Optimal probabilistic placement of facilities using a surrogate model for 3D tsunami simulations

Kenta Tozato[1], Shuji Moriguchi[2], Shinsuke Takase[3], Yu Otake[1], Michael R. Motley[4], Anawat Suppasri[2], and Kenjiro Terada[2]

[1]Department of Civil and Environmental Engineering, Tohoku University, Aza-Aoba 6-6-06, Aramaki, Aoba-ku, Sendai 980-8579, Japan
[2]International Research Institute of Disaster Science, Tohoku University, Aza-Aoba 468-1, Aramaki, Aoba-ku, Sendai 980-8572, Japan
[3]Department of Civil Engineering and Architecture, Hachinohe Institute of Technology, 88-1 Ohbiraki, Myo, Hachinohe, Aomori 031-8501, Japan
[4]Civil and Environmental Engineering, University of Washington, 201 More Hall, Box 352700, Seattle, Washington, 98195-2700, USA.

Correspondence: Kenta Tozato (kenta.tozato.t2@dc.tohoku.ac.jp)

Abstract. This study proposes a framework that can efficiently utilize the information obtained from advanced tsunami simulation for probabilistic tsunami hazard assessment (PTHA) and investigation of optimal facility placement. A series of numerical simulations of the tsunami off the Pacific Coast caused by the 2011 Tohoku Earthquake is performed considering uncertainties of fault parameters. The simulated tsunami force acting on buildings and inundation depth are calculated in the simulations, and they are defined as tsunami hazard indices. Proper orthogonal decomposition (POD) is then applied to the simulated results to extract the characteristic spatial modes, which can be used to construct surrogate models. Monte Carlo simulations (MCS) are then performed at a low computational cost using the surrogate models. Based on the MCS results along with the concept of system failure probability, the optimal placement of facilities is probabilistically investigated with the help of genetic algorithms. The results indicate that the proposed framework enables us to determine the optimal placement of facilities applying different strategies at low computational cost while effectively reflecting the results of advanced tsunami simulations.

1 Introduction

Tsunami is a disaster that has a large impact worldwide, with the recent example of the global Tonga tsunami (Heidarzadeh et al., 2022; Kubota et al., 2022). To reduce the damage caused by tsunamis, it is essential to predict tsunamis by using numerical analysis. The numerical analysis techniques for tsunamis have developed considerably over the years, and high-accuracy hazard assessments and predictions have now become possible (Qin et al., 2018; Xiong et al., 2019). Natural disasters such as tsunamis have numerous uncertainties; hence, conducting probabilistic hazard assessments that consider such factors is important. However, advanced numerical analyses are unsuitable for probabilistic hazard assessment because a large number of calculation cases are generally required. Therefore, it is necessary to develop a framework that effectively uses the few but reliable numerical simulation results to achieve probabilistic evaluations.

Probabilistic hazard assessments of natural disasters have been studied for many years in the seismology fields. Among these, the study by Cornell (Cornell, 1968) is considered groundbreaking. Many research results have been reported on probabilistic seismic hazard analyses (PSHA) (e.g., McGuire, 1977; Ishikawa and Kameda, 1988). Furthermore, probabilistic tsunami hazard analyses (PTHA) have been established based on PSHAs; they have been proposed as a method for understanding the relationship between tsunami heights and exceedance probabilities in a specified period ; see, e.g., Geist and Parsons (2006); Annaka et al. (2007); Fukutani et al. (2015); Mori et al. (2017). Also, some studies have focused on actual regions such as the southeast Aegean Sea (Mitsoudis et al., 2012), the Makran subduction zone (Heidarzadeh and Kinjo, 2011), the Cascadia subduction zone (Park and Cox, 2016; LeVeque et al., 2016; Salmanidou et al., 2021) and Nankai Trough subduction zone (Nakano et al., 2020; Baba et al., 2022). In addition, the slip distributions have been investigated based on probabilistic approaches; see, e.g., LeVeque et al. (2016); Nakano et al. (2020); Scala et al. (2020). Moreover, numerous studies have been conducted to efficiently utilize numerical simulation by constructing surrogate models and utilizing them in PTHA. For instance, response surface-based approaches using polynomial functions (e.g., Kotani et al., 2020), the radial basis function (e.g., Gopinathan et al., 2021), and the Gaussian process regression (e.g., Salmanidou et al., 2021; Alhamid et al., 2022) have been reported. A surrogate model constructed based on the concept of the singular value decomposition is worthy of remark because spatial modes are efficiently utilized; see Fukutani et al. (2021).

Furthermore, the appropriate placement of infrastructure and evacuation facilities is important for minimizing damage caused by tsunamis. Numerous studies have used a probabilistic approach for the optimal placement of network systems and facilities by considering the uncertainties of disasters such as earthquakes or tsunamis. Such research examples include risk assessments for infrastructure system networks (Gomez and Baker, 2019; Miller and Baker, 2015), optimization of relief supply bases and their delivery (Cavdur et al., 2020a, b; Maharjan and Hanapla, 2020), optimal placement of public and evacuation facilities (Park et al., 2012; Zhang and Yun, 2019; Doerner et al., 2008), emergency medical service networks (Mohamadi and Yaghoubi, 2017), and optimal allocation of budgets for disaster countermeasures (Rawls and Turnquist, 2010).

Thus, many research examples on the probabilistic hazard assessments of tsunamis and probabilistic optimal placement of facilities exist; however, few studies have sufficiently used the information obtained from advanced numerical analysis. In this study, we applied the theory of mode decomposition using proper orthogonal decomposition (POD) (Kerschen et al., 2005) for solving such issues. The objectives of POD include extracting data characteristics or reducing data dimensions. POD is often treated as an equivalent of Karhunen–Loeve decomposition developed by Karhunen (1947) or Kosambi (1943); or principal component analysis (PCA) (Jolliffe and Cadima, 2016) developed by Hotelling (1933). POD has numerous application examples in a wide range of fields; furthermore, there are various application examples in the fields of seismic engineering and tsunami engineering. For example, Ha et al. (2008) applied POD for reducing computational costs to construct a tsunami surrogate model. LeVeque et al. (2016) and Melgar et al. (2016) applied the Karhunen–Loeve expansion to consider the fault slip distribution under various scenarios. Furthermore, Nojima et al. (2018) conducted research on the combination of mode decomposition based on singular value decomposition and numerical simulations to predict the distribution of strong motion; Bamer and Bucher (2012) used nonlinear finite element methods to construct a surrogate model, using POD for the prediction of the behavior of buildings. Moreover, Fukutani et al. (2021) constructed a surrogate model that used POD to implement

probabilistic inundation assessments. POD can extract the features of spatial and temporal distributions of risk indices; thus, it is suitable for the construction of surrogate models for disaster hazard assessment.

Given this context, in this study, we constructed a surrogate model by applying POD to advanced three-dimensional (3D) tsunami simulation results; we propose a method that uses this surrogate model to efficiently investigate the optimal placement of facilities such as infrastructure facilities, relief supply bases, and evacuation shelters based on probability theory. Research examples of probabilistic assessments of tsunami hazards using surrogate models of numerical simulations include the previously mentioned approach of using mode decomposition by Fukutani et al. (2021) and the approach of using response surfaces by Kotani et al. (2020)); however, no research examples exist in which the optimal placement of facilities has been investigated based on probability theory using a surrogate model. Therefore, in this study, we constructed a surrogate model using mode decomposition on information obtained from advanced numerical simulations. Moreover, we proposed a method that could efficiently investigate the optimal placement of facilities based on probability theory using this surrogate model.

The structure of this paper is as follows. Section 2 describes the framework and methods used in this study. In Section 3, we applied the proposed method to the 2011 Off the Coast of Tohoku Earthquake (Tsuji et al., 2011) as an application example that considers the actual damage and verified the validity of the constructed surrogate model; we also implemented Monte Carlo simulations to conduct probabilistic risk assessments. Furthermore, the Monte Carlo simulation results and a genetic algorithm were used to investigate the optimal placement of facilities, and the usefulness of the proposed method was discussed. Finally, Section 4 describes the conclusions.

2 Search method for the probabilistic optimal placement using a surrogate model

The proposed method is described in this section. In this study, a combined two-dimensional (2D) and 3D tsunami analysis was first conducted for multiple scenarios with different fault parameters. The maximum tsunami fluid force and the maximum inundation depth were adopted as the assessment indicators. Next, POD was applied to the results obtained from these analyses to extract the spatial modes of the tsunami fluid force and inundation depth; these spatial modes were used to construct a surrogate model; based on this model, the spatial distribution of the tsunami hazard index for an arbitrary scenario could be calculated with low computational cost. Furthermore, this surrogate model and uncertainty parameter fluctuation information were combined to implement a Monte Carlo simulation and compute the probability distribution of the tsunami fluid force and inundation depth at each assessment point; a threshold value was set to create an exceedance probability map. Finally, the Monte Carlo simulation results were used, and a genetic algorithm was applied to investigate the optimal facility placement. Figure 1 shows a flowchart of this study. From the next sub-section, the methods used in each portion are detailed. Of these, for the methods described in Sections 2.1 and 2.2, the methods proposed by Tozato et al. (2022) were adopted.

2.1 Tsunami simulation method

Numerical analyses were first conducted to construct a surrogate model of each tsunami hazard index in the target region. A numerical analysis that combined 2D and 3D analysis was conducted; first, a wide-area 2D tsunami analysis was conducted,

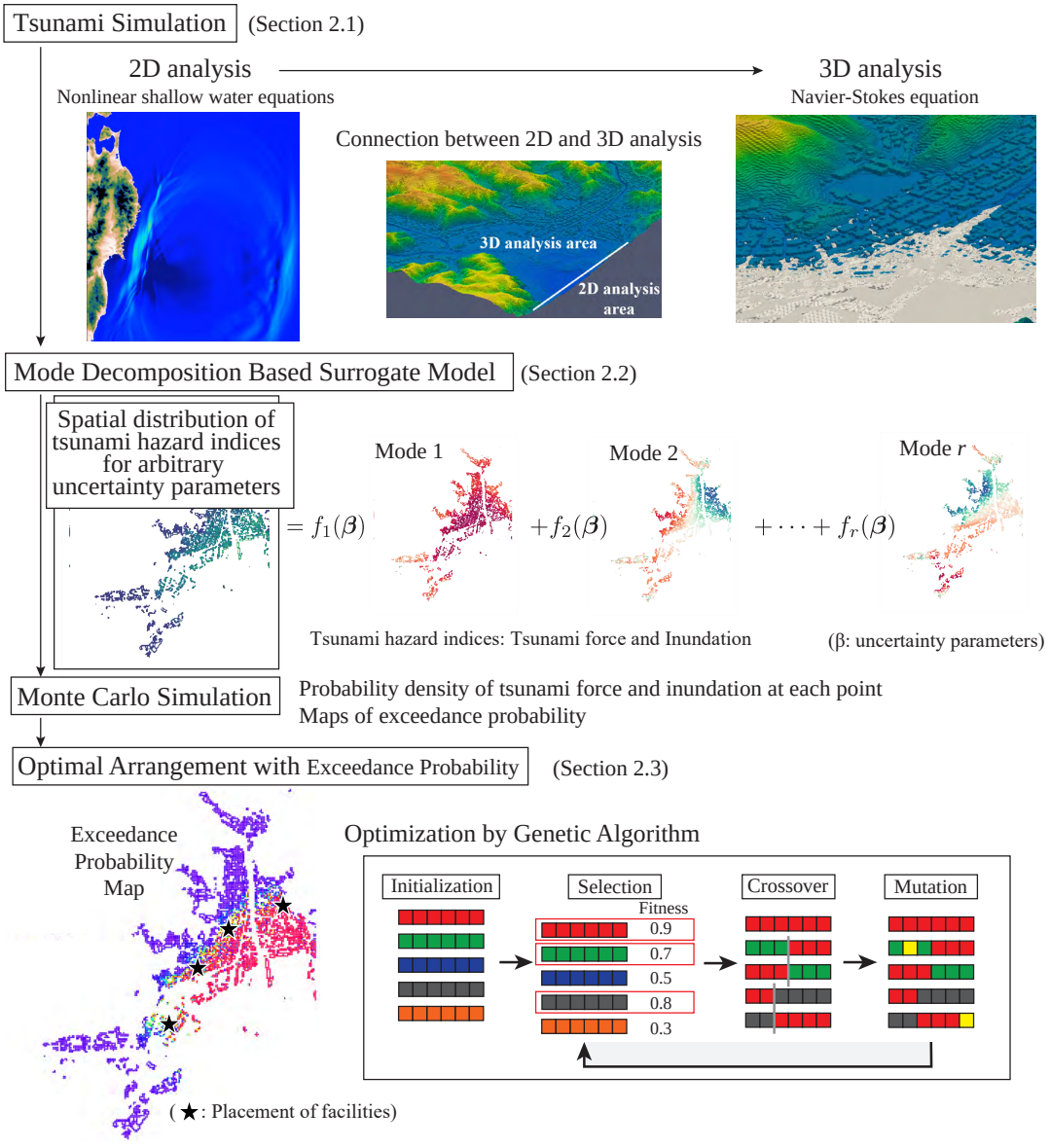


Figure 1. Flowchart of probabilistic tsunami hazard analysis using the mode-decomposition–based surrogate model.

