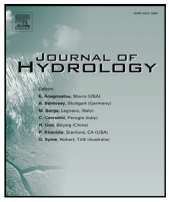
Journal of Hydrology 628 (2024) 130442

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Research papers

Uncertainty in flood risk assessment of linear structures: Why correlation matters

Erlend Briseid Storrøsten *, Luca Piciullo, Farrokh Nadim, Unni Eidsvig

Norwegian Geotechnical Institute, P.O. Box 3930 Ullevål Stadion, Oslo, 0806, Norway

ARTICLE INFO

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Ing Jie Chen, Associate Editor.

Dataset link: https://github.com/norwegian-geotechnical-institute/gp-flood-damage-aggregator

Keywords:
Uncertainty analysis
Flood damage
Risk assessment
Gaussian processes
Damage functions
Linear infrastructures

ABSTRACT

Estimation of the risk posed by inland fluvial floods to critical linear transportation infrastructures requires quantifying the damage inflicted by flooding on roads and railways. The estimated risk is often quantified in terms of the Expected Annual Damage [EAD] of the given linear structure. The aim of this article is to shed light on the estimation of this number and its uncertainty. The damage function models, which describe the degree of damage given flood intensity, represent a source of uncertainty. In this article we argue that this uncertainty needs to be embedded with a correlation structure in order to be evaluated for aggregated values. To this end, a novel methodology based on the use of Gaussian Processes to model the uncertainty associated with the damage function is introduced. Assuming a spatial correlation structure, parameterized by the so called *decorrelation length scale*, this framework is applied to estimate the EAD to roads in Portugal. The study shows that the application of an appropriate decorrelation length scale is decisive to the estimated uncertainty of the EAD. The proposed methodology may also be applicable to other types of hazards, if represented by some hazard intensity parameter associated with a damage function.

1. Introduction

Data from EMDAT (2017) clearly show that the number of disasters caused by natural hazards and the associated economic damages have increased dramatically in the last three decades. This is mainly due to the increase in number and intensity of weather-related events and intense urbanization in areas that are exposed to natural hazards. A robust and resilient transportation infrastructure is important for the critical functions in a society, contributing to its economic and social developments. Weather-related hazards such as extreme temperatures, storms, intense precipitation, floods, erosion, landslides, and forest fires pose significant risks to functionality of transportation networks. Negative impacts include accidents, damage to infrastructure or to its components, delays and malfunctioning of the transportation network. Climate change is anticipated to lead to an escalation of such negative impacts if no countermeasures are taken (Alfieri et al., 2015; Doll et al., 2014). Assessing and managing these risks means that the need for reliable techniques in damage assessment is more urgent than ever before.

Recently, considerable effort has been devoted to the assessment of damage induced by flooding to linear structures such as roads and railways on both global and national scales (Van Ginkel et al., 2021; Koks et al., 2019; Santamaria et al., 2021; Eidsvig et al., 2021). In spite of recent developments, this still remains a challenging task, subject to considerable uncertainty. While quantification of this uncertainty is of importance both to risk modelers and decision-makers, reliable quantification of uncertainties usually depends on the introduction of new and more complex modeling frameworks.

For a detailed analysis of local adverse effects of flooding on an asset, specific simulations can be performed, for example modeling of erosion of an embankment (Tsubaki et al., 2016) or erosion of a bridge foundation (Tanasić et al., 2013). However, for consequence assessment of a larger area, damage functions, which represent an average degree of damage to an asset given the intensity of the hazard, are often used. Damage functions (also called fragility functions or physical vulnerability functions) express the damage degree or probability of damage as a function of the intensity of the hazard and the structural resistance of an asset to the loads/intensity. The structural damage is often described as the degree of loss on a dimensionless 0–1 scale (Eidsvig et al., 2021).

In general, uncertainties are associated both with intensity parameters (e.g., depth, velocity, discharge, duration, and impact pressure in the case of inland flooding) and physical characteristics of the asset (material, design, maintenance history, etc.). Uncertainties related to the distribution of flood intensity parameters stems from

* Corresponding author.
E-mail address: erlend.briseid.storrosten@ngi.no (E.B. Storrøsten).

https://doi.org/10.1016/j.jhydrol.2023.130442
Received 12 September 2022; Received in revised form 14 July 2023; Accepted 20 September 2023
Available online 15 November 2023
0022-1694/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

many sources (Beven and Hall, 2014; Teng et al., 2017). Despite being an active area of research for decades, high resolution accurate flood modeling remains a significant challenge subject to both theoretical and computational challenges. As such, a range of flood modeling techniques exist, and the approach is typically chosen based on the application under consideration (Teng et al., 2017; Bulti and Abebe, 2020). Consequently, uncertainty, sensitivity to model parameters and robustness is an integral part of the field and a subject of active research (Stephens and Bledsoe, 2020; Guerrero et al., 2013; Domeneghetti et al., 2013). Quantification of this uncertainty, which is commonly based on a Monte Carlo framework, is usually associated with high computational costs, and still remains unfeasible for many important applications. Recent approaches to deal with the problem includes the use of statistical emulators (Donnelly et al., 2022; Yan et al., 2021; Kabir et al., 2020). Let us also mention the role of climate change and its impact on precipitation as another major source of uncertainty (Arnell and Gosling, 2016; Asadieh and Krakauer, 2015; Liu et al., 2023).

The work presented herein focuses on the uncertainty in damage function(s) and its effect on the estimated risk. This source of uncertainty can be seen as the lack of knowledge and information about the physical properties of the asset, which affects its resilience to flood-induced damages, or the uncertainty associated with flood intensity parameters not accounted for by the damage function. From a modeling perspective, this uncertainty may be quantified in terms of an unknown factor henceforth referred to as *residual vulnerability*. Despite its fundamental role in the assessment of risk, uncertainty associated with damage functions, is often addressed in a less systematic and more ad hoc manner.

In general, in risk assessment of transportation infrastructure, one is not interested in the pointwise damage, but rather aggregated values such as the EAD over a given geographical region, or maybe EAD for a specific highway. To propagate the pointwise uncertainty of the damage function to assess the uncertainty in aggregated values, it is necessary to embed the uncertainty with a correlation structure, e.g., impose a correlation structure on the residual vulnerability of the assets. There exists a range of different approaches to model the uncertainty associated with damage functions. A common method is to describe the range of possible outcomes, i.e. by assigning the value with so called uncertainty bands (Eidsvig et al., 2014; Jongman et al., 2012; Westerhof et al., 2022). Computations based on uncertainty bands often involves the use of heuristic distributions, like the triangular distribution, or the PERT distribution. This common framework lacks a simple well defined way of embedding spatial correlations. Consequently, most damage assessments based on such descriptions do not embed any kind of spatial correlation except the one imposed by the spatial resolution of the computational procedure. A second method to model uncertainty, which is standard in statistical modeling, is to include noise with a given parametric distribution. Given that the noise is normally distributed, this second type of models allow for the embedding of a spatial correlation structure through the use of Gaussian Processes.

