---|---|---|
| Turning points | 23.63 Million VND | 22.60 Million VND |
| $R^2$ (within)(between) (Overall) | (0.138) (0.645) (0.321) | (0.138) (0.653) (0.324) |
| Adjusted $R^2$ (Overall) | 0.302 | 0.303 |

Table 5 shows the values of GRDP per capita at which damage changed direction from increasing to decreasing as GRDP per capita increased in the two models, as well as the $R^2$ of the two models. These turning points are expressed in GRDP per capita in 2010 constant values. Before GRDP per capita reached these vertices, an increase in GRDP per capita was associated with increased disaster damage. However, when GRDP per capita exceeded these turning points, an increase in GRDP per capita was associated with decreased disaster damage. According to the hybrid model, the turning point was 22.60 million VND nationally, in 2010 constant value.

$R^2$ results showed that the hybrid model had a higher value for overall variation, although there was not a significant difference between the RE and hybrid models. Given our goal of incorporating meteorological, hydrological, and geographic factors and examining both within- and between-effects, we regard the hybrid model as the most suitable choice. Because the between $R^2$ was much higher than the within $R^2$ in the RE and hybrid models, most variation occurred across provinces rather than within provinces over time, and some provinces suffered more damage than others.

## 3. Hybrid model validation

To validate the model fitting, we plotted residuals and their relationships with explanatory variables and fitted values but found no patterns (omitted). We calculated correlation coefficients between the residuals, explanatory variables, and fitted values. These coefficients are shown in Table 6. All the coefficients were close to zero, and their $p$-values were 1 or close to 1, much larger than 0.05. These results confirmed the absence of correlations.

To test the missing variable bias and miss-function form, we implemented a Ramsey RESET for the final hybrid model. The null hypothesis was yhat$^2$ = 0 and yhat$^3$ = 0. The results were Chi-square = 1.41 and $p$-value = 0.493, indicating that there was no model misspecification, no missing variable bias, and that the linear functional form was not a problem.

**Table 6. Correlations between residuals and explanatory variables**

| Variables | Corr. | *p*-value |
|---|---|---|
| Precipitation | 0.013 | 0.763 |
| Pgrdplag | −0.000 | 1.000 |
| Agrdplag | −0.000 | 1.000 |
| dagrdplag | 0.014 | 0.750 |
| magrdplag | −0.000 | 1.000 |
| Fitted value | 0.005 | 0.906 |

## Discussion

Our research aimed to examine changes in natural disaster damage in Vietnam in the context of socioeconomic development. Specifically, we investigated how economic development—proxied by GRDP per capita—and agricultural development—proxied by agricultural GRDP—affected disaster damage, while also accounting for meteorological, hydrological, and geographical factors. As Song & Park (2019) pointed out in their literature review, the same level of climatic event may cause different levels of damage due to variations in social, economic, and geographical conditions. Therefore, understanding how socioeconomic development affects disaster damage is critically important for informed budgetary decisions for national disaster risk management and sustainable development for every country.

This study examined all of Vietnam, utilizing provincial-level data. The amount of damage is the total damage caused by all natural disasters experienced in each region at the provincial level. The response variable is total damage in monetary terms, without disaggregation by disaster type. Effective national disaster risk management must address total damage across all types of disasters. Many previous studies use the total amount of damage. For example, Raschky (2008) analyzed disaster fatalities (death toll) and monetary damage (as a percent of GDP), along with the impacts of GDP, the total number of affected people, population, land area, and investment. Song & Park (2019) analyzed disaster fatalities and damage in US dollars using three socioeconomic factors: GDP, population, and land area.

Our dataset consists of annual data spanning 11 years, for 63 provinces and centrally governed cities in Vietnam, resulting in a short and wide panel structure (large N, small T). Due to the limited time dimension, conducting separate time-series regressions for each province or city would not be reliable, as short time spans hinder the identification of time-related dynamics—such as lags, trends, and autocorrelation—and limit the precision of parameter estimates. As a result, we did not conduct formal comparisons between individual regressions and panel data models using F-tests (Hsiao 2003). Instead, to obtain more robust and generalizable results, we employed panel data models directly, as they allowed us to pool information over multiple years while accounting for individual heterogeneity. The appropriateness of using RE panel models over pooled regressions was supported by formal hypothesis testing, as shown in Table 3. The decomposition of the agricultural GRDP variable into its mean and deviation from the mean in the hybrid model was also supported by the results of the F-test.

