Abstract—We propose a framework that estimates the inundation depth (maximum water level) and debris-flow-induced topographic deformation from remote sensing imagery by integrating deep learning and numerical simulation. A water and debris-flow simulator generates training data for various artificial disaster scenarios. We show that regression models based on Attention U-Net and LinkNet architectures trained on such synthetic data can predict the maximum water level and topographic deformation from a remote sensing-derived change detection map and a digital elevation model. The proposed framework has an inpainting capability, thus mitigating the false negatives that are inevitable in remote sensing image analysis. Our framework breaks limits of remote sensing and enables rapid estimation of inundation depth and topographic deformation, essential information for emergency response, including rescue and relief activities. We conduct experiments with both synthetic and real data for two disaster events that caused simultaneous flooding and debris flows and demonstrate the effectiveness of our approach quantitatively and qualitatively. Our code and data sets are available at https://github.com/nyokoya/dlsim.

Index Terms—Convolutional neural network (CNN), debris-flow mapping, flood mapping, numerical simulation.

I. INTRODUCTION

INFORMATION about inundation depth (maximum water level) and debris-flow-induced topographic deformation is essential for flood and debris-flow emergency response, including rescue and relief activities. Floods and debris flows often occur jointly following torrential rain, and their complexity makes damage assessment challenging [1]. Most remote sensing-based techniques for rapid mapping have been limited to detecting the spatial extent of flood and debris flow [2]–[7]. Numerical simulation models are capable of calculating the realistic maximum water level and topographic deformation; however, they require accurate input data and time-consuming parameter tuning.

In this article, we propose a framework that estimates the maximum water level and topographic deformation after simultaneous rainfall-triggered flood and debris flow [see Fig. 1], using remote sensing imagery and a combination of numerical simulation and deep learning. We synthesize training data comprising triplets of maximum water level, topographic deformation, and binary change maps for various artificial disaster scenarios by simulation and binarization of change information. A regression model bridges simulation and observation, inferring the maximum water level and topographic deformation from ground-surface change information derived from remote sensing imagery and topographic data, i.e., a digital elevation model (DEM). By training this type of inverse estimation model in advance, it is possible to obtain the detailed maximum water level and topographic deformation as soon as remote sensing imagery is available after a disaster.
The advantage of the proposed framework is that it breaks existing limits of remote sensing and enables rapid estimation of detailed disaster information, such as the maximum water level and debris-flow-induced topographic deformation. Using simulations circumvents the need to obtain real training data, which is expensive and complicated by the very nature of disasters, which are rare occurrences in which affected areas are difficult to access. By combining deep learning and numerical simulations, the proposed framework also learns characteristics of the analyzed quantities, such as shape and location. This mitigates the false negatives that are often inevitable in change maps derived by automated remote sensing image analysis. Our contributions are threefold, which are as follows:

1) We propose a framework that integrates remote sensing, deep learning, and numerical simulation to estimate the maximum water level and topographic deformation after floods and debris flows. The framework uses simulations to create training data to estimate unobservable geophysical information such as surface changes after very rare events.

2) We construct two data sets for our task based on real data collected after two complex disasters characterized by floods and debris flows, following torrential rains in Japan, and make these data sets freely available to the community.

3) We evaluate our methodology with two real cases and demonstrate its effectiveness both qualitatively and quantitatively.

The remainder of this article is organized as follows. Section II provides an overview of related work. Section III introduces our methodology. Section IV presents the experimental results, and Section V concludes this article with some remarks and thought about plausible future lines of research.

II. RELATED WORK

A. Flood and Landslide Mapping via Remote Sensing

Flood detection by remote sensing has been well studied, and several systems1 2 3 operate on a global scale. Flood detection at high resolution is mainly based on synthetic aperture radar (SAR) images or optical images and can be broadly divided into unsupervised and supervised approaches. In unsupervised approaches [8]–[10], predisaster and postdisaster images (e.g., intensity images of SAR), index images (e.g., spectral indices of optical data), or their differences are thresholded and then smoothed or masked out to mitigate false positives and false negatives. Supervised approaches [5], [11] identify flooded areas by detecting water in the predisaster and postdisaster images using pixel-wise classification (or semantic segmentation). Wieland and Martinis [6] developed a fully automated system based on a convolutional neural network (CNN) to automatically detect floods from multispectral images. Flood detection in urban areas using SAR images is challenging, and Li et al. [7] tackled this problem with an active self-learning CNN. On the other hand, Ohki et al. [12] used an SAR interferometric phase statistics to estimate the flood segments in built-up areas. Cohen et al. [13] developed a methodology that estimates water level from a flood inundation map and DEM for fluvial floods. The estimation of maximum water level from remote sensing images remains a challenging task when there are false negatives in a flood inundation map and also for flash floods due to the dynamics of water.