While the correlation, of the residual vulnerability at two different locations, could in principle depend on the similarity with respect to properties such as construction type, ground conditions and level of maintenance we assume that the correlation is only dependent on spatial distance. This is partially justified as the mentioned properties are also spatially correlated. Another physical effect leading to spatial correlations is the reduced resilience of a damaged structure. To investigate how the spatial correlation of the residual vulnerability impacts the estimated uncertainty of the EAD, the proposed model is applied to estimate the EAD to roads in Portugal using different decorrelation length scales (See Eq. (2.3)).

Although there do exist some studies that might shed some light on the spatial correlation, e.g. Kellermann et al. (2015), we have not been able to find studies specifically addressing the assessment of the spatial variation of the damage given flood intensity parameters. However, there do exist alternative approaches to the embedding of the spatial correlation. Let us in particular mention the use of the moving split-window technique in Chen et al. (2019). See also the three types of dependence evaluated in Egorova et al. (2008).

The uncertainty related to flood intensity parameters is not considered in this work. However, the framework developed herein is a framework for modeling the uncertainty associated with the damage function, which may naturally be applied in conjunction with other approaches to include the uncertainty associated with flood intensity parameters. The application example within this paper focus on inland flooding. However, the core of the scope and novel methodology proposed in this paper is universal to the hazard type and may be considered as long as one has (i) the spatial distribution of some hazard intensity parameter and (ii) a relation between the hazard intensity parameter and the functional or structural loss of the asset (damage function). In such cases, modeling assumptions regarding the functional form of the damage function, and the suitability of modeling the residual vulnerability using a Gaussian Process are the key conditions which determines the applicability of the proposed methodology.

2. Material and methods

2.1. Estimating damage: The computational framework

To estimate EAD to a road segment due to flooding we apply a damage function/curve associating the intensity of the flood with the inflicted damage. In general, except for electrical installations, one may expect the degree of damage to transportation infrastructure to be dependent on the duration of the flooding. However, to simplify matters, we consider maximal flood intensity I over a given time period (e.g, a year). As mentioned, in addition to flood intensity, the inflicted damage will be dependent on a range of other, unknown, factors related to the asset's resilience to flood damages. To account for these unknown effects, we include a second random field $\mathcal{E}$, associated with the residual vulnerability of the asset so that the (pointwise) damage $D \in [0,1]$ is given as a function $D = D(I, \mathcal{E})$, where $D = 0$ indicates no damage, and $D = 1$ indicates total loss (e.g., nothing left of value). In statistical terms, the field $\mathcal{E}$ is simply a representation of the residuals associated with the damage function. As such, reducing the residual field $\mathcal{E}$ by including relevant features, would improve predictions and reduce uncertainty. The total damage (TD) of a segment S is given by the line integral

$$\mathrm{TD}_S = \int_S D(I(x), \mathcal{E}(x))\, ds, \tag{2.1}$$

and measured in the unit of length. Let ε be a realization of the random field $\mathcal{E}$. The expected (annual) damage (EAD) of the segment S may be expressed as a function of the realization ε according to

$$\mathrm{EAD}_S(\varepsilon) = \mathbb{E}\left[\mathrm{TD}_S \,|\mathcal{E}\right](\varepsilon) = \int_S \int_0^\infty D(i, \varepsilon(x))\, d\mathcal{F}_I(x, i)\, ds, \tag{2.2}$$

where $\mathcal{F}_I(x,i) = P(I(x) \le i)$ is the cumulative distribution function of $I(x)$. The right hand side of (2.2) may be approximated by applying a range of flood maps associated with different return periods (see Appendix). A graphical presentation of the estimation of EAD_S as a random variable associated with Eq. (2.2) is displayed in Fig. 1. To sample realizations ε of the random field $\mathcal{E}$, we assume that $\mathcal{E}(x_1), \dots, \mathcal{E}(x_n)$ is multivariate Gaussian for any finite set of points $\{x_1, \dots, x_n\}$ with a covariance structure specified by a suitable kernel k. To pick a simple interpretable structure we apply the *exponential covariance kernel*

$$k(x, y) = \sigma_f^2 \exp\left(-\frac{|x - y|}{\ell}\right), \tag{2.3}$$

where ℓ is the *spatial decorrelation length scale* and σ_f is the standard deviation of the field.

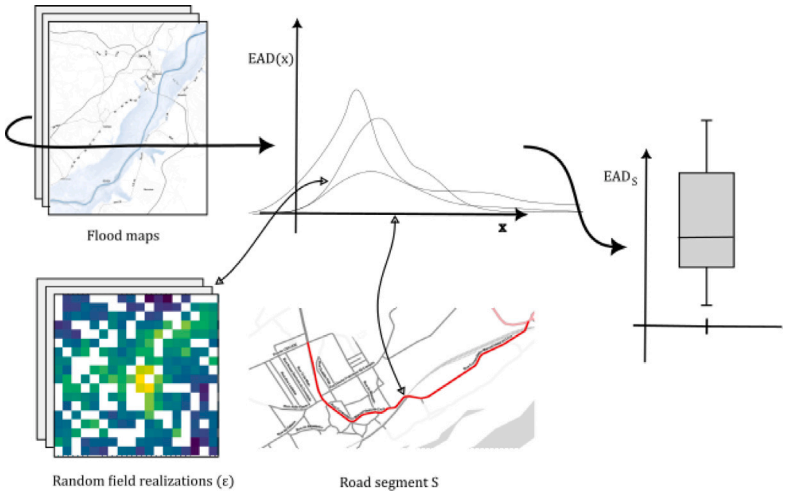


Fig. 1. The computation of expected annual damage according to Eq. (2.2). Each curve at the center represents the expected annual damage per meter at location x (EAD(x)) for a given realization of the random field $\mathcal{E}$.

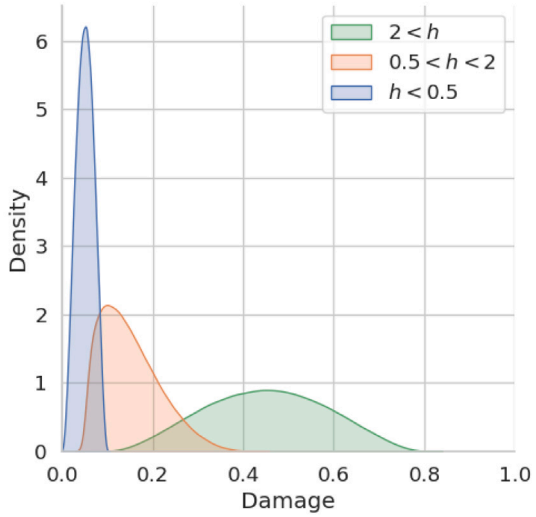


Fig. 2. Distributions of damage sampled according to the PERT distribution using the bounds from Table 1 for the dynamic flood regime i.e, $v > 1$.