and the time history of the tsunami wave height and flow velocity that were observed offshore of the target area were acquired. In this study, analyses were conducted using TUNAMI-N2 (Imamura, 1995; Goto et al., 1997). The following continuity and nonlinear long wave equations are solved in the 2D analysis:

$$\frac{\partial \eta}{\partial t} + \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} = 0 \tag{1}$$

$$\frac{\partial M}{\partial t} + \frac{\partial}{\partial x}\left[\frac{M^2}{D}\right] + \frac{\partial}{\partial y}\left[\frac{MN}{D}\right] + gD\frac{\partial \eta}{\partial x} + \frac{gn^2}{D^{\frac{7}{3}}} M\sqrt{M^2+N^2} = 0 \tag{2}$$

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial x}\left[\frac{MN}{D}\right] + \frac{\partial}{\partial y}\left[\frac{N^2}{D}\right] + gD\frac{\partial \eta}{\partial y} + \frac{gn^2}{D^{\frac{7}{3}}} N\sqrt{M^2+N^2} = 0 \tag{3}$$

where M and N are the flow rates in the x and y directions, respectively, η is the water level, D is the total water depth, g is gravitational acceleration and n is the Manning roughness coefficient.

The obtained tsunami wave height and flow velocity were set as the boundary conditions, and the tsunami reaching the target area was analyzed. The method proposed by Takase et al. (2016) is used for the boundary conditions of the 2D and 3D analyses. The time-series data of the wave height and flow velocity obtained from the 2D wide-area analysis are stored and transferred to the 3D numerical analysis by linear interpolation in space. The interpolated values are given to the 3D analysis as input data.

A 3D analysis was performed in this portion to assess the fluid force acting on the buildings in the target area. We employed the following set of 3D Navier-Stokes and continuity equations in the analysis domain $\Omega_{ns} \in R^3$

$$\rho\left(\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} - \boldsymbol{f}\right) - \nabla\cdot\boldsymbol{\sigma} = 0 \tag{4}$$

$$\nabla\cdot\boldsymbol{u} = 0 \tag{5}$$

where ρ is the mass density, $\boldsymbol{u}$ is the velocity vector, $\boldsymbol{\sigma}$ is the stress tensor, and $\boldsymbol{f}$ is the body force vector. Also, assuming a Newtonian fluid, the stress is calculated as

$$\boldsymbol{\sigma} = -p\boldsymbol{I} + 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) \tag{6}$$

where p is the pressure, μ is the coefficient of viscosity, and $\boldsymbol{\varepsilon}(\boldsymbol{u})$ is the velocity gradient tensor defined as

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T\right) \tag{7}$$

To solve the governing equations of the 3D simulation, the stabilized finite element method (SFEM) is used in this study. The detailed explanation of the numerical method used in this study is described in the relevant paper (Tozato et al., 2022). For the tsunami uncertainty, two fault parameters (described later) were adopted as the uncertainty parameters, and numerical analyses were conducted for several scenarios in which these parameters were changed. Specific analysis area setting conditions are shown in Section 3.

2.2 Construction of the surrogate model using mode decomposition

In this study, proper orthogonal decomposition (POD) was used to extract the spatial characteristic modes from the results of numerical analysis. To apply POD, a column vector $\boldsymbol{x}_i$ (called a data vector) was first defined to accommodate a tsunami hazard index, for which we selected the spatial distribution of either the maximum impact force acting on the buildings or the maximum inundation depth obtained from a numerical simulation for scenario i. Here, if there are n evaluation points, $\boldsymbol{x}_i$ has n components. Then, a data matrix was created by arranging all the data vectors according to a certain rule to be used for a target of POD.

$$\boldsymbol{X} = \begin{pmatrix} | & & | \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_N \\ | & & | \end{pmatrix} \tag{8}$$

where N refers to the number of scenarios. Furthermore, the vertical line in Eq. (8) was used to indicate that the data vector is a column vector. Using this matrix, the covariance matrix of the data was expressed in the form of $\boldsymbol{C} = \boldsymbol{X}\boldsymbol{X}^T$; the eigenvalues represent the variances, and the eigenvectors represent the spatial modes. In this study, we assumed that the eigenvalues were arranged in descending order from the first mode, and the eigenvalue and eigenvector corresponding to the j-th mode were expressed as λ_j and $\boldsymbol{u}_j$, respectively. Furthermore, in POD, the contribution rate of each mode is often used as a criterion for determining the number of dimensions to be reduced. The contribution rate is an index that shows how much each mode explains the data, and the contribution rate of the j-th mode d_j is expressed as follows, using the eigenvalues $\lambda_k (k = 1, ..., n)$.

$$d_j = \frac{\lambda_j}{\sum_{k=1}^{n} \lambda_k} \tag{9}$$

Furthermore, singular value decomposition was used to express the data matrix as follows using the eigenvalue λ_j and the eigenvector $\boldsymbol{u}_j$.

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \begin{pmatrix} | & & | \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_p \\ | & & | \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_p} \end{pmatrix} \begin{pmatrix} - & \boldsymbol{v}_1 & - \\ & \vdots & \\ - & \boldsymbol{v}_p & - \end{pmatrix} = \boldsymbol{U}\boldsymbol{A}^T \tag{10}$$

Here, $\boldsymbol{U}$ is a matrix in which the modes are arranged in the column direction, $\boldsymbol{\Sigma}$ is a matrix in which the square roots of the eigenvalues are arranged in diagonal terms, $\boldsymbol{V}$ is a matrix in which the eigenvectors of $\boldsymbol{X}^T\boldsymbol{X}$ are arranged, p is the number of eigenvalues that is greater than zero, and $\boldsymbol{A}^T = \boldsymbol{\Sigma}\boldsymbol{V}^T$ is a matrix in which the POD coefficients are arranged. The relationship of singular value decomposition for the result of one scenario is given as follows:

$$\boldsymbol{x}_i = \sum_{k=1}^{p} \alpha_{ik}\boldsymbol{u}_k = \alpha_{i1}\boldsymbol{u}_1 + \cdots + \alpha_{ip}\boldsymbol{u}_p \tag{11}$$

Here, α_{ik} donates ik component of the matrix $\boldsymbol{A}$ in which the POD coefficients are arranged. When removing modes with low explainability for the data, the data vector is expressed approximately as a linear combination excluding the modes with a low contribution rate as follows by determining the number of modes r to be reduced from the contribution rate or other indices.

$$\boldsymbol{x}_i \simeq \sum_{k=1}^{r} \alpha_{ik} \boldsymbol{u}_k = \alpha_{i1}\boldsymbol{u}_1 + \cdots + \alpha_{ir}\boldsymbol{u}_r \tag{12}$$

However, notably, a reduction in dimensions will result in a loss of the information contained in the omitted modes. Here, if this data $\boldsymbol{x}_i$ is the result of the uncertainty parameter $\boldsymbol{\beta}_i$, then the POD coefficients of any uncertainty parameter $\boldsymbol{\beta}$ can be expressed. Therefore, next, the POD coefficients α_{ik} are expressed as functions $f_k(\boldsymbol{\beta})(k = 1, ..., r)$ of the uncertainty parameter. The surrogate model can be expressed as follows by expressing this as a function of the uncertainty parameter for each corresponding mode.

$$\hat{\boldsymbol{x}}(\boldsymbol{\beta}) = \sum_{k=1}^{r} f_k(\boldsymbol{\beta}) \boldsymbol{u}_k \tag{13}$$

In this study, the radial basis functions (RBF) (Buhmann, 1990) were used as the interpolation functions. RBF interpolation can be used to handle cases where the analysis scenarios are not evenly arranged in the parameter space. The function $f_k(\boldsymbol{\beta})$ corresponding to mode k can be expressed as follows:

$$f_k(\boldsymbol{\beta}) = \sum_{i=1}^{N} w_i \phi(\boldsymbol{\beta}, \boldsymbol{\beta}_i) = \sum_{i=1}^{N} w_i \exp(-\gamma ||\boldsymbol{\beta} - \boldsymbol{\beta}_i||^2) \quad (k = 1, ..., r) \tag{14}$$

Here, $\boldsymbol{\beta}_i$ is the input parameter group for scenario i, w_i is the weight, and $\phi(\boldsymbol{\beta}, \boldsymbol{\beta}_i)$ is the basis function; γ is a parameter that determines the smoothness of the function. The weights of RBF interpolation can be obtained in the following form by substituting the correspondence between the known input parameter $\boldsymbol{\beta}_i$ and the POD coefficient α_{ik} that expresses the output result.

$$\begin{pmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{Nk} \end{pmatrix} = \begin{pmatrix} \phi(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1) & \cdots & \phi(\boldsymbol{\beta}_1, \boldsymbol{\beta}_N) \\ \vdots & & \vdots \\ \phi(\boldsymbol{\beta}_N, \boldsymbol{\beta}_1) & \cdots & \phi(\boldsymbol{\beta}_N, \boldsymbol{\beta}_N) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} \quad (k = 1, ..., r) \tag{15}$$

They can be expressed in their respective bold forms as follows:

$$\boldsymbol{\alpha}_k = \boldsymbol{\Phi} \boldsymbol{w}_k \quad (k = 1, ..., r) \tag{16}$$

Here, $\boldsymbol{\alpha}_k$ is a vector in which the coefficients of the k-th mode are arranged, and $\boldsymbol{w}_k$ is a vector in which the weights of $f_k(\boldsymbol{\beta})$ are arranged. The function using the weight obtained in Eq. (15) is expressed as an interpolation passing through all the referenced data points. However, cases in which the referenced data points change or oscillate at a local level can result in interpolation with little physical meaning. To resolve such issues, we introduced a regularization term for computing the weights. Specifically, we introduced L2 regularization called ridge regression (Hoerl and Kennard, 1970), and weight $\boldsymbol{w}_k$ was obtained by solving the following optimization problem.

$$\arg\min_{\boldsymbol{w}_k} (||\boldsymbol{\alpha}_k - \boldsymbol{\Phi} \boldsymbol{w}_k||_2^2 + \lambda ||\boldsymbol{w}_k||_2^2) \tag{17}$$

This process is generally used in the field of machine learning to prevent overfitting; λ indicates the degree of regularization. Introducing the regularization term allows to suppress the local oscillations and enables smooth interpolation. However, care must be taken for cases in which regularization is introduced because this may not pass through all data points. Furthermore, the accuracy of interpolation depends on the RBF parameter γ and regularization parameter λ; hence, these need to be appropriately determined. In this study, a combination of these parameters was determined by cross-validation (Stone, 1947) and Bayesian optimization (Močkus, 1975).