Detection of landslides, including debris flows, is another common topic in remote sensing image analysis for disaster mapping. As with flood detection, change detection using predisaster and postdisaster images is a typical approach. The use of spectral indices from multispectral images (e.g., normalized vegetation and soil index) is the simplest and effective method of detecting landslides, particularly in vegetated mountainous areas [14]–[18]. Landslide detection using SAR intensity imagery is an alternative to optical image-based approaches in adverse weather conditions. However, its accuracy is limited by the presence of layovers and shadowing, particularly for narrow debris flows [19]–[22]. Interferometric SAR has been demonstrated to be advantageous in detecting large-scale, slowly moving landslides [23]. Research on landslide detection using machine learning from optical and SAR images has recently gained popularity [24]–[28], but the collection of training data is costly. An effective means of estimating more detailed damage information, such as the amount of soil runoff and deposition, is to analyze the topography before and after the disaster using LiDAR [29]. However, LiDAR measurements are costly and thus usually not available from the initial observations of disaster areas by aircraft and helicopters. Therefore, it remains challenging to estimate debris-flow-induced topographic deformation from emergency observations.

In this work, the synergistic use of deep learning and numerical simulation provides a solution to the abovementioned problems, enabling the estimation of both maximum water level and topographic deformation after floods and debris flows from remote sensing imagery in seconds to minutes.

B. Flood and Debris-Flow Simulation

The simulation of flood hazards has been well studied and is already a common technique for estimating flood risk. Traditionally, most simulation methods require inflow to the area of interest as a boundary condition, but some methods have been developed to predict from observable rainfall data by integrating with the rain runoff process [30], [31]. However, such methods, which deal only with water, are insufficient to simulate debris flow that consists of water and sediment materials.

Simulation methods for debris flow have also been developed by several research groups. The most typical method tracks the debris flow from a certain inflow point based on the fluid dynamics method [32]–[38]. In these methods, the location and flow discharge should be given in order for simulations to be conducted; however, these data are normally based on observational information, such as debris-flow

\footnotesize{1 https://floodobservatory.colorado.edu/index.html 
2 https://floodmap.modaps.eosdis.nasa.gov/
3 https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-12939/22596_read-51634/}
trace, which can be obtained in postdisaster terms. Therefore, the usability of these simulations is not high enough for prediction purposes.

In contrast, methods that estimate both the transportation and development of debris flow [39]–[41] can be applied only from the initial location of the slope failures. By connecting with a statistical landslide prediction, predictive simulation that requires no debris-flow traces has also been proposed [42]. In this work, we employed this predictive method to generate several scenarios of rainfall-triggered flood and debris-flow damages.

C. Synergy of Deep Learning and Simulation

The collection of training data is a challenge for deep learning in all fields, including remote sensing [43]–[45]. Enormous efforts have been made to create synthetic data for training through simulation in various fields, including computer vision, bioinformatics [46], natural language processing [47], [48], and remote sensing [49]. In computer vision, for example, simulation-generated synthetic data are widely used as training data for basic tasks, such as depth estimation [50], optical flow [51], semantic segmentation [52], [53], and object detection [54]. Full-scale simulation environments have been used to create indoor and outdoor scenes for autonomous driving [55], [56], robotics [57], [58], and aerial navigation [59], [60]. Research on domain adaptation is also underway to more efficiently utilize models learned from synthetic data for the analysis of real data [52], [54].

Collecting training data for very rare events such as disasters is challenging. In particular, it is difficult to collect dense, detailed disaster information from real measurements, such as the inundation depth and topographic deformation. Inspired by the abovementioned research, this work proposes to generate training data for flood and debris flow by numerical simulation. In the event of a disaster, a deep model can then rapidly estimate the detailed damage information, using a change detection map obtained by conventional remote sensing image analysis as input.

III. METHODOLOGY

To break the limits of current remote sensing approaches, we combine two technologies: numerical simulation and deep learning. The former can generate a sufficient amount of synthetic data for training, and the latter is capable of solving complex inverse problems from a significant amount of training data. The proposed methodology comprises four modules: 1) image analysis to detect changes from bitemporal remote sensing data; 2) simulation of flood and debris flow to synthesize training data of target variables (i.e., maximum water level and topographic deformation); 3) binarization of change information to link numerical simulation and remote sensing (or synthetic and real data); and 4) regression of target variables from a binary change map and DEM based on CNNs. Fig. 2 shows an overview of our methodology’s concept.