Table 1
Range of the induced damage on roads as a function of depth h [m] and velocity v [m/s].

Velocity	$v < 1$			$v > 1$		
depth	$h < 0.5$	$0.5 < h < 2$	$2 < h$	$h < 0.5$	$0.5 < h < 2$	$2 < h$
Minimum	0.0000	0.001	0.05	0.0000	0.050	0.10
Mode	0.0001	0.010	0.10	0.0500	0.100	0.45
Maximum	0.0500	0.050	0.20	0.1000	0.450	0.80

2.2. A continuous spatial stochastic damage function

Next, we assess the problem of specifying a damage function D. Suppose we have a set of observed values of the flood depth (height) h_i, velocity v_i and associated damage d_i for $x_1, \dots, x_N$ different spatial locations. The problem of specifying D is then a typical regression problem. First we note that there are certain properties that the damage function should satisfy:

1. D has to take values in $[0, 1]$.
2. D has to be monotone in both parameters.
3. D is zero when height and velocity are zero.

Property 1 is not suitable for a linear regression and Property 3 is not suitable with regards to additive noise and so we have to make some adjustments in order to arrive at a suitable model. To ensure Property 1 we transform the target according to

$$d_i = \frac{l_i}{1 + l_i} \quad \Leftrightarrow \quad l_i = \frac{d_i}{1 - d_i}. \tag{2.4}$$

Then we try to establish a linear model I predicting the transformed target l as a function of h, v and possibly some nonlinear terms hv (momentum) and hv^2 (kinetic energy). To make sure that zero is being mapped to zero we set the intercept to zero, cf. Property 3. Making sure coefficients are non-negative and using a multiplicative non-negative noise it follows that properties 1,2 and 3 are satisfied. To be more specific, we assume the following function form for l:

$$l_i = I(h_i, v_i)e^{\varepsilon_i}, \text{ where } I(h, v) = \beta_1 h + \beta_2 v + \beta_3 hv + \beta_4 hv^2. \tag{2.5}$$

To fit the model I we minimize $\sum \varepsilon_i^2$, cf. Section 2.2.1. Given that ε_i is sufficiently close to normal, we may try to fit a Gaussian Process $\mathcal{E}$ by maximizing the likelihood of $\mathcal{E}(x_i) = \varepsilon_i$. It follows that we arrive at a functional form for the damage function $D = D(I, \mathcal{E})$ given by

$$D(I, \mathcal{E}) = \frac{Ie^{\mathcal{E}}}{Ie^{\mathcal{E}} + 1} \tag{2.6}$$

As we will explain below, we will not be able to fit the covariance structure of $\mathcal{E}$ due to lack of spatially structured data.

2.2.1. Fitting the damage function

To assess a suitable damage function D we have relied on vulnerability relations available in the literature. Vulnerability relations are used to express the degree of physical loss/material damage to an asset. The Swiss Road Authorities provides guidelines for assessment of material damage of roads caused by flooding (Bernard et al., 2012; Oberndorfer et al., 2020). The guidelines divide the flood conditions into six different categories according to depth and velocity. The damage degree of the road segment is specified as a range, indicated by a minimum, a mode (i.e, the most likely value) and a maximum associated with each category of flood conditions. The values and categories are presented in Table 1. The dynamic part of the table ($v > 1$ m/s) is taken from (Oberndorfer et al., 2020, A6), while the static part ($v < 1$ m/s) of the damage table was chosen in agreement with the values found in Bernard et al. (2012, p. 81).

In order to obtain a continuous stochastic damage function D on the form described by Eqs. (2.6) and (2.5), training data were sampled based on the bounds from Table 1. To this end we applied the PERT distribution, which is conveniently defined by the minimum, mode and maximum value. Fig. 2 shows the densities of the PERT distribution

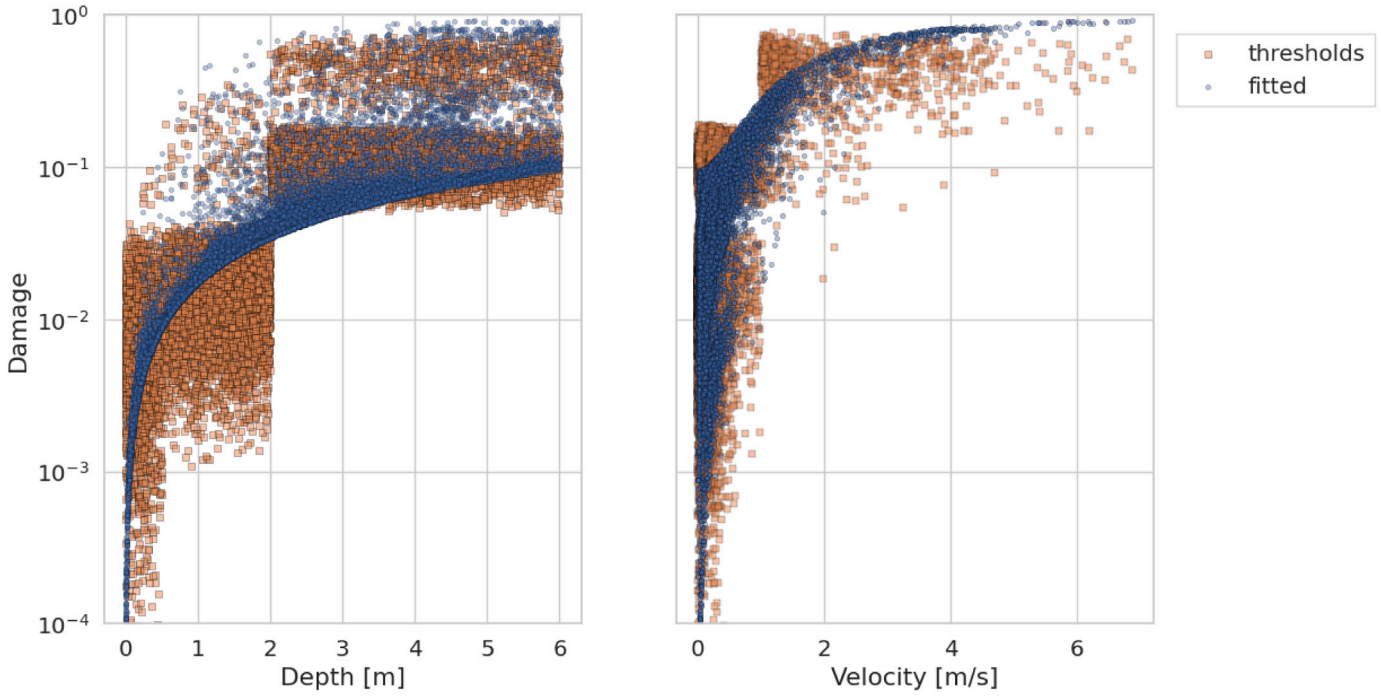


Fig. 3. Predictions of damage based on the continuous model (fitted) compared with sampled values based on Table 1 (thresholds) as a function of depth and velocity respectively.