2.3 Search for the optimal placement using a genetic algorithm

A genetic algorithm (Holland, 1992) was used in this study for investigating optimal placement. Genetic algorithms search for approximate solutions of data, where multiple individuals whose solution candidates are expressed with genes are prepared, individuals with high fitness are preferentially selected, and solutions are searched for by repeating operations such as crossover and mutation. The problem targeted in this study includes an extremely large number of assessment points; furthermore, checking all combinations is extremely inefficient. Hence, we adopted the genetic algorithm for the combination optimization problem.

Figure 2 shows an overview of the genetic algorithm. In this study, the assessment point number was placed in the component of each individual, and optimization of those combinations was performed with a genetic algorithm. First, the number of individuals was determined, and a combination of points to be selected was randomly determined for the initial individuals. Next, the fitness was calculated for the generated individuals. Two individuals that are to be the parents of the next generation were then selected according to the obtained fitness. The parent selection method involves selecting individuals with high fitness as parents; low-fitness individuals are thus eliminated. The next generation of individuals is generated by randomly exchanging each component for the two selected parents. The location of exchange and the number of exchanges are randomly determined. This is repeated until the number of individuals in the next generation reaches the initially set number. In this study, we adopted an elite conservation strategy as a method to avoid deterioration of fitness during generational change, with settings such that some of the top individuals with high fitness could be passed on to the next generation as is. The final next-generation individuals were determined by mutating each component of each individual with a certain probability. In this process, the point number may be duplicated within one individual, and in such cases, the duplicated point is randomly re-selected. This process was repeated until the fitness converged, and an optimal point combination was determined.

3 Application to cases assuming an actual tsunami

The method described in the previous section was applied to a problem in which an actual tsunami was assumed. In this study, we conducted a series of numerical analyses that considered the uncertainties with the 2011 Off the Coast of Tohoku Earthquake as the target event. We applied mode decomposition on these results to construct a surrogate model of the numerical analysis, and we implemented Monte Carlo simulations to investigate the optimal placement of facilities probabilistically.

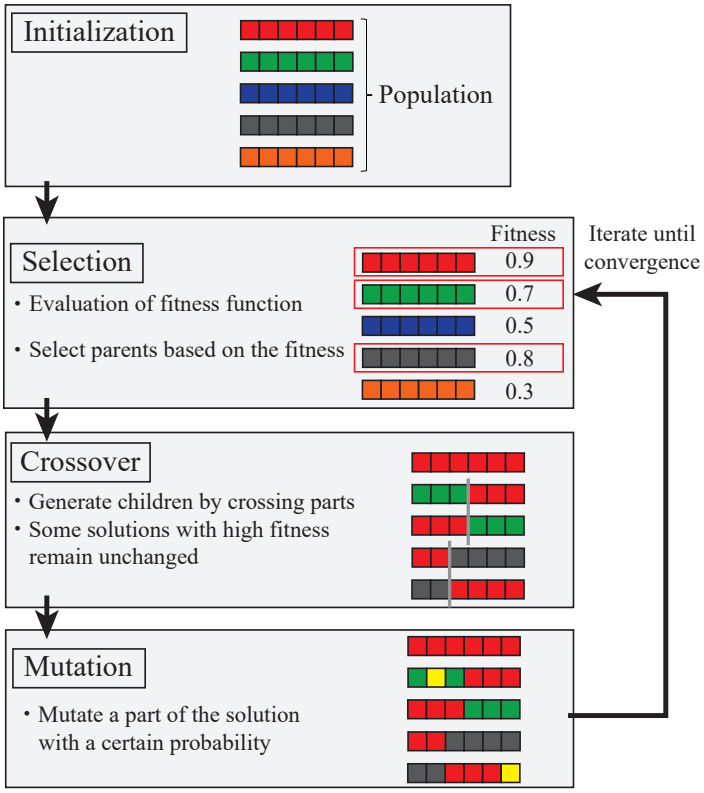


Figure 2. Flow of the genetic algorithm.

3.1 Tsunami simulation

We used the same numerical analysis results as those conducted in previous research (Tozato et al., 2022). The reader is referred to the study by (Tozato et al., 2022) for details regarding the computational conditions. This study only considers the slip and rake as uncertainty parameters shown in Fig. 3 because these parameters are supposed to have a deep relationship with the characteristics of fault stagger. The parameters at the time of the Off the Coast of Tohoku Earthquake were set as mean values, with the slip varying between 0.7 and 1.4 times, and the rake varying between $-20°$ to $+25°$. The illustration of the fault parameters is shown in Fig. 3. Table 1 shows all analysis cases. In this study, the validity of the constructed surrogate model was verified by using $\pm 10°$ rake cases among the 50 cases shown in Table 1; the remaining 40 cases were used to construct the surrogate model.

Figures 4 and 5 show the target area and snapshots of the analysis results of the inundation area for the mean case (S3R5), respectively. To confirm the validity of the numerical simulations, the simulation result for case S3R5, which corresponds to the actual tsunami condition, is compared with the observed data (The 2011 Tohoku Earthquake Tsunami Joint Survey Group, 2012; Mori et al., 2012). Figure 6 shows the simulated and observed inundation depths at the points A to H indicated in Fig.

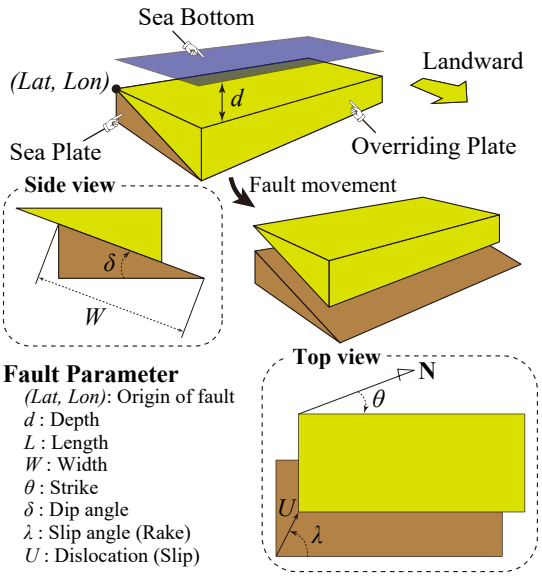


Figure 3. Illustration of the fault parameters. (borrowed from Kotani et al. (2020)).

Table 1. Names of the calculation cases. (Tozato et al. (2022))

			Rake									
		[°]	-20	-15	-10	-5	0	+5	+10	+15	+20	+25
	[%]	Normalized value	0.753	0.815	0.877	0.938	1	1.062	1.123	1.185	1.247	1.309
Slip	70	0.7	S1R1	S1R2	S1R3	S1R4	S1R5	S1R6	S1R7	S1R8	S1R9	S1R10
	85	0.85	S2R1	S2R2	S2R3	S2R4	S2R5	S2R6	S2R7	S2R8	S2R9	S2R10
	100	1	S3R1	S3R2	S3R3	S3R4	S3R5	S3R6	S3R7	S3R8	S3R9	S3R10
	120	1.2	S4R1	S4R2	S4R3	S4R4	S4R5	S4R6	S4R7	S4R8	S4R9	S4R10
	140	1.4	S5R1	S5R2	S5R3	S5R4	S5R5	S5R6	S5R7	S5R8	S5R9	S5R10

4. These results show that although there were some differences between the numerical analysis results and observed results at locations far away from the shore, the observation values were generally reproduced in the areas around the shore. The difference between the numerical analysis results and the observation values in the locations far away from the shore was considered to be because the outflows of the buildings were not considered in this study; hence, the waves did not reach the locations far away from the shore.

Figure 4. Boundary between 2D and 3D analyses areas. Points A to H are used to compare the inundation depths between the observational data and simulation results. (©Google Maps, Tozato et al. (2022))

In this study, the impact force acting on the buildings and the inundation depth were adopted as tsunami hazard indices. The effects of the tsunami fluid force have been considered even in recent design criteria (American Society of Civil Engineers, 2017; Nakano, 2017); therefore, we considered the fluid force acting on buildings as a tsunami hazard index. A 3D simulation was conducted to construct a surrogate model of the tsunami fluid force; however, the tsunami fluid force is strongly influenced by the direction that the building is facing, and it is difficult to assess the fluid force for each point. Therefore, the tsunami fluid force was assessed with a 2D mesh consisting of approximately 10 m $\times$ 10 m grids in this study, each of which is a unit for force evaluation. The tsunami impact force is calculated by synthesizing all the pressures acting on the surfaces of buildings within each grid in the two horizontal directions and averaged over the grid. Then, hereafter, each grid is regarded as a point associated with this averaged force. In addition, for each point, the maximum impact force is represented by evaluating the maximum value over the analysis time. An image of a mesh for evaluating the tsunami force is shown in Fig. 7. The target area was $2,145 \times 2,600$ m; hence, the number of assessment points in the POD was $n = 214 \times 260 = 55,640$.

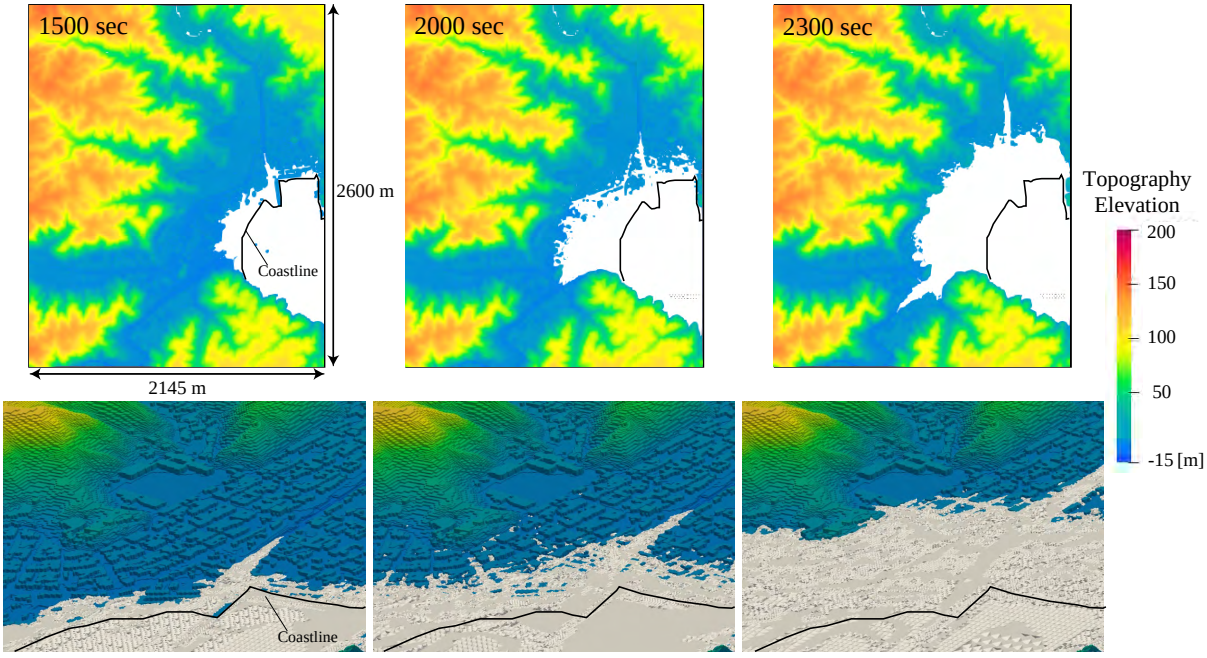


Figure 5. Snapshots of the tsunami runup obtained through 3D analysis. (Tozato et al. (2022))