The second and third modules deductively create output and input of training samples, respectively, for various artificial scenarios of floods and debris flows. The fourth module learns a nonlinear mapping inductively to solve the inverse problem from binary change information together with DEM to the maximum water level and topographic deformation. In a real scenario, the outcome of the first module is used as the input in the inference phase of the fourth module. Sections III-A–III-D detail the four modules.

A. Image Analysis

There are various approaches for flood and landslide (including debris flow) detection that use either optical or SAR data, as previously reviewed in Section II. In this work,
we select one of the simplest methods, based on spectral indices derived from bitemporal optical images, to ease the third module and automate the whole processing chain. We use the normalized difference vegetation index (NDVI) if a near-infrared band is available. If only an RGB image is available, which is often the case for airborne emergency observation, we use the visible atmospherically resistant index (VARI). NDVI and VARI are calculated as follows:

\[ \text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})} \]
\[ \text{VARI} = \frac{(\text{Green} - \text{Red})}{(\text{Green} + \text{Red} - \text{Blue})}. \] (1)

We calculate either NDVI or VARI from predisaster and postdisaster optical images and detect areas where vegetation coverage decreases with changes due to floods and debris flows. Hard thresholding is used in this work, and the threshold values to judge whether a pixel is vegetated or not are empirically set to 0.7 for NDVI and 0 for VARI. We assume that debris flows occur in vegetated mountainous areas. Note that the method used cannot detect the inundation of nonvegetation areas and also narrow debris flows occluded by tree crowns. Therefore, the change detection result provides only partial information on the flood and debris-flow extent areas, with possible missing (i.e., false negatives). Regression models (see Section III-D) learn to inpaint such missing information from synthetic data created by simulation (see Section III-B) and binarization (see Section III-C).

### B. Numerical Simulation

In this work, the simulation methods developed in [62] were used. Dynamics of the debris flow can be described by the governing equations based on shallow water equations that take erosion and deposition processes into consideration. When erosion takes place, the water and sediment at the ground/river bed are retrieved into the flow body. On the other hand, both the sediment and water are trapped in the bed when deposition takes place. To express these processes, the following equations are employed:

\[ \frac{\partial S}{\partial t} + \frac{\partial E}{\partial x} + \frac{\partial F}{\partial y} = \mathbf{S} \] (2)

where \( \mathbf{S} \) is the conservative variable, flux for \( x \)- and \( y \)-directions, and source vectors, respectively. The term \( h \) is the flow depth, \( u \) and \( v \) are the velocity of \( x \)- and \( y \)-directions, respectively, \( C \) is the sediment concentration of flow body, \( z_b \) is the river/ground bed elevation, \( g \) is the gravity acceleration, and \( \epsilon \) is the eddy momentum diffusivity. The terms \( S_{0x} \) and \( S_{0y} \) are the topographical gradients for \( x \)- and \( y \)-directions, respectively. The terms \( S_{fx} \) and \( S_{fy} \) are the frictional gradients for \( x \)- and \( y \)-directions, respectively, which are calculated by different equations for three flow modes, stony debris flows, hyperconcentrated flow, and water flows, depending on the concentration of flow body \( C \) [63]. The term \( i \) is the erosion/deposition velocity, which is calculated by the balance of the equilibrium concentration obtained by the function of water surface gradients. It is also calculated by different functions for the three flow modes. In addition, in this study, fluidization rate \( \gamma \) is introduced to consider the transformation of the fine solid material to fluid. The effective specific weight of fluid material \( \rho \) and its concentration in deposited material \( C_\text{s} \) are modified by the following equation:

\[ \rho = \frac{\gamma \sigma C_\text{d} + \rho_0 (1 - C_\text{d})}{\gamma C_\text{d} + (1 - C_\text{d})}, \quad C_\text{s} = C_\text{d} (1 - \gamma) \] (4)

where \( C_\text{d} \) is the original sediment concentration in the deposited material and \( \rho_0 \) is the specific weight of water.

For numerical modeling, we used the MacCormack scheme with artificial viscosity, which is categorized into a two-step scheme in finite-difference methods (FDMs). The code is parallelized by employing both MPI and OpenMP and therefore can be simulated on large-scale supercomputers [42].