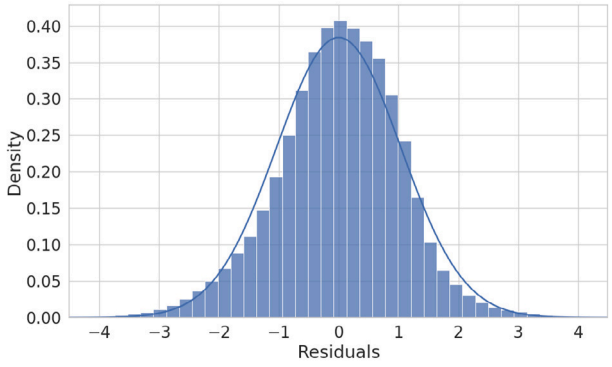


Fig. 4. Scaled histogram with fitted normal distribution of the residuals ϵ_i associated with the fitting of the continuous damage model to the sampled values from Table 1.

applied to the three different categories $0 < h < 0.5$, $0.5 < h < 2$ and $2 < h$ for the case $v > 1$ from Table 1. Comparing the PERT distribution with a triangular distribution, it puts some more weight on the mode value. This is clearly visible for the case $(0.5 < h < 2)$. The depth h_i and velocity v_i values were sampled from a flood map with a 1000 year return period covering Portugal, cf. Section 2.3. To fit the function based on depth values relevant to the flooded areas, the data set was truncated at $h_i > 6$. The associated damage values d_i was evaluated by sampling from the PERT distribution according to Table 1 as described above. In order to fit the coefficients we minimized the mean square loss $\Sigma_i \epsilon_i^2$. Comparing the loss of different models, we decided to reduce the flexibility and ensure positive coefficients by setting $\beta_3 = 0$. For the remaining parameters we arrived at

$$\beta_1 = 0.0176,\ \beta_2 = 0.0026,\ \beta_4 = 0.0476. \tag{2.7}$$

In Fig. 3, values sampled according to Table 1, labeled thresholds, are compared with the fitted values of the continuous model. As expected the continuous model impose a high degree of continuity with respect to depth and velocity. It is worth noting that the fitted values of the continuous model take on more extreme values at extreme flood regimes. In particular, for high velocity values (i.e., $v > 2\,\mathrm{m/s}$) the continuous model predicts higher damage values. This does not appear unnatural from a physical perspective, cf. Kreibich et al. (2009).

A histogram displaying the density of the associated residuals ϵ_i are shown in Fig. 4. We note that the density is close to normal except from being slightly more peaked at zero and a little skew. The standard deviation is estimated to 1.03. Is it justified to use the standard deviation of these residuals for the standard deviation σ_f (See Eq. (2.3)) of $\mathcal{E}$? The residuals are both due to the discrepancy between the continuous model and the thresholds in Table 1 and the variation imposed by the uncertainty represented in the damage table. However it does appear rational that the artificial big jumps present in the damage table may also be accounted for in terms of uncertainty in the continuous model. In Fig. 5 values sampled from the continuous model using the estimated variance of the residuals are compared with values sampled from the threshold model. The continuous model appears to replicate the uncertainty in a satisfactory way, although maybe less so regarding the more extreme values.

2.3. Estimating damage to roads in Portugal

As part of the Safeway project (Santamaria et al., 2021), the above framework was implemented and applied to a set of flood maps covering Portugal. Fig. 6 shows a diagram of the implemented computational flow.

The flood maps were downloaded through the webpage (SNIAmb, 2022). The data set consists of three different return periods 20, 100 and 1000, each with associated intensity parameters velocity and depth. The flood maps have a spatial resolution of 30 m by 30 m.

The data on the linear infrastructures (i.e., roads) were downloaded from the open access web platform Open Street Map database. The OpenStreetMap (OSM) project was founded in the United Kingdom in 2004 and is aimed at creating a free, world-wide geographic data set. The focus is mainly on transport infrastructure (streets, paths, railways, rivers), but OpenStreetMap also collects a multitude of points

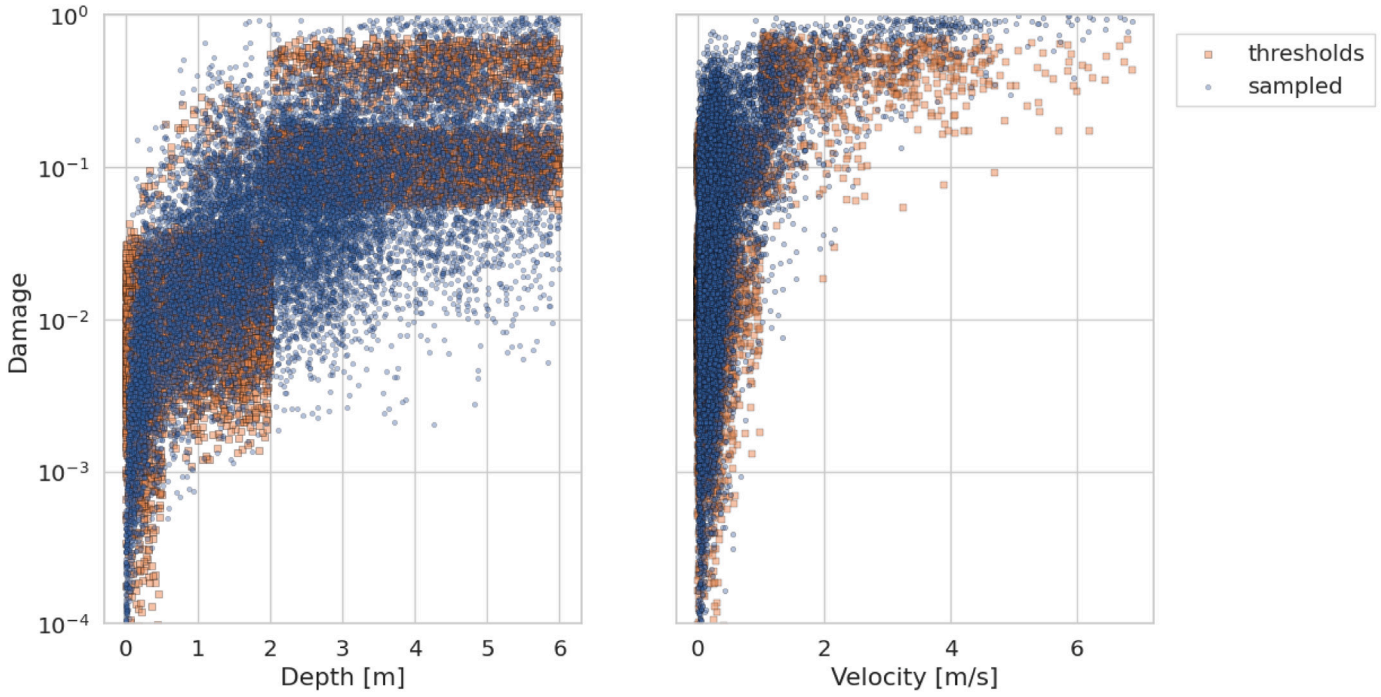


Fig. 5. Sampled values of damage based on continuous model with noise (sampled) compared with sampled values based on Table 1 (thresholds).