This research analyzed the influences of human factors such as GRDP per capita and agricultural GRDP factors affecting natural disaster damage by including meteorological, hydrological, and geographic causes, such as precipitation, TRI, and elevation, which distinguishes it from many previous studies. Precipitation is an important meteorological and hydrological indicator, and annual precipitation was used in this study. There is room for improvement if panel data on short-term maximum precipitation and annual maximum wind speed become available. The FE model was not adopted because we wanted to incorporate geographic factors, which are time-invariant variables. The RE model results were presented instead, but this choice was not based on the Hausman test. The Hausman test was not used as evidence against the FE model due to concerns about its size distortion problem (Guggenberger 2010). Two models, RE and hybrid, showed similar results for precipitation and GRDP per capita in their signs and significance levels, but their magnitudes were different. We can conclude that (1) precipitation was identified as an important contributor to damage, (2) higher agricultural GRDP contributed to increased disaster damage, and (3) GRDP per capita, indicating economic development, showed an inverted-U-shape relation with natural disaster damage in monetary terms, changing from positive to negative. Such an inverted-U-shaped relationship has been found in other studies examining similar relationships (Schumacher & Strobl 2011, Raschky 2008). The turning points in GRDP per capita for the two models differed, but the one in the hybrid model was 22.60 million VND in constant 2010 values. In 2010, 17 provinces and cities had a GRDP per capita higher than this value. This number has since increased, reaching 59 provinces and cities by 2020. Of course, the concrete value of the turning point needs to be explained with caution. It presents the situation at the national level and is not intended to be applied to individual provincial or regional levels.

The hybrid model showed that provinces with higher agricultural GRDP tended to suffer higher damage, as indicated by the coefficient of the mean of the agricultural GRDP variable. This coefficient, 1.339 (significant at the 0.1% level), is greater than 1, indicating an elastic impact. However, within-province agricultural development, as shown by the coefficient of deviation from the mean variable (*dagrdplag*) —4.192—did not show a significant association at the 5% statistical level. The asymmetric 95% confidence interval of deviation from the mean variable around zero suggests that the coefficient was more likely to be positive than negative, indicating that the development of agriculture in a province or city is probably associated with increasing damage. The results concerning agricultural GRDP in the RE model also confirmed this point. According to Allison (2009), the coefficient on the within-individual component (e.g., *dagrdplag*) is often interpreted as a short-run effect, reflecting how deviations from an individual's average influence the outcome. In contrast, the between-individual component (e.g., *magrdplag*) captures the long-run association, reflecting differences across individuals in their average values over time. As 4.192 is higher than 1.339, these findings suggest that temporary changes in agricultural GRDP may have a stronger immediate impact on disaster damage than structural, long-term differences between individual regions. In other words, the difference in regions is a structural and long-term issue; therefore, mitigating damage while promoting agricultural development within each region should be the focus for businesses and policymakers.

This finding, which highlights the pressure from agricultural development to increase disaster risk, aligns with the understanding that both crop farming and the aquaculture industry are more susceptible to natural disasters due to their greater exposure to natural elements. This also implies that regions with a high proportion of agriculture in the economy are less resilient to natural disasters. These findings offer valuable insights for policymakers for sustainable economic development. Considering the results regarding GRDP per capita and the limited capacity of residents with low average incomes, we recommend that rural areas with low incomes be a key focus of disaster risk reduction in national public policy. In Vietnam, maintaining food security while mitigating disaster damage remains a challenging task, and measures for disaster mitigation are vital for sustainable agricultural development.

Several other variables were tested in the models but later removed. First, no significant time trend was observed during our research period (2010-2020), despite plotting current values suggesting an increasing trend. To account for price changes, we used constant 2010 values for disaster damage instead of current values. Second, population and land area were not found to contribute to natural disaster damage. Raschky (2008) finds that, contrary to his explanation regarding damage potential, population has a statistically significant adverse effect, but land area is not significant. In contrast, Song & Park (2019) used three explanatory variables—GDP, population, and land area—and found that all are significant, with GDP and population having positive effects and land area having a negative one. The negative effect for land area is contrary to our expectations, as it typically serves as an indicator of the potential size of damage from a disaster. In our findings, neither population nor land area are statistically significant, implying that the magnitude of damage is not necessarily related to the size of the region. Third, the model results did not confirm the expectation that forest coverage and plantation forest area played a clear role in mitigating disaster damage.