The location of the initiation points of a debris flow [see Fig. 3(c)] is required in order to conduct the simulation and can be obtained only after the disaster event. The statistically predicted initiation points can substitute for the actual data; however, the predicted damage is not uniquely derived [42]. Therefore, the method can generate many possible damage results.

In this method, in order to generate the artificial damage data, we use the distribution of a probability, which is obtained by logistic regression employing the actual disaster data and topographical data. In the regression, local slopes, accumulation, and tangential and plan curvatures are selected as explanatory variables. The obtained probability is shown in Fig. 3(b). In order to use these data for simulation inputs, sets of pseudorandom numbers were used to convert to the binary point distribution shown in Fig. 3(c) as an example. From elevation [see Fig. 3(a)] and points [see Fig. 3(c)], the simulation calculates the transport of debris and water flow temporally and spatially. We select the maximum water level and final terrain deformation as the target variables, which represents the damage from the hazards, as shown in Fig. 3(d) and (e).

### C. Binarization of Change Information

One key objective of the proposed methodology is to ensure the transferability of regression models trained on synthetic data to real data. Based on the physical reasoning behind the change detection method presented in Section III-A, we attempt to create synthetic binary change maps that resemble a real change detection map obtained by remote sensing.
image analysis. In addition, inspired by domain randomization [52], we try to make the distribution of the synthetic binary data sufficiently wide and varied so that regression models trained on the synthetic data work robustly with real data.

The binary change map obtained from the observation contains false positives and false negatives due to the characteristics of the detection method and the observed data, and therefore, it is necessary to perform binarization of change information (i.e., maximum water level and topographic deformation) simulating these errors. When using the change detection method based on vegetation-related spectral indices from optical images, it should be noted that floods and debris flows are not detected in areas that satisfy one of the following two conditions: 1) areas with low NDVI before the disaster and 2) areas where occlusion occurs due to the effects of tree crowns and incident angles. In the binarization, the former can be easily synthesized by using the NDVI/VARI image derived from the predisaster optical imagery. In the latter case, more detailed 3-D information is required to perform the model-based simulation. For simplicity, we perform morphological erosion to synthesize false-negative pixels. Pixels having decreased in vegetation due to other reasons than a disaster are erroneously recognized as flood or debris-flow areas. Since it is difficult to synthesize such phenomenon based on a model, we randomly add noise to synthesize false positive pixels. Furthermore, the simulation results include floods and debris flows everywhere, and there are very few negative examples in which no flood or debris flow has occurred. We use a cut out [47] for data augmentation to enforce regression models to output 0 if there is no change.

D. CNN-Based Regression

CNNs have achieved great success in regression problems [64], [65]. By estimating the maximum water level and tomographic changes separately by individual CNNs, we are able to account for their different physical properties.

Regression models learn a nonlinear mapping $f_\theta$ from input $x$, which is composed of binary change map and slope images, to output $y$ (maximum water level or topographic deformation): $f_\theta : x \rightarrow y$. 

---

**Fig. 3.** Input data for the simulation and one example of the output. (a) Elevation of the target region. (b) Distribution of the probability to be initiation points obtained by the logistic regression. (c) Example of the initiation points generated by (b) and pseudorandom numbers. (d) and (e) Example of results on case 54 for maximum water level and terrain deformation, respectively. The size of points is enlarged to be displayed in (c) [42].
The smooth $L_1$ loss (or Huber loss) is used for robust regression of target variables

$$\mathcal{L}_\delta(y, f_\theta(x)) = \begin{cases} 
\frac{1}{2} (y - f_\theta(x))^2, & |y - f_\theta(x)| \leq \delta \\
\delta |y - f_\theta(x)| - \frac{1}{2} \delta^2, & \text{otherwise}
\end{cases}$$

(5)

The smooth $L_1$ loss combines the advantage of the $L_2$ loss (gradient decreases when the loss gets close to local minima) and the $L_1$ loss (less sensitive to outliers). We set $\delta$ to 1.

For the regression models, we investigate variations of two architectures, Attention U-Net [61] and LinkNet [66], which have consistently shown high performance in semantic segmentation [67]. We adopt model fusion to further improve accuracy by taking the average of the outputs of the two models as the final output. Fig. 4 shows the architectures of Attention U-Net and LinkNet investigated in this work with the size of feature maps. The size of input images is $256 \times 256$ pixels.