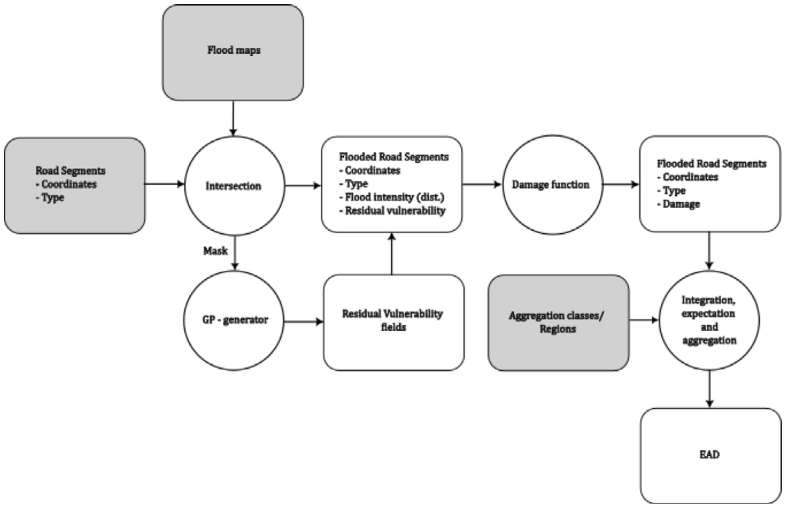


Fig. 6. A schematic drawing of the implemented computational flow associated with the calculation of EAD according to the framework described in Section 2.1. The boxes represents data objects while the circles represent operations. The shaded boxes represents input data.

of interest, buildings, natural features and land use information, as well as coastlines and administrative boundaries (Geofabrik GmbH, 2020). The data are collected by project members using their GPS devices and entered into the central database. The OSM can be considered one of the most accurate open access databases on linear infrastructure. Some papers stated that OSM, was ca. 80% complete in January 2016 (Meijer et al., 2018; Barrington-Leigh and Millard-Ball, 2017). The OSM database classifies roads into 5 types: highway, primary, secondary, tertiary, other roads. Highways comprise all major highways and trunk roads. Primary roads are represented by all major regional roads. Secondary roads as all major provincial and sub-national roads. Tertiary roads are important local roads, often linking secondary or primary roads with each other. In this study we have considered roads classified as highway, trunk, primary, secondary and tertiary. The roads classified as "other" have not been included in the analyses. Road and railway segments classified as bridge in the attribute table, have been excluded in the analyses. On the Geofabrik download server OSM data are collected in large files, called extracts, organized by region (Geofabrik GmbH, 2020). Processing of the OSM-extracts downloaded from Geofabrik was carried out using the Osmium Library (Topf, 2022).

To evaluate the damage over road segments we needed to sample realizations of the random field $\mathcal{E}$. To this end we applied the exponential covariance kernel defined in Eq. (2.3). Sampling from a large random field is computationally expensive. To minimize the computational expense, we made sure to only sample values along the roads exposed to flooding and use a pixel size of 200 m by 200 m. In Fig. 7 two samples of the random field $\mathcal{E}$ with different decorrelation length scales are displayed. It is clear from the figure that spatially close points tend to have similar values. As expected, the sample generated using a spatial decorrelation length scale of 1000 m displays much slower variation

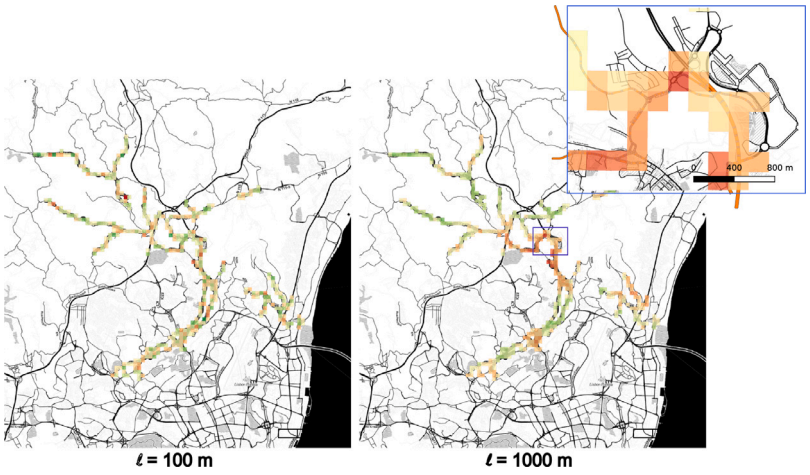


Fig. 7. Detail of two sampled random fields with different decorrelation length scale ℓ over a road network. The random fields are displayed by a green to red color scale representing values in the range $(-3, 3)$. At the center of the maps are the city of Loures.

than the sample generated with a spatial decorrelation length scale of 100 m. Note that the low resolution of the sampled field implies a slightly higher correlation between close points in the calculations compared with the mathematical model. This bias is in particular affecting the case $\ell = 100$ m. For each segment, damage was evaluated at each coordinate as downloaded from the OSM database. While the distance between the coordinates of each segment vary considerably, most are below 100 m. Consequently, in addition to a finer resolution on the sampled field, it would be necessary to refine the subdivision of the segments by adding additional coordinates to study the effect of decorrelation length scales below 100 m.

3. Results and discussion

Fig. 8 displays box-plots of expected annual damage (EAD) measured in meter for Portugal (8(a)) and the District of Santarém (8(b)). EAD is estimated using different decorrelation length scales for the residual vulnerability field $\mathcal{E}$ and classified according to road type. A considerably larger uncertainty is associated with a larger decorrelation length scale. Further, the distribution is increasingly skew. While the mean remains stable, the mode tends to decrease to compensate for an increasingly heavier tail.

Regarding the difference between the levels of aggregation, the dependency on the decorrelation length scale ℓ in the chosen range is more pronounced for the District of Santarém than for Portugal. The District of Santarém has a maximal diameter of about 140 km, indicating that for $\ell = 10$ km we may not expect a large averaging effect. This is in particular true as the most affected areas are relatively concentrated. Portugal has a maximal diameter of about 600 km and one should expect a substantial averaging effect even for $\ell = 10$ km, making the uncertainty less dependent on the length scales in the selected range. This is clearly the case when considering relative uncertainty, i.e., scaled by the mean or mode.