While forests are well known for regulating the water cycle, preventing soil erosion, and moderating both global and microclimates, young trees may not yet be strong enough to perform these functions effectively. Moreover, forest structure and condition are critical factors. Additionally, deforestation along riverbanks can increase flood risks, even if tree-planting efforts are carried out elsewhere. Forests play a crucial role in disaster risk reduction, while Chisan (forest conservation) activities, such as afforestation, slope stabilization, and erosion control, are essential for reinforcing vulnerable sites against disasters. However, the effectiveness of forests in mitigating disasters depends not only on their coverage or plantation area but also on their structure, health, and spatial distribution. Therefore, traditional indicators like forest coverage and plantation area may not accurately capture their disaster risk reduction functions. To address this gap, it is necessary to develop indicators that consider forest structure, condition, and location, particularly at the provincial level, where targeted disaster mitigation strategies can be implemented effectively. Last, exceptionally heavy rainfall can still trigger landslides, flash floods, and flooding, even in forested areas. Nevertheless, forests are expected to contribute to disaster mitigation, and the conversion of forest land in areas at risk of disaster to farmland or other uses should be carefully regulated.

This paper presents the comprehensive results of the study on the impact of economic development on natural disaster prevention and mitigation. Economic development enhances the capacity of infrastructure through national and provincial investment, and it also raises the awareness of individuals and the entire society. This, in turn, leads to disaster prevention and mitigation behaviors among individuals, various institutions, and the private sector. Even though some economic activities, such as road construction and urbanization, may potentially cause disaster damage, this research found that, as a whole, in the period from 2010 to 2020 in Vietnam, economic development contributed to natural disaster prevention and mitigation after passing some level of economic status, as proxied by GRDP per capita.

As stated in the first section, natural disasters also affect economic development. We used lag variables of GRDP per capita, and this effect was well considered. Through model validation, we found no correlations between residuals and explanatory variables. Therefore, we conclude that our results are reliable.

In conclusion, natural causes, such as precipitation, elevation, and terrain ruggedness index (TRI), play important roles. On the other hand, economic development contributes to the prevention and mitigation of natural disasters. However, agricultural development still tends to increase damage from natural disasters. Central and provincial governments are expected to play a more significant role in allocating financial resources for disaster prevention and mitigation, particularly in vulnerable rural areas.

## Acknowledgements

We thank GSO for providing us with data on natural disaster damage from 2010 to 2020. This research was financially supported by the Project on Development of Technologies to enhance the functions of forests for disaster prevention and mitigation in developing countries, Forestry Agency, Japan.

## References

Allison, P. A. (2009) *Fixed effects regression models*. Thousand Oaks, SAGE Publications, London, UK.

Bahinipati, C. S. & Patnaik, U. (2020) Does development reduce damage risk from climate extremes? Empirical evidence for floods in India. *Water Policy*, **22**, 748-767.

Baltagi, B. H. (2021) *Econometric analysis of panel data* (6th ed). Springer Nature, Cham, Switzerland.

Bocchini, F. (2014) One typhoon after another: Viet Nam in Fall 2013. http://labos.ulg.ac.be/hugo/wp-content/uploads/sites/38/2017/11/The-State-of-Environmental-Migration-2014-31-47.pdf. Accessed on 12 June 2025.

Center for Excellence in Disaster Management & Humanitarian Assistance (CFE-DM) (2025) Vietnam Disaster Management Reference Handbook. www.cfe-dmha.org. Accessed on 12 June 2025.

Centre for Research on the Epidemiology of Disasters (CRED) (2025) EM-DAT: The International Disaster Database. Université catholique de Louvain. https://public.emdat.be/data. Accessed on 11 June 2025.

Coronese, M. et al. (2019) Evidence for sharp increase in the economic damages of extreme natural disasters, *Proc. Natl. Acad. Sci.*, **116**, 21450-21455.

Cuaresma, J. C. et al. (2008) Natural disasters as creative destruction? Evidence from developing countries. *Econ. Inq.*, **46**, 214-226.

de la Fuente, A. (2010) Public responses to natural disasters: What has been done and what else could be done. *In* Fuentes-Nieva, R. & Seck, P. A. (eds.) Risk, shocks, and human development, Palgrave Macmillan, 342-373.

GADM (2024) GADM maps and data. https://gadm.org/download_country.html. Accessed on 8 May 2024.

General Statistical Office (GSO), Vietnam (2018) Statistical Yearbook of Vietnam 2017. https://www.nso.gov.vn/wp-content/uploads/2019/10/Nien-giam-2017-pdf.pdf. Accessed on 13 June 2021.