U-Net is an encoder–decoder structure with skip connections, which shares the information learned by the encoder with the decoder through concatenation [68]. We adopt a modified version of the architecture proposed in [61], which incorporates a self-attention mechanism in U-Net with contextual information extracted at a coarser scale. Each encoder block is composed of two convolutional layers, each followed by batch normalization and a rectified linear unit (ReLU). The attention gate module emphasizes the features that are valid for a given task and suppresses irrelevant features when concatenating features extracted by the encoder with those of the decoder through the skip connections. Fig. 5 shows the design of our attention gate module modified from [61].

The original LinkNet also has an encoder–decoder structure with residual blocks and skip connections, but it shares the information learned by the encoder with the decoder through additive operations. In this work, we do not use residual blocks but use the same one as Attention U-Net, which leads to a more compact model.

We use the Adam solver [69] for optimization with a learning rate of 0.0001. Xavier initialization is used to initialize the weights. The batch size is 32 and the number of epochs is 200. Our model is implemented in the PyTorch framework [70] with four NVIDIA Tesla V100 16-GB GPUs.

IV. EXPERIMENTS

In this section, we demonstrate the usefulness of the proposed framework for estimating the maximum water level and topographic deformation using data from two complex disasters where floods and debris flows occurred simultaneously. The performance of the proposed CNN-based methods—Attention U-Net, LinkNet, and the fusion—is evaluated quantitatively and qualitatively with both synthetic and real data experiments.

A. Data Sets

The experiments focus on the analysis of floods and debris flows caused by two disaster events: 1) torrential rain in July 2017 in Northern Kyushu, Japan, and 2) torrential rain in July 2018 in Western Japan. Hereinafter, we refer to data sets for these events as Northern Kyushu 2017 and Western Japan 2018. Fig. 6 shows the study scenes for the two events. Each of them covers $8 \text{ km} \times 9 \text{ km}$ with $1600 \times 1800$ pixels at a ground sampling distance of 5 m, including a variety of geographically different areas (e.g., urban, agriculture fields, and mountain forests). We used the DEM released by the
Geospatial Information Authority of Japan (GSI) for the simulation. Details of each data set are given next.

1) Northern Kyushu 2017: Torrential rains hit northern Kyushu on July 5 and 6, 2017, and damaged many houses (336 were completely destroyed and 1096 partially destroyed) and caused human casualties. In this work, we analyze the area surrounding the city Asakura, where debris flows and inundation occurred simultaneously.

For remote sensing images, we use a predisaster Sentinel-2 image and aerial photographs released by the GSI soon after the disaster [see Fig. 6(a)]. The mosaic aerial imagery was made from two different observations. We use a debris-flow and inundation extent map manually created by experts from aerial photographs as reference data for evaluation. Clouds and missing areas included in the aerial photographs used in the analysis were masked out manually and were not included in the evaluation.

By the simulation method presented in Section II-B, we generated 60 sets of input points by changing the random seed and conducted simulations for the 60 cases simultaneously using the K computer installed in the RIKEN Center for Computational Science in Japan. By using the reference map of the flood and debris-flow extent map created by the visual interpretation as the input of the simulation, we obtained the maximum water level and topographic deformation with high accuracy and used it as reference data for evaluation. We refer to this simulation result as the reference simulation.

2) Western Japan 2018: The second data set was collected before and after the floods and debris flows caused by torrential rains in western Japan during the period June 28–July 8, 2018. These rains caused river flooding, inundation, flash floods, and debris flows in many areas of western Japan, with the death toll exceeding 200. In this work, we analyze an area in Higashihiroshima, which was severely affected by the floods and debris flows.

We use the predisaster and postdisaster Sentinel-2 images [see Fig. 6(b)] as input for remote sensing image analysis. A debris-flow extent map created by experts’ visual interpretation from aerial photographs is used as a reference for evaluation of disaster extent detection. For the Western Japan 2018 data set, predisaster and postdisaster LiDAR-derived digital terrain models (DTMs) are also available for the study area. The difference in the DTMs is used as a reference for numerical evaluation of topographic deformation. Note that we mask out (or set to zero) the values of the topographic deformation reference at unchanged pixels in the debris-flow extent reference map.

We generated ten sets of input points and conducted the simulation, changing the ten values of \( \gamma \) from 0 to 0.6 (i.e., \( \gamma = 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6 \)) for each set of input points; therefore, 100 results were generated. By using the debris-flow initiation points manually annotated by experts from Hiroshima University for the input of the simulation, we obtained ten reference simulations of the maximum water level and topographic deformation for the ten parameter sets. We quantitatively evaluate the accuracy of the ten reference simulations of topographic deformation in comparison with the LiDAR-derived reference and use the best one as the reference simulation. The accuracy of the reference simulation for topographic deformation is discussed in Section IV-D.