Some understanding of the impact of the scale of the spatial decorrelation length on the estimated uncertainty of the EAD may be derived from a formal limit argument. In particular, why an increased spatial decorrelation length is associated with a larger uncertainty. Consider a network of roads in a region R and split the network into N "small" segments S_i of equal length. Let $x_i \in S_i$ for $1 \le i \le N$. Then

$$\mathrm{EAD}_R = \mathbb{E}\left[\mathrm{TD}_R \,|\mathcal{E}\right] \approx \sum_{i=1}^{N} \frac{\mathcal{L}}{N} \underbrace{\int_0^\infty D(i, \mathcal{E}(x_i)) d\mathcal{F}_I(x, i)}_{D(x_i)}, \tag{3.1}$$

where $\mathcal{L}$ denotes the total length of the structure, $D(x) \in [0, 1]$ is the pointwise expected annual damage at location x and N is sufficiently large. It follows that the variance may be approximated by

$$\begin{aligned} \mathrm{Var}\left(\mathrm{EAD}_R\right) &\approx \frac{\mathcal{L}^2}{N^2} \sum_{i,j=1}^{N} \mathrm{Cov}\left(D(x_i), D(x_j)\right) \\ &= \frac{\mathcal{L}^2}{N^2} \left(\sum_i \mathrm{Var}\left(D(x_i)\right) + \sum_{i \neq j} \mathrm{Cov}\left(D(x_i), D(x_j)\right) \right). \end{aligned} \tag{3.2}$$

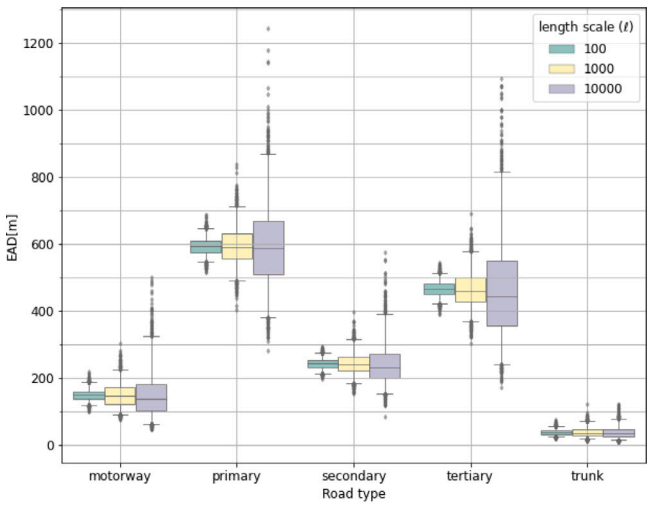
Note that the first sum grows linearly in N while the second sum grows as N^2. Hence, if $D(x_i)$ is independent of $D(x_j)$ whenever $x_i \neq x_j$ the variance would vanish as $N \to \infty$, effectively reducing the estimated uncertainty to zero. A fact known as *the law of large numbers*. This would be the effect of modeling $\mathcal{E}$ as a Gaussian white noise. A modeling approach achieved by sampling independent ε_i for each segment. Observe that the estimated uncertainty in this case will be heavily dependent on the subdivision of the network quantified by N. Increasing the scale of the spatial decorrelation length will increase the terms in the second sum and hence also the variance of EAD_R and effectively increase the estimated uncertainty also in the limit $N \to \infty$.

4. Conclusion

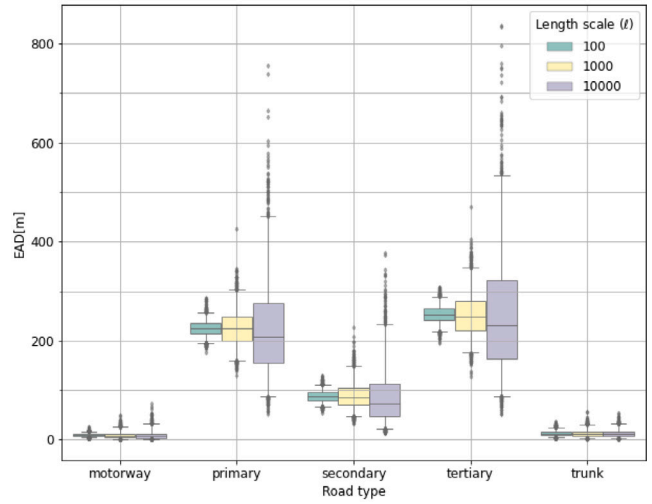
In this article we have established a continuous spatial stochastic model for damage to linear structures due to flooding. The fundamental parameter introduced in the model through the covariance kernel (2.3) is the so-called spatial decorrelation length scale. The spatial decorrelation length scale determines the degree of correlation between the residual vulnerability of the linear structure measured at two different locations and their spatial distance.

Applying this model to assess the expected annual damage (EAD) to roads in Portugal we have observed that the relative size of the decorrelation length scale is decisive to the relative uncertainty of the EAD. That is, if the uncertainty of the damage function is associated with a short decorrelation length scale compared with the region of interest, applying a highly uncertain damage function still yields a relatively certain estimate of the EAD. Conversely, a small uncertainty in the damage function, associated with a large relative decorrelation length scale may lead to a considerable uncertainty in the EAD. Even though the parameter is important for uncertainty estimates, it has received little attention in the literature.

Due to the lack of spatially structured damage data, we have not been able to assess the suitability of the applied model and in particular its spatial structure. The collection of damage data in combination with hydrodynamical back analysis of flooding events would enable a



(a) Portugal



(b) District of Santarém

Fig. 8. Box-plots of expected annual damage (EAD) classified according to road type. The figures are generated using 2000 samples of the random field $\mathcal{E}$ with decorrelation length scale ℓ equal to 100,1000 and 10 000 meters. Whiskers are placed at the 2.5 and 97.5 percentile.

statistical assessment of the model parameters. An appropriate treatment, allowing uncertain model parameters and the inclusion of both data and prior knowledge to assess their values, may be achieved in a Bayesian framework (Gelman et al., 2015; Jalayer et al., 2023). However, the use of prior knowledge to restrict the range of plausible values for the spatial decorrelation length scale depends on a clear understanding of the effects it represents. As mentioned before, we expect the reduced resilience to flooding associated with damaged structures to lead to spatial correlations. Another contribution stems from properties affecting the assets resilience to flood-induced damages or spatial correlations associated with hazard intensity parameters, not considered in the damage function, like duration of the flooding event. While the first effect is most likely associated with a shorter decorrelation length scale, the second effect, which may be reduced by the inclusion of relevant properties in the damage function, is likely to be associated with larger decorrelation length scales. To account for different contributions, associated with different decorrelation length scales, it is possible to replace the single kernel in Eq. (2.3) with a weighted sum, representing the contribution of the different effects.