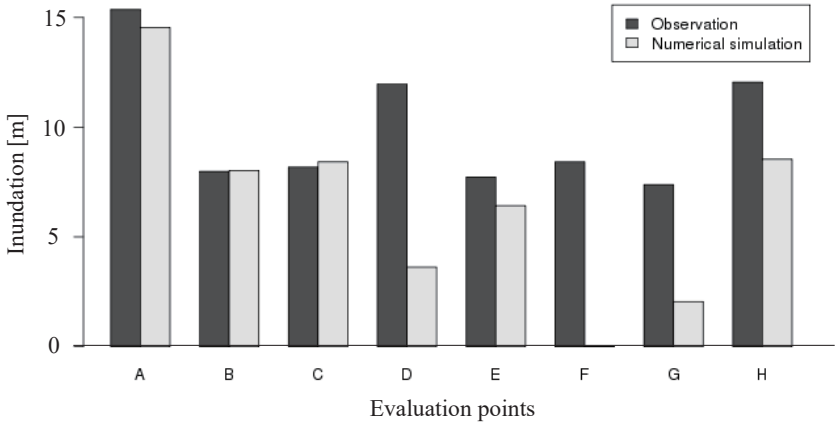


Figure 6. Comparison of the inundation between observational data and numerical simulation. (Observation data are provided by field survey results (The 2011 Tohoku Earthquake Tsunami Joint Survey Group, 2012)). (Tozato et al. (2022))

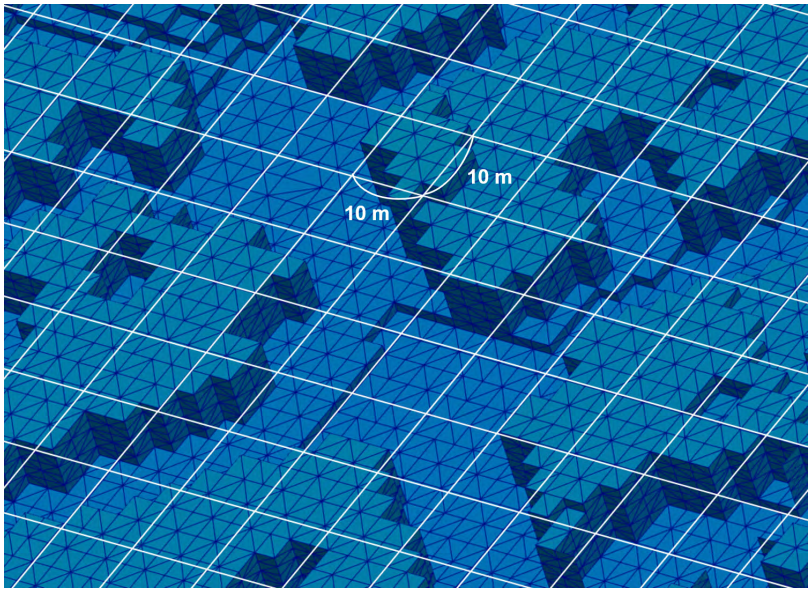


Figure 7. Image of a mesh for evaluating tsunami force. (Tozato et al. (2022))

3.2 Construction of a surrogate model with POD

POD was applied to the numerical analysis results to construct a surrogate model. When applying the POD, the data at each point were normalized in advance to mean values of 0 and standard deviations of 1. Figure 8 shows the spatial modes extracted by the POD from the first mode to the third mode. The values shown in the figure are values of the eigenvectors $\boldsymbol{u}_j$, and these were adjusted so that the maximum absolute value of the components in the eigenvectors was 1. The characteristics of the spatial distribution could be read for each tsunami hazard index from the extracted modes. A comparison of the spatial modes of the maximum impact force and maximum inundation depth confirmed that the three modes shown in Fig. 8 have similar spacial characteristics. For example, in the first mode, the sign was the same overall, and the value on the coast side was large; thus, the mode showed an overall tendency where the coast side was the most affected by the tsunami, with the effect becoming smaller moving away from the coast. In the second mode, a tendency could be seen where the maximum impact force and maximum inundation depth were opposite at the east and west sides. In the third mode, a tendency could be seen where the maximum impact force and maximum inundation depth were opposite at the north and south sides. These modes were considered to be related to the inflow direction of the tsunami. Furthermore, higher-dimension modes included the characteristics of local sections. Figure 9 shows the contribution rates of the modes. The contribution rate of the first mode was high for both the maximum impact force and maximum inundation depth.

Next, the coefficients of each mode were expressed as a function of the uncertainty parameters. As previously mentioned, RBF interpolation shown in Eq. (14) was used, and the regularization shown in Eq. (17) was introduced in the calculation to obtain the weight. The accuracy of the surrogate model changed according to the RBF smoothness parameter and regularization

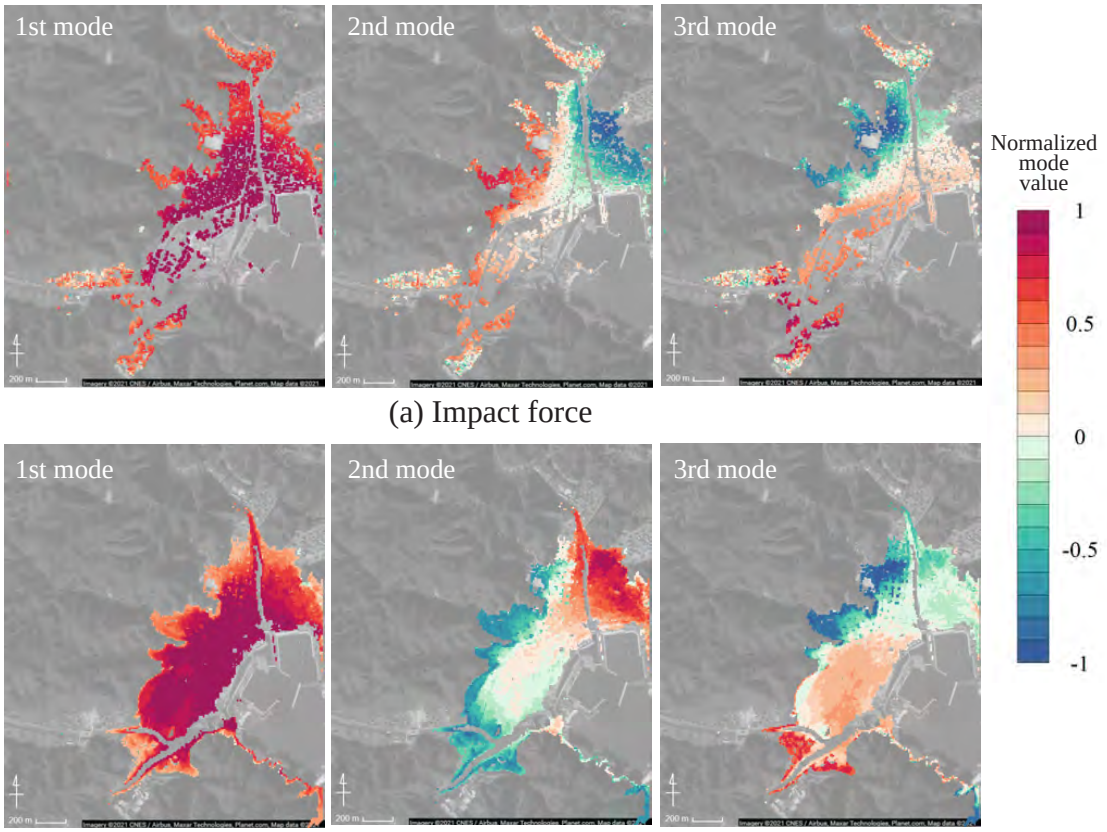


Figure 8. Spatial modes for each tsunami hazard index extracted by POD. (left: 1st mode, middle: 2nd mode, right: 3rd mode) (©Google Maps)

parameter; thus, it is important to appropriately determine these. In this study, these parameters were determined using cross-validation.

The learning and verification cases used for cross-validation were obtained by taking a total of 40 cases used in the construction of the surrogate model, removing four cases that correspond to the corners of the parameter space (S1R1, S1R10, S5R1, S5R10), and dividing the remaining 36 cases between learning and verification cases for cross-validation. The corner data were not used for the verification cases because these data were extrapolated. In this example, the number of divisions between the learning and verification cases was set to 12. In other words, the model was constructed using 37 cases for a single validation, and the validation is conducted using three cases. Furthermore, the cross-validation error was calculated by comparing the reconstructed results using the spatial mode and taking the ratio of the mean absolute error to the mean value as shown in the

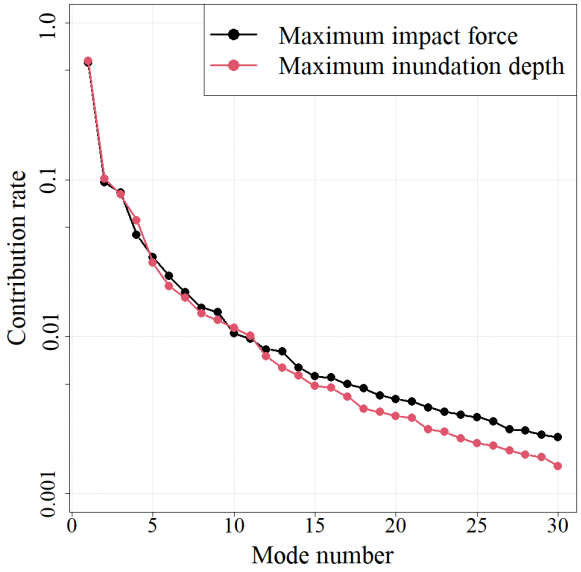


Figure 9. Contribution rates for each tsunami hazard index.

following equation.

$$e_r = \frac{\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}|x_{ij} - \hat{x}_{ij}|}{\frac{1}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}x_{ij}} \tag{18}$$

Here, n is the number of assessment points, N is the number of scenarios, and x_{ij} is the numerical analysis result for scenario i and point j. Furthermore, $\hat{x}_{ij}$ is the value for scenario i and point j when reconstructed with the surrogate model, and e_r is the error when the number of modes is r.

Figure 10 shows the cross-validation error of the maximum impact force and maximum inundation depth for each number of modes. Cases in which no regularization term is present (Eq. (15)) are also shown in Fig. 10. Bayesian optimization was used for the search in the parameter space, the number of searches was set to 80, and the upper confidence bound (UCB) strategy was used for the acquisition function. The error comparison confirmed that the accuracy of the surrogate model was improved by introducing the regularization term. By introducing the regularization term, a robust surrogate model can be constructed.