B. Evaluation Metrics

We use three metrics for quantitative evaluation: 1) the root-mean-square error (RMSE); 2) the intersection over union (IoU); and 3) the log-scaled histogram intersection (LSHI).

The pixel-wise accuracy of the predictions obtained by the proposed method is evaluated by RMSE using the reference simulation of maximum water level and topographic deformation. We use RMSE in both synthetic and real data experiments. In the real data experiment with the Western Japan 2018 data set, RMSE is computed for topographic deformation using the LiDAR-derived reference, which is denoted as RMSE (real).

For evaluating the accuracy of detecting affected areas, we calculate IoU by comparing the binary change map obtained by simple thresholding from the maximum water level and topographic deformation with the reference map of the flood and debris-flow extent map obtained by visual interpretation. IoU is used only in the experiments with real data where visual interpretation results are available. We adopt the best threshold for all results for fairness.

In addition to the spatial (pixel-by-pixel) details of the disaster, it is also important to understand the overall scale of the disaster. We adopt LSHI to measure the accuracy in terms of the overall scale of flood and debris flow. We calculate LSHI in both synthetic and real data experiments. As with

Fig. 6. (a) (Left) Predisaster Sentinel-2 imagery and (Right) Postdisaster aerial RGB imagery over the study area of Northern Kyushu 2017. (b) (Left) Predisaster and (Right) Postdisaster Sentinel-2 images over the study area of Western Japan 2018.
RMSE, LSHI is computed for topographic deformation using the LiDAR-derived reference for the Western Japan 2018 data, which is denoted as LSHI (real).

For the comparison of RMSE and LSHI, the baseline is the average of all simulation outcomes for each set of parameters, which can be regarded as the expected value of target variables with Monte Carlo simulations. For the Western Japan 2018 data set, we have the ten average simulations and use the best one as the baseline for each test case. Note that the average simulation is more accurate than any individual simulation. For IoU, the binary change detection result of remote sensing image analysis and the reference simulation are also compared. For RMSE (real) and LSHI (real) of the topographic deformation in the Western Japan 2018 real data experiment, we evaluate the reference simulation as well.

### C. Synthetic Data Experiment

In the synthetic data experiments, we evaluate the generalization ability of the regression models for unseen data. We split the simulation cases into training and test sets as follows.

1) **Northern Kyushu 2017**: We use 40 cases for training and 20 cases for testing; 1750 patches with nonoverlapping sampling are used for training.

2) **Western Japan 2018**: We use 70 cases for training and 30 cases for testing; 3010 patches with nonoverlapping sampling are used for training.

The average simulation of the training data is compared to the baseline.

1) **Quantitative Results**: Table I shows the accuracy of the estimated maximum water level and topographic deformation calculated by the regression models based on Attention U-Net, LinkNet, and their fusion. All the models outperform the average simulation in RMSE for both data sets, indicating that the networks successfully learn a nonlinear mapping with respect to different binary change maps rather than outputting a simple average of training data. The superiority of the proposed methods in RMSE is more evident with the Western Japan 2018 data set. One possible reason for the large RMSE of the average simulation in the Western Japan 2018 data set is that the average simulation, which is the sample average, deviates from the true expected value since only about seven cases are included in the training data for each parameter set. The RMSE values of our results for the Western Japan 2018 data set are larger than those for the Northern Kyushu 2017 data set due to the large variation in the simulation results. The average simulation shows better LSHI values for Western Japan 2018. The reason for this result is that the average simulation (the best among the ten average simulations) uses the same parameters as the test case and the scale of the disaster is very similar to each test case, while the regression model uses all the training data generated with the different parameters and the scale of the disaster is not optimized.

2) **Qualitative Results**: Fig. 7 shows four example patches of the binary change map of the input, the average simulation, our result with model fusion, and the reference simulation used as ground truth for both maximum water level and topographic deformation. Our results resemble the reference simulation and outperform the average simulation by exploiting the hints of disaster locations from the binary change map. The effectiveness of the proposed fusion method compared to the average simulation is noticeable in the places indicated by the green circles in the figure. False positives and false negatives are inevitable for the average simulation since it only represents the expected value of the target variable based on the Monte Carlo simulation, whereas the proposed fusion method accurately estimates the target variables for each different scenario of the test set, based on the disaster locations included in the binary change detection map. The third and fourth columns in Fig. 7(a) and all examples in Fig. 7(b) clearly demonstrate that the proposed fusion method can estimate the target variables in the missing parts of the binary change map.