CRediT authorship contribution statement

Erlend Briseid Storrøsten: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft. **Luca Piciullo:** Data acquisition, Writing – original draft. **Farrokh Nadim:** Writing – review & editing. **Unni Eidsvig:** Project administration, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Only publically available data has been used. The code associated with this work is available at https://github.com/norwegian-geotechnical-institute/gp-flood-damage-aggregator.

Funding

The research leading to the results in this paper has received funding from the European Community's H2020 Program MG-7-1-2017, Resilience to extreme (natural and human-made) events, Norway, under Grant Agreement number: 769255—"GIS-based infrastructure management system for optimized response to extreme events of terrestrial transport networks (SAFEWAY)". The support is gratefully acknowledged.

Disclaimer

The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the European Union. Neither the Innovation and Networks Executive Agency (INEA) nor the European Commission are responsible for any use that may be made of the information contained herein.

Appendix. Approximating expected annual damage

In Section 2.1, Eq. (2.2) the EAD of a segment S is expressed as an integral with respect to the cumulative distribution function $\mathcal{F}_I$ of the maximal (annual) flood intensity I. To approximate the integral we apply a set of flood maps associated with a range of return periods $n_1 < \cdots < n_J$. Let $\{i_{n_j}\}_{j=1}^{J}$ be the associated flood maps satisfying

$$P(I(x) > i_n(x)) = \frac{1}{n}, \tag{A.1}$$

for (almost) all x and all $n \in \{n_1, \ldots, n_J\}$ so that $i_n(x)$ is the intensity associated with the return period n at location x. By definition of the cumulative distribution function

$$\mathcal{F}_I(x, i_n) = 1 - P(I(x) > i_n(x)) = 1 - \frac{1}{n}. \tag{A.2}$$

Let us split the integral in Eq. (2.2) according to

$$\int_0^\infty D(i, \varepsilon(x))\, d\mathcal{F}_I(x, i) = \int_0^{i_{n_1}} D(i, \varepsilon(x))\, d\mathcal{F}_I(x, i) + \int_{i_{n_1}}^{i_{n_J}} D(i, \varepsilon(x))\, d\mathcal{F}_I(x, i) + \int_{i_{n_J}}^\infty D(i, \varepsilon(x))\, d\mathcal{F}_I(x, i)$$

$$=: \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3. \tag{A.3}$$

Given that the return period n_1 is sufficiently small, the associated intensity i_{n_1} is assumed to be so small that the damage is negligible. This lower limit is typically set with a reference to the assets general flood protection. Consequently $\mathcal{I}_1$ is assumed to be zero. Similarly $\mathcal{I}_3 < 1/n_J$ so that by taking n_J sufficiently large $\mathcal{I}_3$ may also be neglected. To evaluate $\mathcal{I}_2$ we apply the trapezoidal rule to obtain

$$\begin{aligned}\int_{i_{n_1}}^{i_{n_J}} D(i,\varepsilon(x))\, d\mathcal{F}_I(x,i) &= \sum_{j=1}^{J-1} \int_{i_{n_j}}^{i_{n_{j+1}}} D(i,\varepsilon(x))\, d\mathcal{F}_I(x,i) \\ &\approx \sum_{j=1}^{J-1} \frac{1}{2}(D(i_{n_j},\varepsilon(x)) + D(i_{n_{j+1}},\varepsilon(x)))(\mathcal{F}_I(x,i_{n_{j+1}}) - \mathcal{F}_I(x,i_{n_j})),\end{aligned} \tag{A.4}$$

where we note that

$$(\mathcal{F}_I(x,i_{n_{j+1}}) - \mathcal{F}_I(x,i_{n_j})) = \frac{1}{n_j} - \frac{1}{n_{j+1}}. \tag{A.5}$$

References

Alfieri, L., Burek, P., Feyen, L., Forzieri, G., 2015. Global warming increases the frequency of river floods in Europe. Hydrol. Earth Syst. Sci. 19 (5), 2247–2260. http://dx.doi.org/10.5194/HESS-19-2247-2015.

Arnell, N.W., Gosling, S.N., 2016. The impacts of climate change on river flood risk at the global scale. Clim. Change 134 (3), 387–401. http://dx.doi.org/10.1007/S10584-014-1084-5/FIGURES/3.

Asadieh, B., Krakauer, N.Y., 2015. Global trends in extreme precipitation: Climate models versus observations. Hydrol. Earth Syst. Sci. 19 (2), 877–891. http://dx.doi.org/10.5194/HESS-19-877-2015.

Barrington-Leigh, C., Millard-Ball, A., 2017. The world's user-generated road map is more than 80% complete. PLoS One 12 (8), 1–20. http://dx.doi.org/10.1371/journal.pone.0180698.

Bernard, G., Marguerite, T.M., Marc, H., Philippe, A., Urs, V., Urban, R., Arthur, S., Hugo, R., Luuk, D., Thomas, E., Josef, E., Reto, K., Hans, K., Michel, D., Hans-Heini, U., Bernhard, P., 2012. ASTRA 89001-Natural Hazards on National Roads: Risk Concept. Technical Report, URL www.astra.admin.ch.

Beven, K., Hall, J., 2014. Applied Uncertainty Analysis for Flood Risk Management. IMPERIAL COLLEGE PRESS, http://dx.doi.org/10.1142/p588.

Bulti, D.T., Abebe, B.G., 2020. A review of flood modeling methods for urban pluvial flood application. Model. Earth Syst. Environ. 6 (3), 1293–1302. http://dx.doi.org/10.1007/s40808-020-00803-z.

Chen, W., Wang, X., Deng, S., Liu, C., Xie, H., Zhu, Y., 2019. Integrated urban flood vulnerability assessment using local spatial dependence-based probabilistic approach. J. Hydrol. 575, 454–469. http://dx.doi.org/10.1016/J.JHYDROL.2019.05.043.

Doll, C., Klug, S., Enei, R., 2014. Large and small numbers: Options for quantifying the costs of extremes on transport now and in 40 years. Nat. Hazards 72 (1), 211–239. http://dx.doi.org/10.1007/S11069-013-0821-9/TABLES/6.

Domeneghetti, A., Vorogushyn, S., Castellarin, A., Merz, B., Brath, A., 2013. Probabilistic flood hazard mapping: Effects of uncertain boundary conditions. Hydrol. Earth Syst. Sci. 17 (8), 3127–3140. http://dx.doi.org/10.5194/HESS-17-3127-2013.

Donnelly, J., Abolfathi, S., Pearson, J., Chatrabgoun, O., Daneshkhah, A., 2022. Gaussian process emulation of spatio-temporal outputs of a 2D inland flood model. Water Res. 225, 119100. http://dx.doi.org/10.1016/j.watres.2022.119100.