Finally, the validity of the surrogate model was verified by comparing the numerical simulation results for the scenarios that were not used for the constructed surrogate model with the results of the constructed surrogate model. Figure 11 shows a comparison between the numerical simulation results and the surrogate model results for the S3R3 scenario. Regarding the number of modes used in the surrogate model, the maximum impact force was set to 8 and the maximum inundation depth was set to 11. Also, Fig. 12 shows the mean absolute errors calculated by Eq. (18) in calculations for 10 scenarios used for validation. As can be seen from Fig. 11, the constructed surrogate models can roughly represent the targeted spatial distribution of the maximum impact force and maximum inundation depth. However, Fig. 12 shows that 10 % or higher errors occur in

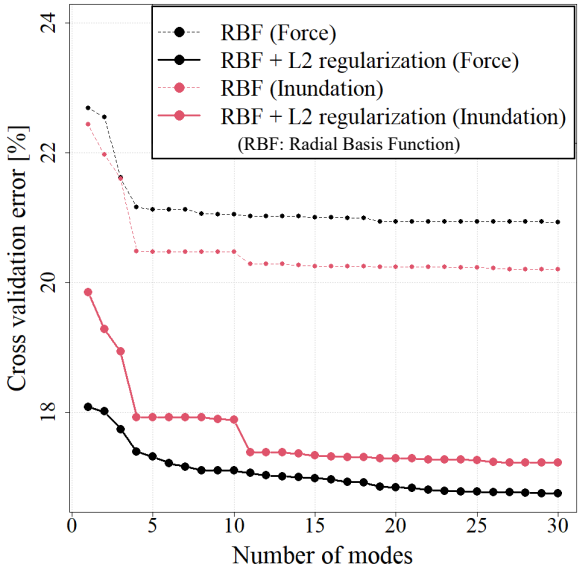


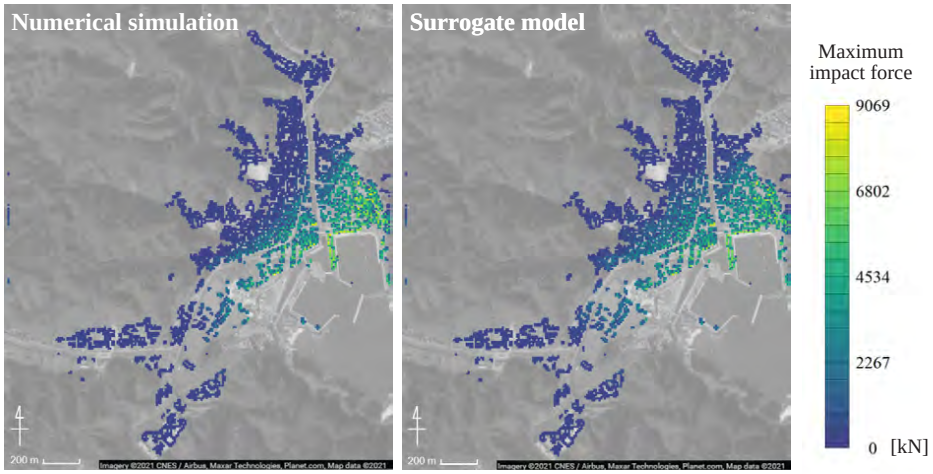
Figure 10. Cross-validation error for each number of modes.

some validation scenarios, and the areas of large error are localized. This is because there is a possibility that spatial modes used for the surrogate models cannot properly capture the local characteristics.

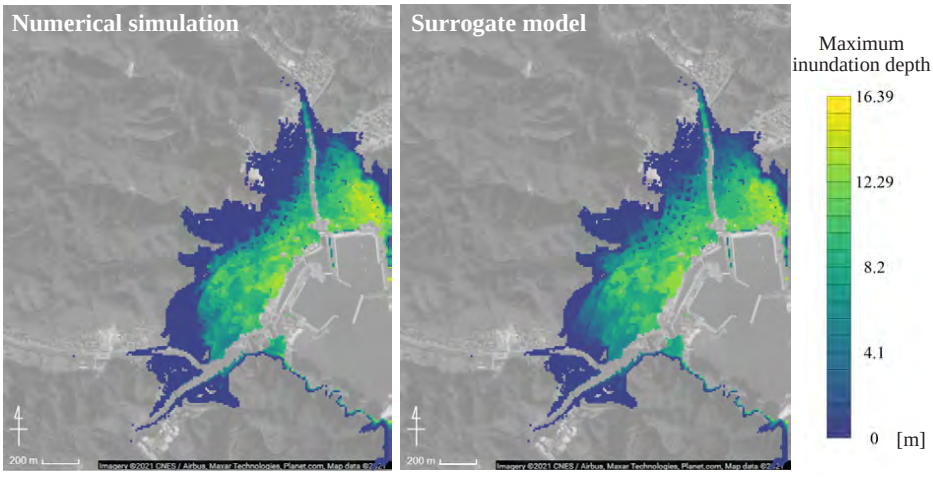
3.3 Monte Carlo simulation

We conducted probabilistic tsunami hazard assessments using the constructed surrogate model. The surrogate model enables the computation of the spatial distribution of hazard indices at a low computational cost; hence, many trials can be secured at a relatively low computational cost, and a risk assessment that efficiently utilizes the advanced numerical simulation results can be conducted.

In this study, we conducted a probabilistic assessment of the tsunami hazard by applying Monte Carlo simulations. The variation of uncertainty parameters must be quantitatively assessed probabilistically for Monte Carlo simulations. We assumed that the slip and rake followed a normal distribution, and the probability distribution parameters were set as shown in Table 2. For the mean values, as mentioned in the previous section, normalized values were used as the input parameters, so these were set as 1.0. Furthermore, for the standard deviation of the slip, a value of 0.1 was used, which indicates a standard deviation value that is 10% of the mean value. For a normal distribution, the spread of approximately three times the given standard deviation is present; thus, the value of three times the standard deviation was set as this value so that the range of Table 1 in the previous section was covered. Furthermore, for the standard deviation of the rake, Japan Society of Civil Engineering (2011) conducted probabilistic assessments where the rake was varied by $\pm 10°$; therefore, in this study as well, we considered this



(a) Maximum impact force



(b) Maximum inundation depth

Figure 11. Comparison between results of numerical simulation and the surrogate model. (Scenario: S3R3) (©Google Maps)

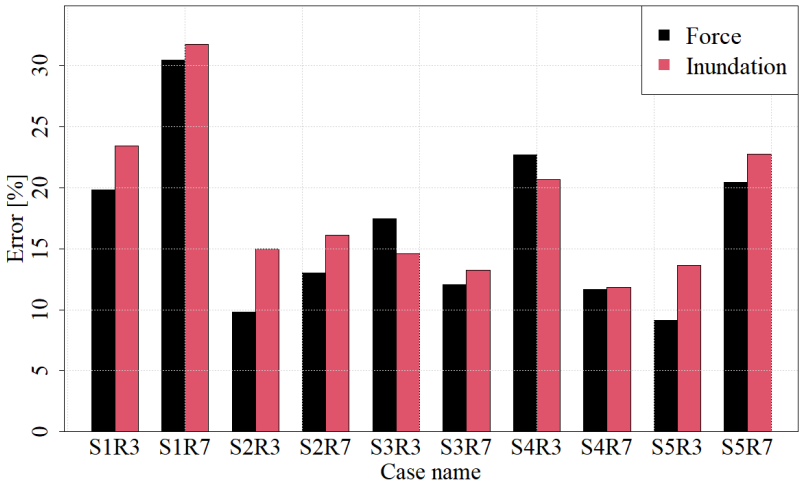


Figure 12. Errors between the numerical simulation and surrogate model for each validation scenario.

degree of variation, and a value of 0.04 was used, which was considered to result in a variation of approximately three times the standard deviation.

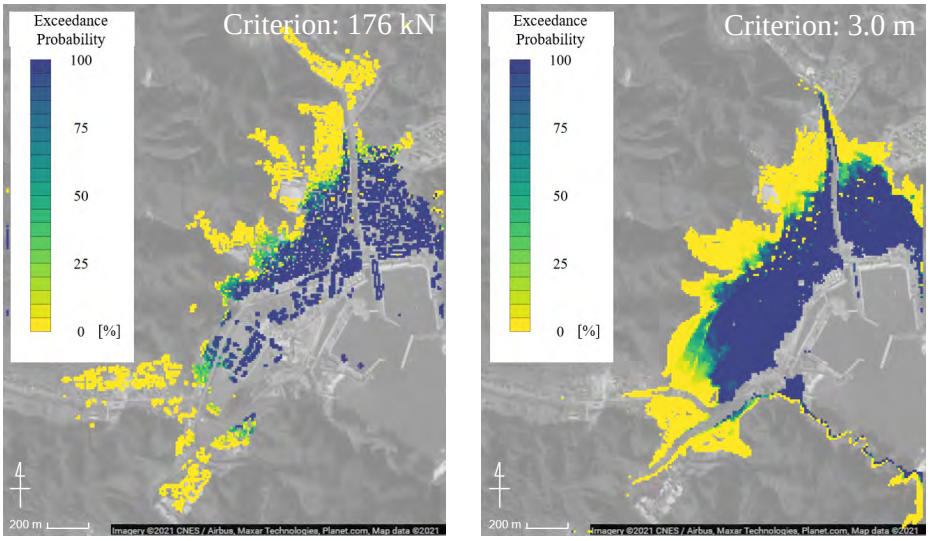
Table 2. Information on the variation of uncertainty parameters.

Parameter	Mean	Standard deviation
Slip	1.0	0.1
Rake	1.0	0.04

Monte Carlo simulations were conducted using the uncertainty parameters listed in Table 2 and the surrogate model. Specifically, the information of each uncertainty parameter was used to randomly generate a combination of the slip and rake; this value was assigned to the mode coefficient functions of the surrogate model, and the spatial distribution of the hazard index was calculated by multiplying the coefficients and the modes. This was repeated for the number of trials, which was set to 10,000 in this study, and the probability density distributions of the maximum impact force and maximum inundation depth at each point were calculated.

Maps of exceedance probability are obtained from the results of Monte Carlo simulations. The exceedance probability at each evaluation point is calculated assuming the failure that can be defined by the criteria of each hazard index. Based on the previous studies (Suppasri et al., 2013, 2019), we defined the criteria as 176 kN for the maximum impact force and 3 m for the maximum inundation depth. The obtained exceedance probability maps for both hazard indices are shown in Fig. 13. In both maps, there is a tendency that high exceedance probabilities arise near the coast and low exceedance probabilities occur farther away from the coast. On the other hand, there are some differences between the maps locally. For example, the exceedance

probability of the maximum impact force tends to be high in the areas where there are many buildings. Since the computational cost of surrogate models is quite low, such probabilistic maps can be easily obtained.



(a) Impact force (b) Inundation

Figure 13. Spatial distribution for the exceedance probability. (©Google Maps)

3.4 Optimal placement of facilities using genetic algorithms

The results of the Monte Carlo simulation are used to investigate an optimal placement of the facilities. The optimal placement is examined based on the concepts of both parallel and series systems, and the results are compared. In the case of a parallel system, the system failure is considered to occur when all of the facilities have failed. On the other hand, regarding a series system, the system is considered to be failed when any one of the facilities is failed. The failure of the facility is defined as exceeding the failure criterion for each hazard index.

In this study, the optimization problem is defined to select the components of the system from the evaluation points for the hazard indices. The areas that have 25% or higher exceedance probabilities for each failure criterion are set as the target areas. The reason why the points of the system were selected only in these high risk areas is that a point with an exceedance probability of 0% should be always selected if there are such points in the target area, and this does not result in an optimal problem. This condition is not so realistic, but we employed it as a calculation condition to clearly represent the performance of the proposed method.

For the settings of the genetic algorithm, the number of individuals was set to 200, and the mutation probability was set to 10% for each component of each individual. Furthermore, we adopted an elite conservation strategy, and the solutions that have high fitness values are likely to be selected for the crossover. In this study, the solution convergence was set as that when

the solution individuals do not change over 2,000 steps. We investigated cases in which the number of the facilities is 4 for each failure criterion and each system. Since the solution may depend on on the initial conditions, we hence conducted three trials under the same conditions for each case.

The results of optimal placements based on the concepts of the parallel and series systems for each hazard index are shown in Fig. 14. In these figures, the results are also compared with the placements determined based on a simple strategy, in which the placement of the facilities is selected in order of lower exceedance probability. Red points shown in Fig. 14 indicate the placement selected by the genetic algorithm, and black points show the placement selected in order of lower exceedance probability. System failure probabilities for each placement are shown in Table 3. The system failure probabilities are expressed as the probability of failure of all facilities for the parallel system and the probability of failure of one or more facilities for the series system.