### D. Real Data Experiment

In the real data experiments, for our proposed methods, we used all the synthetic data for training and the remote sensing-derived change detection maps as input for the inference. We again adopt the average simulation of the training data as the baseline of simulation results. Furthermore, the remote sensing-derived change detection maps are regarded as the baseline of the remote sensing results. We process densely overlapping patches with a stride of 8 and take the average as our final output to mitigate patch-wise boundary artifacts.

1) **Quantitative Results**: Table II summarizes the quantitative evaluation results for the real data experiment. The

---

**Table I**

**Numerical Evaluation Results in Synthetic Data Experiments**

| Approach | Northern Kyushu 2017 | Western Japan 2018 |
|----------|----------------------|---------------------|
|          | RMSE | LSHI | RMSE | LSHI | RMSE | LSHI | RMSE | LSHI |
| Average Sim. | 0.1567 | 0.9602 | 0.1499 | 0.8746 | 0.3524 | 0.9772 | 0.3585 | 0.9701 |
| Att. U-Net | 0.1322 | 0.9674 | 0.1302 | 0.9222 | 0.2837 | 0.9727 | 0.3193 | 0.9132 |
| LinkNet | 0.1293 | 0.9582 | 0.1319 | 0.9183 | 0.3007 | 0.9312 | 0.3425 | 0.8992 |
| Fusion | **0.1246** | 0.9652 | **0.1246** | 0.9191 | **0.2784** | 0.9546 | **0.3161** | 0.9176 |
proposed methods show the best accuracy for all the evaluation metrics with the Northern Kyushu 2017 data set and also for IoU, RMSE (real), and LSHI (real) that are based on the real reference data for the Western Japan 2018 data set. The results demonstrated that by integrating remote sensing, simulation, and deep learning, it is possible to extract physically semantic disaster information that cannot be obtained by either remote sensing image analysis or simulation alone.

In the Northern Kyushu 2017 experiment, the IoU scores achieved by our approach clearly outperform the change detection result of remote sensing image analysis, indicating that the synergistic use of simulation and deep learning

| Approach        | Northern Kyushu 2017 | Western Japan 2018 |
|-----------------|----------------------|---------------------|
|                 | Water Level | Topographic Def.  | Water Level | Topographic Def. |
|                 | RMSE  | IoU  | LSHI | RMSE  | IoU  | LSHI | RMSE  (real) | IoU  | LSHI (real) | LSHI |
| Detection       | 0.3558 | 0.3054 | 0.7739 | 0.6419 | 0.1024 | 0.9256 | 0.5328 | 0.3946 | 0.0889 | 0.6261 | 0.8684 |
| Average Sim.    | 0.3649 | 0.2872 | 0.5119 | 0.2365 | 0.2378 | 0.7391 | 0.2303 | 0.4006 | 0.3018 | 0.7129 | 0.5137 |
| Att. U-Net      | 0.3889 | 0.4433 | 0.7811 | 0.1597 | 0.3765 | 0.8910 | 0.7186 | 0.3160 | 0.6586 | 0.2276 | 0.4020 | 0.3003 | 0.6984 | 0.5070 |
| LinkNet         | 0.3899 | 0.4420 | 0.7924 | 0.1688 | 0.3536 | 0.8404 | 0.7121 | 0.3414 | 0.7155 | 0.2276 | 0.4020 | 0.3003 | 0.6984 | 0.5070 |
| Fusion          | 0.3852 | 0.4454 | 0.7840 | 0.1654 | 0.3790 | 0.8575 | 0.7131 | 0.3199 | 0.6854 | 0.2266 | 0.4000 | 0.3045 | 0.6903 | 0.4987 |
| Reference Sim.  | 0.4070 | 0.3054 | 0.7739 | 0.2266 | 1.0000 | 0.0000 | 0.4497 | 0.0000 | 0.1000 | 0.6576 | 1.0000 |
Fig. 8. (a), (c), and (e) Maximum water level and (b), (d), and (f) topographic deformation of (a) and (b) reference simulation, (c) and (d) average simulation, and (e) and (f) our method with (g) reference map of (blue) flood and (brown) debris-flow extent and (h) binary change detection map. Mask for clouds and no data is shown in gray.

successfully improved the detection of disaster extent areas. The fact that our IoU scores are comparable to or even better than those of the reference simulation suggests that high accuracy has been achieved for flood and debris-flow detection. The RMSE and LSHI of our results are better than those of the average simulation with large
margins, which supports the effectiveness of the proposed framework.