General Statistical Office (GSO), Vietnam (2022) Total disaster damage in money from 2010 to 2020. Unpublished data.

González, F. A. I. (2022) Natural disasters and economic growth: a synthesis of empirical evidence. Noesis. *J. Soc. Sci. Humanit.*, **31**, 155-173.

Guggenberger, P. (2010) The impact of a Hausman pretest on the size of a hypothesis test: The panel data case. *J. Econometrics*, **156**, 337-343.

Hausman, J. (1978) Specification tests in econometrics. *Econometrica*, **46**, 1251-1271.

Hijmans, R. (2024) *raster: Geographic Data Analysis and Modeling*. R package version 3.6-32. https://github.com/rspatial/raster.

Hsiao, C. (2003) *Analysis of Panel Data* (2nd ed.). Cambridge University Press. London, UK.

International Federation of Red Cross and Red Crescent Societies (Red Cross) (2017) *DREF Operation Update Viet Nam: Typhoon Damrey*. https://adore.ifrc.org/Download.aspx?FileId=179622. Accessed on 11 June 2025.

Kahn, M. E. (2005) The death toll from natural disasters: The role of income, geography, and institutions. *Rev. Econ. Stat.*, **87**, 271-284.

Kellenberg, D. K. & Mobarak, A. M. (2008) Does rising income increase or decrease damage risk from natural disasters? *J. Urban Econ.*, **63**, 788-802.

Kim, S. Y. & Kim, B. H. S. (2017) The effect of urban green infrastructure on disaster mitigation in Korea. *Sustainability*, **9**, 1026.

Klomp, J. & Valckx, K. (2014) Natural disasters and economic growth: A meta-analysis. *Glob. Environ. Change*, **26**, 183-195.

Loayza, N. V. et al. (2012) Natural disasters and growth: Going beyond the averages. *World Dev.*, **40**, 1317-1336.

Mundlak, Y. (1978) On the pooling of time series and cross section data. *Econometrica*, **46**, 69-85.

Noy, I. & Vu, T. B. (2010) The economics of natural disasters in a developing country: The case of Vietnam. *J. Asian Econ.*, **21**, 345-354.

Padli, J. et al. (2018) The impact of human development on natural disaster fatalities and damage: panel data evidence. *Econ. Res.-Ekon. Istraž.*, **31**, 1557-1573.

Provincial Statistical Office (PSO), Vietnam (2013) Provincial statistical yearbook for all 63 provinces and cities.

Provincial Statistical Office (PSO), Vietnam (2016) Provincial statistical yearbook for all 63 provinces and cities.

Provincial Statistical Office (PSO), Vietnam (2020) Provincial statistical yearbook for all 63 provinces and cities.

R Core Team (2024) *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing.

Ramsey, J. B. (1969) Tests for specification errors in classical linear least squares regression analysis. *J. R. Stat. Soc., Ser. B*, **31**, 350-371.

Raschky, P. A. (2008) Institutions and the losses from natural disasters. *Natl. Hazards Earth Syst. Sci.*, **8**, 627-634.

Riley, S. J. et al. (1999) A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity. *Intermt. J. Sci.*, **5**, 23-27.

Schumacher, I. & Strobl, E. (2011) Economic development and losses due to natural disasters: The role of hazard exposure. *Ecol. Econ.*, **72**, 97-105.

Skidmore, M. & Toya, H. (2002) Do natural disasters promote long-run growth? *Econ. Inq.*, **40**, 664-687.

Song, Y. S. & Park, M. J. (2019) Development of damage prediction formula for natural disasters considering economic indicators. *Sustainability*, **11**, 868.

StataCorp (2023) Stata statistical software: Release 18. College Station, TX: StataCorp LLC.

Tasri, E. S. et al. (2020) The effect of economic factors on the level of disasters that occur in the area of West Sumatra. *IOP Conf. Ser.: Earth Environ. Sci.*, **485**, 012110.

Trevisani, S. et al. (2023) Hacking the topographic ruggedness index, *Geomorphology*, **439**, 108838.

Wada, et al. (2014) The role of a macro-economic model for disaster risk reduction policy in developing countries. *IDRiM J.*, **4**, 12-29.

World Bank (2024) https://data.worldbank.org/. Accessed on 10 May 2024.

Yamamura, K. (1999) Transformation using (x + 0.5) to stabilize the variance of populations. *Res. Popul. Ecol.*, **41**, 229-234.