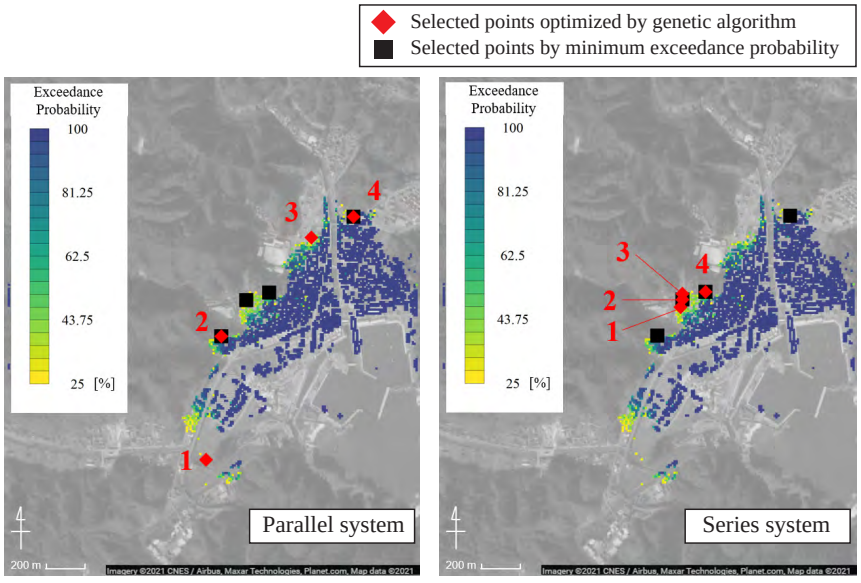
Table 3. System failure probability for each hazard index and each system.

	Impact force		Inundation	
	Parallel	Series	Parallel	Series
Minimum failure probability [%]	20.94	29.62	22.50	28.03
Genetic algorithm [%]	19.40	28.09	0.10	25.50

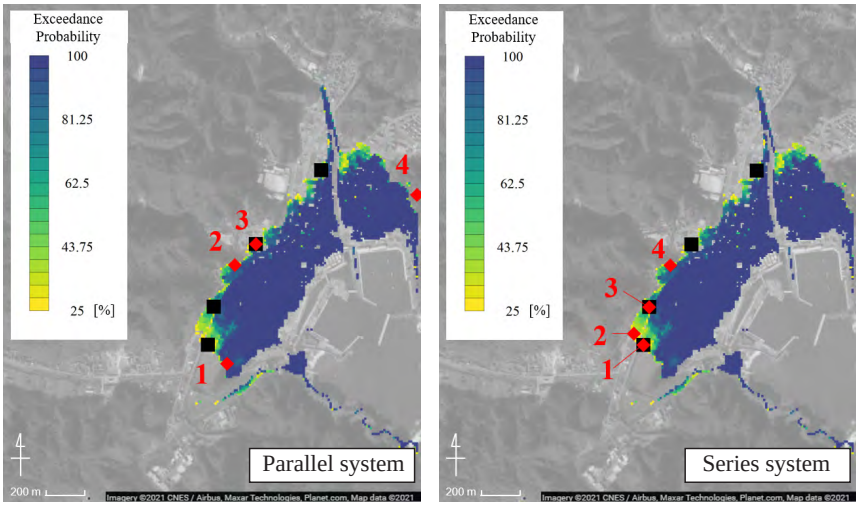
According to the results shown in Fig. 14, it is found that the selected points are placed away from the coastline. This tendency comes from the fact that low failure probability points are generally located away from the coastal region. Different placements were obtained for the maximum impact force and inundation depth. Comparing the parallel and series systems, it can be seen that the points are selected by concentrating on similar locations in the series system, while the points are spatially distributed in the parallel system. Regarding the series system, where all failures of one or more facilities would result in the failure of the system, the points are concentrated in such an arrangement that the probability of exceedance is small. On the contrary, one or more of the facilities need only be safe in a parallel system. That is why the selected points are spatially distributed.

The comparison between the placement selected by the genetic algorithms and the placement selected in order of lower exceedance probabilities shows that some selected locations are different from each other. The system failure probabilities shown in Table 3 indicate that lower failure probability tends to be obtained by the genetic algorithm. Thus, it can be confirmed that the genetic algorithm is suitable in this particular optimal placement problem.

A detailed analysis of the placements selected for each system was performed. Figure 15 shows the scatter plots of the uncertainty parameters, slip and rake, considered in the Monte Carlo simulation, and the number of failure points are colored. The placement for the parallel system optimizes around a smaller light blue area and the placement for the series system optimizes around a larger black area. According to Fig. 15, the slip mainly contributes to the failure of each point and the system for all placements because no failure occurs in smaller-slip cases and the failure occurs in the high-slip case. The rake

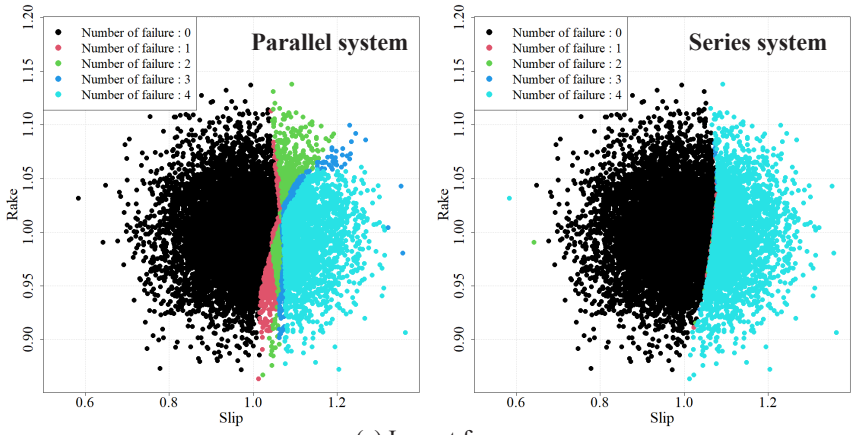


(a) Impact force

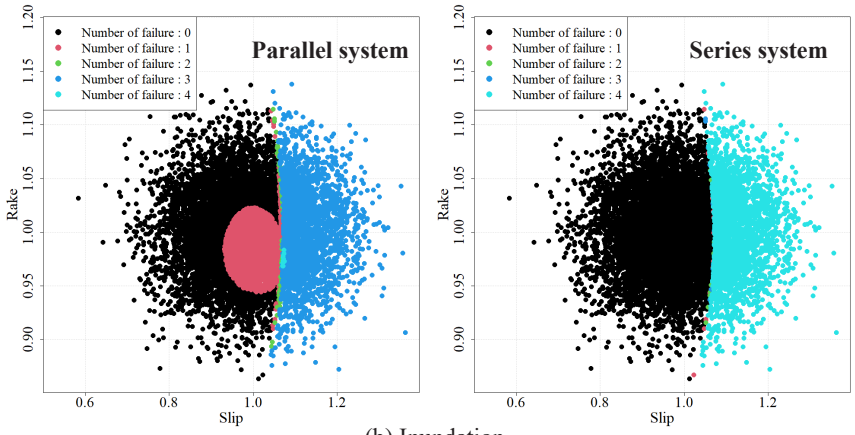


(b) Inundation

Figure 14. Optimal placements obtained from the genetic algorithm for each hazard index. (©Google Maps) Selected points by minimum exceedance probability show the results of placements determined in order of decreasing exceedance probability at each evaluation point. Only search areas are colored in this figure.



(a) Impact force



(b) Inundation

Figure 15. The scatter plot of the uncertainty parameters colored by the number of failure points for each placement.

contributes to the parallel system placement because the number of failure points changes in the same slip values for the parallel system.

In the series system, the scenarios are separated into two situations; all safe and all failed. In contrast, the parallel system shows a more complex placement pattern. Figure 16 shows a colored scenario in the uncertainty parameter space. As can be seen in Fig. 16, regarding the maximum impact force, it can be seen that the optimal points are selected such a way that they are not failed depending on the rake. On the other hand, in the case of the inundation depth, it can be confirmed that point 4 has an unique tendency compared to the other points. As we can understand from this numerical example, the optimal placement is efficiently discussed by the method proposed in this study.

Using the results of the Monte Carlo simulation using the surrogate models, the optimal facility placements can be probabilistically investigated based on information from the advanced numerical simulation. This method could be used to solve the problem of optimal placement of facilities such as relief bases, shelters, and infrastructure facilities during disasters.

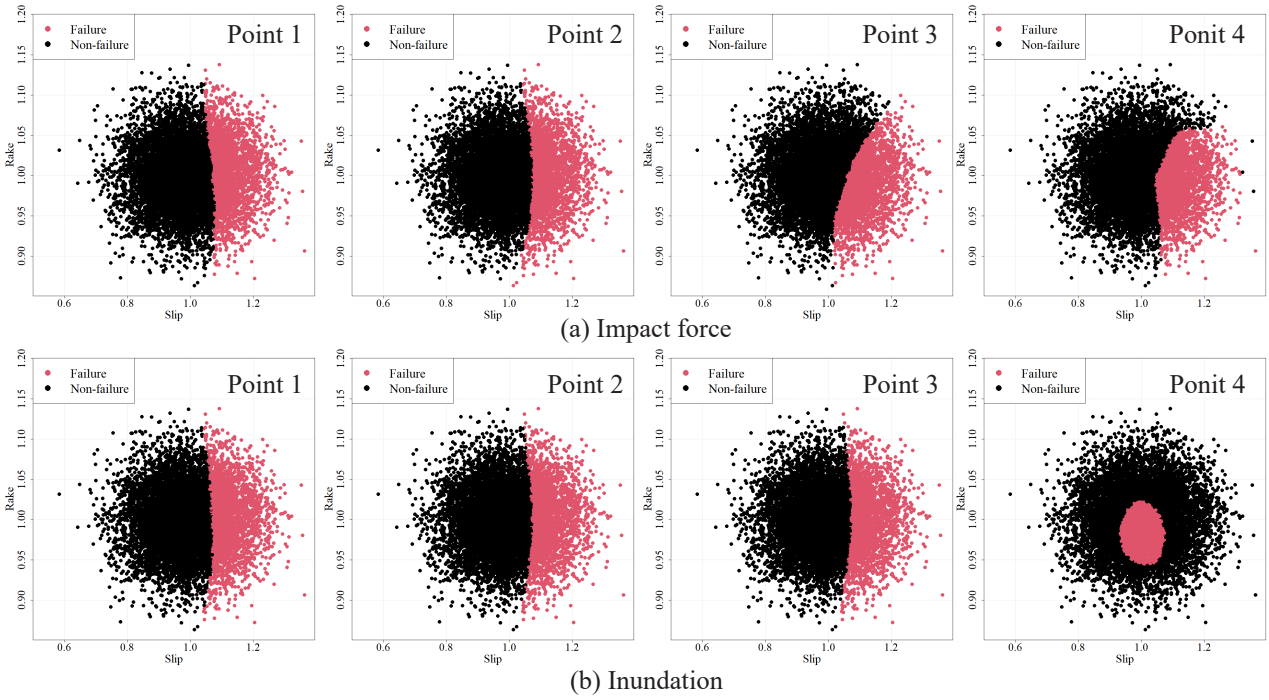


Figure 16. The scatter plot of the uncertainty parameters colored by failure or non-failure for each placement in inundation case.

4 Conclusions

In this study, we used /redthe results obtained by 2D-3D coupled simulation to construct a surrogate model, and we used this model to propose a method that enabled the efficient investigation of the optimal placement of facilities with the probabilistic approach. We constructed a surrogate model by applying POD to 3D tsunami simulation results, and Monte Carlo simulations that used this surrogate model were conducted to show that it was possible to assess the probability density distribution of the hazard indices at all points within the target area and the spatial distribution of exceedance probability with a low computational cost. Furthermore, we show that applying the genetic algorithm to Monte Carlo simulation results enabled the search for the optimal placement of parallel and series systems, which minimized failure probability. In this way, the proposed method enabled the investigation of the optimal placement of facilities probabilistically by efficiently utilizing the information obtained from advanced numerical simulations.