Egorova, R., van Noortwijk, J.M., Holterman, S.R., 2008. Uncertainty in flood damage estimation. Int. J. River Basin Manage. 6 (2), 139–148. http://dx.doi.org/10.1080/15715124.2008.9635343.

Eidsvig, U.M.K., Papathoma-Köhle, M., Du, J., Glade, T., Vangelsten, B.V., 2014. Quantification of model uncertainty in debris flow vulnerability assessment. Eng. Geol. 181, 15–26. http://dx.doi.org/10.1016/j.enggeo.2014.08.006.

Eidsvig, U., Santamaría, M., Galvão, N., Tanasic, N., Piciullo, L., Hajdin, R., Nadim, F., Sousa, H.S., Matos, J., 2021. Risk assessment of terrestrial transportation infrastructures exposed to extreme events. Infrastructures 6 (11), 163. http://dx.doi.org/10.3390/INFRASTRUCTURES6110163, 2021, Vol. 6, Page 163.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., Rubin, D., 2015. Bayesian Data Analysis, third ed. Chapman and Hall/CRC, New York, http://dx.doi.org/10.1201/b16018.

Geofabrik GmbH, 2020. Geofabrik. https://www.geofabrik.de/geofabrik/geofabrik.html. Accessed: 2022-03-02.

Guerrero, J.-L., Westerberg, I.K., Halldin, S., Lundin, L.-C., Xu, C.-Y., 2013. Exploring the hydrological robustness of model-parameter values with alpha shapes. Water Resour. Res. 49 (10), 6700–6715. http://dx.doi.org/10.1002/wrcr.20533.

Jalayer, F., Ebrahimian, H., Trevlopoulos, K., Bradley, B., 2023. Empirical tsunami fragility modelling for hierarchical damage levels. Nat. Hazards Earth Syst. Sci. 23 (2), 909–931. http://dx.doi.org/10.5194/nhess-23-909-2023.

Jongman, B., Kreibich, H., Apel, H., Barredo, J.I., Bates, P.D., Feyen, L., Gericke, A., Neal, J., Aerts, J.C.J.H., Ward, P.J., 2012. Comparative flood damage model assessment: Towards a European approach. Nat. Hazards Earth Syst. Sci. 12 (12), 3733–3752. http://dx.doi.org/10.5194/nhess-12-3733-2012.

Kabir, S., Patidar, S., Xia, X., Liang, Q., Neal, J., Pender, G., 2020. A deep convolutional neural network model for rapid prediction of fluvial flood inundation. J. Hydrol. 590, 125481. http://dx.doi.org/10.1016/j.jhydrol.2020.125481.

Kellermann, P., Schöbel, A., Kundela, G., Thieken, A.H., 2015. Estimating flood damage to railway infrastructure - the case study of the march river flood in 2006 at the Austrian northern railway. Nat. Hazards Earth Syst. Sci. 15 (11), http://dx.doi.org/10.5194/nhess-15-2485-2015.

Koks, E.E., Rozenberg, J., Zorn, C., Tariverdi, M., Vousdoukas, M., Fraser, S.A., Hall, J.W., Hallegatte, S., 2019. A global multi-hazard risk analysis of road and railway infrastructure assets. Nature Commun. 10 (1), http://dx.doi.org/10.1038/s41467-019-10442-3.

Kreibich, H., Piroth, K., Seifert, I., Maiwald, H., Kunert, U., Schwarz, J., Merz, B., Thieken, A.H., 2009. Is flow velocity a significant parameter in flood damage modelling? Nat. Hazards Earth Syst. Sci. 9 (5), 1679–1692. http://dx.doi.org/10.5194/NHESS-9-1679-2009.

Liu, W., Feng, Q., Engel, B.A., Yu, T., Zhang, X., Qian, Y., 2023. A probabilistic assessment of urban flood risk and impacts of future climate change. J. Hydrol. 618, 129267. http://dx.doi.org/10.1016/j.jhydrol.2023.129267.

Meijer, J.R., Huijbregts, M.A.J., Schotten, K.C.G.J., Schipper, A.M., 2018. Global patterns of current and future road infrastructure. Environ. Res. Lett. 13 (6), 064006. http://dx.doi.org/10.1088/1748-9326/aabd42.

Oberndorfer, S., Sander, P., Fuchs, S., 2020. Multi-hazard risk assessment for roads: Probabilistic versus deterministic approaches. Nat. Hazards Earth Syst. Sci. 20 (11), http://dx.doi.org/10.5194/nhess-20-3135-2020.

Santamaria, M., Arango, E., Jafari, F., Sousa, H., the SAFEWAY consortium, 2021. Dynamic Risk-based Predictive Models. Deliverable 5.1, SAFEWAY.

SNIAmb, 2022. SNIAmb, sistema nacional de informacao de ambiente. https://sniamb.apambiente.pt/. Accessed: 2022-03-02.

Stephens, T.A., Bledsoe, B.P., 2020. Probabilistic mapping of flood hazards: Depicting uncertainty in streamflow, land use, and geomorphic adjustment. Anthropocene 29, 100231. http://dx.doi.org/10.1016/J.ANCENE.2019.100231.

Tanasić, N., Ilić, V., Hajdin, R., 2013. Vulnerability assessment of bridges exposed to scour. Transp. Res. Rec. 2360 (1), 36–44. http://dx.doi.org/10.3141/2360-05.

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., Kim, S., 2017. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environ. Model. Softw. 90, 201–216. http://dx.doi.org/10.1016/j.envsoft.2017.01.006.

Topf, J., 2022. Osmcode. https://osmcode.org/. Accessed: 2022-03-02.

Tsubaki, R., David Bricker, J., Ichii, K., Kawahara, Y., 2016. Development of fragility curves for railway embankment and ballast scour due to overtopping flood flow. Nat. Hazards Earth Syst. Sci. 16 (12), 2455–2472. http://dx.doi.org/10.5194/NHESS-16-2455-2016.

Van Ginkel, K.C., Dottori, F., Alfieri, L., Feyen, L., Koks, E.E., 2021. Flood risk assessment of the European road network. Nat. Hazards Earth Syst. Sci. 21 (3), 1011–1027. http://dx.doi.org/10.5194/NHESS-21-1011-2021.

Westerhof, S.G., Booij, M.J., Van den Berg, M.C.J., Huting, R.J.M., Warmink, J.J., 2022. Uncertainty analysis of risk-based flood safety standards in the Netherlands through a scenario-based approach. Int. J. River Basin Manag. 1–16. http://dx.doi.org/10.1080/15715124.2022.2060243.

Yan, X., Mohammadian, A., Khelifa, A., 2021. Modeling spatial distribution of flow depth in fluvial systems using a hybrid two-dimensional hydraulic-multigene genetic programming approach. J. Hydrol. 600, 126517. http://dx.doi.org/10.1016/j.jhydrol.2021.126517.