The uncertainty parameters in this study were limited to two when conducting assessments; however, there are more uncertainties than these in actual phenomena. Thus, it is important to conduct a probabilistic hazard assessment that considers this fact. Furthermore, the extent to which each uncertainty parameter will fluctuate must be assessed in advance to determine cases in which numerical simulations or Monte Carlo simulations are to be conducted. The surrogate model that was constructed with the proposed method is capable of assessments with sufficient accuracy for interpolation (within the range of uncertainty parameters for which numerical simulations were conducted); however, the accuracy often decreases for cases of extrapolation

(outside the range of uncertainty parameters for which the numerical simulations were conducted); thus, an assessment of uncertainty fluctuation is important. In addition, since the accuracy of the surrogate model changes according to the number of training data and the number of spatial modes used in the surrogate model, it's necessary to establish a way of properly determining them in future studies. In this study, we used the simulation data set that was created in the previous study and therefore could not investigate their effect for surrogate modeling. Nevertheless, an appropriate number of training data should be carefully determined in light of accuracy. To address this problem, it may be necessary to consider adopting an approach like Adaptive Surrogate Modeling (e.g., Wang et al., 2014; Gong et al., 2016) to check and improve the accuracy of surrogate models. It is also noted that we calculated the exceedance probability with the failure criterion as a constant when investigating the exceedance probability and optimal placement; however, the destruction criteria vary according to the building material. Thus, the use of this information would enable a more advanced probabilistic risk assessment and optimal placement of facilities.

Appendix A: Verification of the validity of the numerical analysis method

In this study, we conducted a comparison with the experimental results of the study by Winter et al. (2020) and verified the validity of the 3D analysis method adopted in this study. In the study by Winter et al. (2020), experiments were conducted on the fluid force acting on the structure while changing the structural placement; in this study, we conducted a comparison between the experiment results under the conditions shown in Fig. A1 among the aforementioned changes in structure placement with the analysis results. Details of the experiment are as shown in the study by Winter et al. (2020), and a comparison was performed with the numerical analysis results regarding the temporal changes in the water level on the front side of the structure and the fluid force acting on the structure. Fig. A2 shows the comparison results for each case. As shown in the figure, the validity of the adopted numerical analysis could be confirmed because the fluid force and water depth tendencies were generally captured. Furthermore, Fig. A3 shows a snapshot of the simulation results.

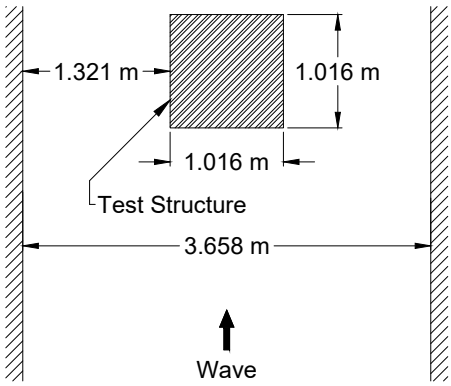


Figure A1. Configuration of the test structure. (borrowed from Winter et al. (2020))

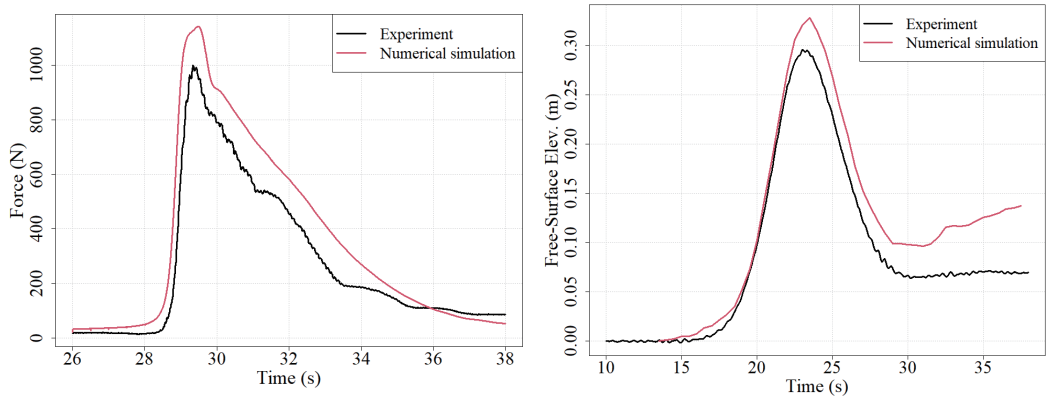


Figure A2. Comparison between the experimental results and numerical simulation results. ((a) Force, (b) Free-surface elevation)

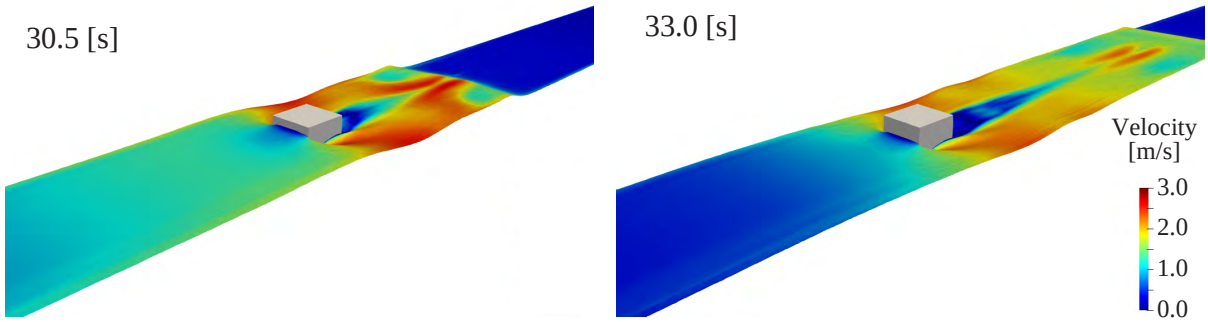


Figure A3. Snapshot of the numerical simulation result.

Code and data availability. Source code and details of the tsunami simulation were sourced from Kotani et al. (2020). Outputs of the simulations are available the Zenodo open-access repository at https://doi.org/10.5281/zenodo.6394294 (Tozato , 2022).

Author contributions. KTo contributed to methodology, analyses, and preparing the original manuscript. SM contributed to the conceptualization and methodology, and prepared the manuscript. ST contributed to the numerical simulations. YO and AS contributed to the methodology and conceptualization. MM contributed to the investigation and reviewing the manuscript. KTe contributed to the supervision and reviewing the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

References

Alhamid, A. K., Akiyama, M., Ishibashi, H., Aoki, K., Koshimura, S., Frangopol, D. M.: Framework for probabilistic tsunami hazard assessment considering the effects of sea-level rise due to climate change, Structural Safety, 94, 102152, https://doi.org/10.1016/j.strusafe.2021.102152, 2022.

Annaka, T., Satake, K., Sakakiyama, T., Yanagisawa, K., Shuto, N.: Logic-tree approach for probabilistic tsunami hazard analysis and its applications to the Japanese coasts. Pure Appl. Geophys. 164(2), 577–592, https://doi.org/10.1007/s00024-006-0174-3, 2007.

Baba, T., Kamiya, M., Tanaka, N., Sumida, Y., Yamanaka, R., Watanabe, K., Fujiwara, H.: Probabilistic tsunami hazard assessment based on the Gutenberg–Richter law in eastern Shikoku, Nankai subduction zone, Japan. Earth Planets Space, 74, 156, https://doi.org/10.1186/s40623-022-01715-1, 2022.

Bamer, F., Bucher, C.: Application of the proper orthogonal decomposition for linear and nonlinear structures under transient excitations. Acta Mechanica 223(12), 2549–2563, https://doi.org/10.1007/s00707-012-0726-9, 2012.

Buhmann, M.D.: Multivariate cardinal interpolation with radial-basis functions. Constr. Approxim. 6(3), 225–255, https://doi.org/10.1007/BF01890410, 1990.

Cavdur, F., Kose-Kucuk, M., Sebatli, A.: Allocation of temporary disaster-response facilities for relief-supplies distribution: A stochastic optimization approach for after disaster uncertainty, Natural Hazards Review, 22(1), 05020013, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000416, 2020.

Cavdur, F., Sebatli-Saglam, A., Kose-Kucuk, M.: A spreadsheet-based decision support tool for temporary-disaster-response facilities allocation. Saf. Sci. 124, 104581, https://doi.org/10.1016/j.ssci.2019.104581, 2020.

Cornell. C. A.: Engineering seismic risk analysis. Bull. Seismol. Soc. Am. 58(5), 1583–1606, https://doi.org/10.1785/BSSA0580051583, 1968.

Doerner, K.F., Gutjahr, W.J., Nolz, P.C.: Multi-criteria location planning for public facilities in tsunami-prone coastal areas. OR Spectrum 31, 651–678, https://doi.org/10.1007/s00291-008-0126-7, 2008.

Gong, W., Duan, Q., Li, J., Wang, C., Di, Z., Ye, A., Miao, C., and Dai, Y.: Multiobjective adaptive surrogate modeling-based optimization for parameter estimation of large, complex geophysical models, Water Resour. Res., 52, 1984– 2008, https://doi.org/10.1002/2015WR018230, 2016.

Fukutani, Y., Moriguchi, S., Terada, K., Otake, Y.: Time-dependent probabilistic tsunami inundation assessment using mode decomposition to assess uncertainty for an earthquake scenario. Journal of Geophysical Research: Oceans, 126, e2021JC017250, https://doi.org/10.1029/2021JC017250, 2021.

Fukutani, Y., Suppasri, A., Imamura, F.: Stochastic analysis and uncertainty assessment of tsunami wave height using a random source parameter model that targets a Tohoku-type earthquake fault. Stoch. Environ. Res. Risk Assess 29(7), 1763–1779, https://doi.org/10.1007/s00477-014-0966-4, 2015.

Geist, E.L., Parsons, T.: Probabilistic analysis of tsunami hazards*. Nat. Hazards 37(3), 277–314, https://doi.org/10.1007/s11069-005-4646-z, 2006.

Gomez, C., Baker, J.W.: An optimization-based decision support framework for coupled pre- and post-earthquake infrastructure risk management. Struct. Saf. 77, 1–9, https://doi.org/10.1016/j.strusafe.2018.10.002, 2019.

Gopinathan, D., Heidarzadeh, M., Guillas S.: Probabilistic quantification of tsunami current hazard using statistical emulation, Proc. R. Soc. 477, 20210180, http://doi.org/10.1098/rspa.2021.0180, 2021.

Goto, C., Ogawa, Y., Shuto, N., Imamura, F.: Numerical method of tsunami simulation with the leap-frog scheme. IUGG/IOC TIME Project, IOC Manual and Guides, 35, 1–126, 1997.

Ha, D.M., Tkalich, P., Chan, E.S.: Tsunami forecasting using proper orthogonal decomposition method. J. Geophys. Res. Oceans113(C6), https://doi.org/10.1029/2007JC004583, 2008.

Heidarzadeh, M., Kijko, A.: A probabilistic tsunami hazard assessment for the Makran subduction zone at the northwestern Indian Ocean. Nat. Hazards, 56 (3), 577-593. https://doi.org/10.1007/s11069-010-9574-x, 2011.

Heidarzadeh, M., Gusman, A., Ishibe, T., Sabeti, R., Šepić, J.: Estimating the eruption-induced water displacement source of the 15 January 2022 Tonga volcanic tsunami from tsunami spectra and numerical modelling. Ocean Engineering, 261, 112165. https://doi.org/10.1016/j.oceaneng.2022.112165, 2022.

Hoerl, A.E., Kennard, R.W.: Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12(1), 55–67, https://doi.org/10.1080/00401706.1970.10488634, 1970.

Holland, J.H.: Adaptation in Natural and Artificial Systems, second edition, University of Michigan Press, Ann Arbor, MI,1992, 1992.

Hotelling, H.: Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 25(6), 417–441, https://psycnet.apa.org/doi/10.1037/h0071325, 1933.

Imamura, F.: Review of tsunami simulation with a finite difference method, in Long-Wave Runup Models, edited by H. Yeh, P. Liu, and C. Synolakis, World Scientific Publishing, Hackensack, N. J, 25–42, https://doi.org/10.1142/9789814530330, 1995.

Ishikawa, Y., Kameda, H.: Hazard-consistent magnitude and distance for extended seismic risk analysis. Proceedings of the 9th World Conference on Earthquake Engineering. Tokyo-Kyoto, 89–94, 1988.

Japan Society of Civil Engineering. The method of probabilistic tsunami hazard analysis (in japanese), 2011.

Jolliffe, I.T., Cadima, J.: Principal component analysis: a review and recent developments. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374(2065), 20150202, https://doi.org/10.1098/rsta.2015.0202, 2016.

Karhunen, K.: Uber lineare methoden in der wahrscheinlichkeitsrechnung. Annals of Academic Science Fennicae, Series A1 Mathematics and Physics, 37, 3–79, 1947.

Kerschen, G., Golinval, J.C., Vakakis, A.F., Bergman, L.A.: The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: An overview. Nonlinear Dyn. 41(1), 147–169, https://doi.org/10.1007/s11071-005-2803-2, 2005.

Kosambi, D.D.: Statistics in function space. J. Indian Math. Soc. 7, 76–88, 1943.

Kotani, T., Tozato, K., Takase, S., Moriguchi, S., Terada, K., Fukutani, Y., Otake, Y., Nojima, K., Sakuraba, M., Choe, Y.: Probabilistic tsunami hazard assessment with simulation-based response surfaces. Coast. Eng. 160, 103719, https://doi.org/10.1016/j.coastaleng.2020.103719, 2020.

Kubota, T., Saito, T., Nishida, K.: Global fast-traveling tsunamis driven by atmospheric Lamb waves on the 2022 Tonga eruption, Science, 377(6601), 91-94, https://doi.org/10.1126/science.abo4364, 2022.

LeVeque, R.J., Waagan, K., Gonz´alez, F.I., Rim, D., Lin, G:. Generating random earthquake events for probabilistic tsunami hazard assessment. Pure Appl. Geophys. 173(12), 3671–3692, https://doi.org/10.1007/s00024-016-1357-1, 2016.

Maharjan, R., Hanaoka, S.: A credibility-based multi-objective temporary logistics hub location- allocation model for relief supply and distribution under uncertainty. Socio-Econ. Plan. Sci. 70, 100727, https://doi.org/10.1016/j.seps.2019.07.003, 2020.

McGuire, R.K.: Seismic design spectra and mapping procedures using hazard analysis based directly on oscillator response. Earthq. Eng. Struct. Dyn. 5(3), 211–234, https://doi.org/10.1002/eqe.4290050302, 1977.

Melgar, D., LeVeque, R.J., Dreger, D.S., Allen, R.M.: Kinematic rupture scenarios and synthetic displacement data: An example application to the Cascadia Subduction Zone. J. Geophys. Res. Solid Earth 121(9), 6658–6674, https://doi.org/10.1002/2016JB013314, 2016.

Miller, M., Baker, J.: Ground-motion intensity and damage map selection for probabilistic infrastructure network risk assessment using optimization. Earthq. Eng. Struct. Dyn. 44, 1139–1156, https://doi.org/10.1002/eqe.2506, 2015.

Minimum Design Loads and Associated Criteria for Buildings and Other Structures. American Society of Civil Engineers, ASCE/sei 7-16 edition, 2017.

Mitsoudis, D.A., Flouri, E.T., Chrysoulakis, N., Kamarianakis, Y., Okal, E.A., Synolakis, C.E.: Tsunami hazard in the Southeast Aegean Sea. Coast. Eng. 60, 136–148, https://doi.org/10.1016/j.coastaleng.2011.09.004, 2012.

Močkus, J.: On bayesian methods for seeking the extremum, Springer, Berlin, Heidelberg, https://doi.org/10.1007/3-540-07165-2_55, 1975.

Mohamadi, A., Yaghoubi, S.: A bi-objective stochastic model for emergency medical services network design with backup services for disasters under disruptions: An earthquake case study. Int. J. Disaster Risk Reduc. 23, 204–217, https://doi.org/10.1016/j.ijdrr.2017.05.003, 2017.

Mori. N., Takahashi, T., THE 2011 Tohoku Earthquake Tsunami Joint Survey Group: Nationwide Post Event Survey and Analysis of the 2011 Tohoku Earthquake Tsunami, Coastal Engineering Journal, 54(1), 1250001-1-1250001-27, https://doi.org/10.1142/S0578563412500015, 2012.

Mori, N., Goda, K., Cox, D.: Recent process in probabilistic tsunami hazard analysis (ptha) for mega thrust subduction earthquakes. The 2011 Japan Earthquake and Tsunami: Reconstruction and Restoration, pp. 469–485, https://doi.org/10.1007/978-3-319-58691-5_27, 2017.

Nakano, Y.: Structural design requirements for tsunami evacuation buildings in Japan. ACI Symposium Publication, 313, 2017.

Nakano, M., Murphy, S., Agata, R., Igarashi, Y., Okada, M., Hori, T.: Self-similar stochastic slip distributions on a non-planar fault for tsunami scenarios for megathrust earthquakes. Prog Earth Planet Sci, 7, 45, https://doi.org/10.1186/s40645-020-00360-0, 2020.

Nojima, N., Kuse, M., Duc, L.Q.: Mode decomposition and simulation of strong ground motion distribution using singular value decomposition. JAEE 18(2), 95–114, https://doi.org/10.5610/jaee.18.2_95, 2018.

Park, H., Cox, D.T.: Probabilistic assessment of near-field tsunami hazards: Inundation depth, velocity, momentum flux, arrival time, and duration applied to seaside, Oregon. Coast. Eng. 117, 79–96, https://doi.org/10.1016/j.coastaleng.2016.07.011, 2016.

Park, S., van de Lindt, J.W., Gupta, R., Cox, D., 2021. Method to determine the locations of tsunami vertical evacuation shelters. Nat. Hazards 63, 891–908, https://doi.org/10.1007/s11069-012-0196-3, 2012.

Pearson K.: F.R.S.: LIII. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11), 559–572, https://doi.org/10.1080/14786440109462720, 1901.

Qin, X., Motley, M.R., Marafi, N.A.: Three-dimensional modeling of tsunami forces on coastal communities. Coast. Eng. 140, 43–59, https://doi.org/10.1016/j.coastaleng.2018.06.008, 2018.

Rawls, C.G., Turnquist, M.A.: Pre-positioning of emergency supplies for disaster response. Transp. Res. B Methodol. 44, 521–534, https://doi.org/10.1016/j.trb.2009.08.003, 2010.

Salmanidou, D. M., Beck, J., Pazak, P., Guillas, S.: Probabilistic, high-resolution tsunami predictions in northern Cascadia by exploiting sequential design for efficient emulation, Nat. Hazards Earth Syst. Sci., 21, 3789–3807, https://doi.org/10.5194/nhess-21-3789-2021, 2021.

Scala, A., Lorito, S., Romano, F., Murphy, S., Selva, J., Basili, R., Babeyko, A., Herrero, A., Hoechner, A., Løvholt, F., Maesano, F. E., Perfetti, P., Tiberti, M. M., Tonini, R., Volpe, M., Davies, G., Festa, G., Power, W., Piatanesi, A., Cirella, A.: Effect of Shallow Slip

Amplification Uncertainty on Probabilistic Tsunami Hazard Analysis in Subduction Zones: Use of Long-Term Balanced Stochastic Slip Models, Pure Appl. Geophys. 177, 1497–1520, https://doi.org/10.1007/s00024-019-02260-x, 2020.

Stone, M.: Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society. Series B (Methodological), 36(2), 111–147, https://doi.org/10.1111/j.2517-6161.1974.tb00994.x, 1947.

Suppasri, A., Mas, E., Charvet, I., Gunasekera, R., Imai, K, Fukutani, Y., Abe, Y., Imamura, F.: Building damage characteristics based on surveyed data and fragility curves of the 2011 great east japan tsunami. Nat. Hazards 66, 319–341, https://doi.org/10.1007/s11069-012-0487-8, 2013.

Suppasri, A., Pakoksung, K., Charvet, I., Chua, C.T., Takahashi, N., Ornthammarath, T., Latcharote, P., Leelawat, N., Imamura, F.: Load-resistance analysis: An alternative approach to tsunami damage assessment applied to the 2011 great east Japan tsunami. Nat. Hazards Earth Syst. Sci. 19, 1807–1822, https://doi.org/10.5194/nhess-19-1807-2019, 2019.

Takase, S., Moriguchi, S., Terada, K., Kato, J., Kyoya, T., Kashiyama, K., Kotani, T.: 2D-3D hybrid stabilized finite element method for tsunami runup simulations. Comput. Mech. 58(3), 411–422, https://doi.org/10.1007/s00466-016-1300-4, 2016.

The 2011 Tohoku Earthquake Tsunami Joint Survey Group. Field survey results, official survey data, http://www.coastal.jp/tsunami2011/index.php?Field%20survey%20results (last access: 21 March 2022), 2012.

Tozato, K., Takase, S., Moriguchi, S., Terada, K., Otake, Y., Fukutani, Y., Nojima, K., Sakuraba, M., Yokosu, H.: Rapid tsunami force prediction by mode-decomposition-based surrogate modeling. Nat. Hazards Earth Syst. Sci. Disc. 22, 1267–1285, https://doi.org/10.5194/nhess-22-1267-2022, 2022.

Tozato, K.: K-Tozato/3D_tsunami_simulation: (Dataset_for_NHESS), Zenodo [data set], https://doi.org/10.5281/zenodo.6394294, 2022.

Tsuji, Y., Satake, K., Ishibe, T., Kusumoto, S., Harada, T., Nishiyama, A., Kim, H. Y., Ueno, T., Murotani, S., Oki, S., Sugimoto, M., Tomari, J., Heidarzadeh, M., Watada, S., Imai, K., Choi, B. H., Yoon, S. B., Bae, J. S., Kim, K. O., Kim, H. W.: Field surveys of tsunami heights from the 2011 Off the Pacific Coast of Tohoku, Japan, earthquake. Bulletin of the Earthquake Research Institute University of Tokyo 86, 29–279, 2011. (in Japanese with English abstract)

Wang, C., Duan, Q., Gong, W., Ye, A., Di, Z., Miao, C.: An evaluation of adaptive surrogate modeling based optimization with two benchmark problems, Environmental Modelling & Software, 60, 167-179, https://doi.org/10.1016/j.envsoft.2014.05.026, 2014.

Winter, A.O., Alam, M.S., Asce, S.M., Shekhar, K., Motley, M.R, Asce, M., Eberhard, M.O., Barbosa, A.R., Asce, A.M., Lomonaco, P., Arduino, P., Cox, D.T.: Tsunami-like wave forces on an elevated coastal structure: Effects of flow shielding and channeling. J. Waterw. Port Coast. Ocean Eng. 146, 04020021, https://doi.org/10.1061/(ASCE)WW.1943-5460.0000581, 2020.

Xiong, Y., Liang, Q., Park, H., Cox, D., Wang, G.: A deterministic approach for assessing tsunami-induced building damage through quantification of hydrodynamic forces. Coast. Eng. 144, 1–14, https://doi.org/10.1016/j.coastaleng.2018.11.002, 2019.

Zhang, W., Yun, Y.: Multi-scale accessibility performance of shelters types with diversity layout in coastal port cities: A case study in Nagoya City, Japan. Habitat Int. 83, 55–64, https://doi.org/10.1016/j.habitatint.2018.11.002, 2019.