In the Western Japan 2018 experiment, our IoU scores are comparable to that of the change detection by remote sensing image analysis because the improvement in (or inpainting of) false negatives was offset by the increase in false positives. RMSE and LSHI results in the Western Japan 2018 experiment must be carefully interpreted due to the inaccuracy of the reference simulation, as is apparent in its limited scores in RMSE (real) and LSHI (real) for topographic deformation. In other words, RMSE (real) and LSHI (real) are the most important metrics, and our method shows its superiority over the best-effort simulation.

The computational time of the inference for each study area (an area of 72 km$^2$) is 799 s with densely overlapping patch processing (i.e., stride is 8), while it takes only 1.4 s with nonoverlapping patches. The conventional simulation-based approach takes days to produce the final output due to the high computational work needed to simulate on a basin scale [71] and the time required to obtain accurate input data (e.g., starting points of debris flows), which can be made by visual interpretation of remote sensing images, and to tune the parameters. Therefore, the proposed method is much more rapid than the conventional method.

2) Qualitative Results: Fig. 8 shows the visual results of the Northern Kyushu 2017 experiment by comparing the reference simulation, the average simulation, and our estimation results of the maximum water level and topographic deformation together with the visual interpretation reference and the flood and debris-flow detection map obtained by image analysis.
It can be visually observed that the results of the proposed fusion method achieve a more accurate estimation than the average simulation, which is consistent with the quantitative evaluation results. The average simulation includes overestimation in the northern and southern regions, whereas the proposed fusion method has fewer such errors. The proposed fusion method successfully detects the inundation areas along the rivers, which could not be detected by the simple remote sensing image analysis. Even if the flood area is missing in the change detection results, our method succeeded in completing physically semantic disaster information in the missing area by using the change detection results of the surrounding area, such as debris flows in a mountainous area. Our approach failed to detect pixels having small absolute values and spatially small shapes in the reference simulation. When the changes are not detected by remote sensing image analysis due to occlusion by the tree canopy, it is difficult to estimate target variables that have very small patterns, such as a small/narrow debris flow, because each change (e.g., debris flow) occurs independently.

Fig. 9 shows the visual results of topographic deformation estimation for the Western Japan 2018 experiment. The advantage of our fusion method is much more evident. Both the reference simulation and the average simulation overestimate topographic deformation, while our result best resembles the LiDAR-derived reference. From Fig. 9(c) and (d), it can be observed that the estimated locations of debris flows are accurate and the overall scale of the disaster is similar, which was also confirmed by the quantitative evaluation in Table II. We can find the inpainting capability in the green circles of Fig. 9. Sediment deposition is not detected in remote sensing image analysis and is overestimated in the reference simulation. Our fusion method successfully estimates the sediment deposition using the hints of debris-flow detection in upper streams. The estimation of sediment deposition is challenging as shown in the enlarged blue squares. The extent to which sediment is washed downstream depends largely on the disaster scenario. If the synthetic data do not include cases that are at least local but close to the real data scenario, our method may fail to estimate long-range displaced sediment deposition.

It should be noted that the reference simulation and the visual interpretation reference do not always match, as shown in Figs. 8 and 9. This is evident numerically from the fact that the IoU of the reference simulation is much lower than 1 for both data sets. For example, the area indicated by the green circle in Fig. 8 is annotated as a debris-flow area by human experts; however, the area is underestimated in the reference simulation. Since the proposed method uses the results of remote sensing image analysis as input and finds the corresponding topographical changes, our fusion method estimates that a large amount of sediments flowed out at places where debris flows were detected. This result is consistent with the visual interpretation reference but does not match the simulation reference data, leading to higher IoU and RMSE. Fig. 9(a), (d), and (e) clearly shows the gap between the reference simulation and the visual interpretation reference and the LiDAR-derived reference for topographic deformation. The gap implies the difficulty of making a numerical simulation that matches the observation, even when the human-annotated debris-flow starting points are used for the simulation due to the complex dynamics of mixed water and debris. Our framework overcomes this limitation by taking advantage of remote sensing, simulation, and deep learning.

The experimental results on real data demonstrate that the models learned from the simulated data are applicable to real data. The proposed framework can be extended to other regions or countries. However, it is worth noting that the trained networks with the presented two data sets are not directly applicable to other places because the characteristics of flood and debris-flow disasters are different among different geographic regions. Our recommendation is to perform simulations in advance over areas of interest and train models with respect to different geographic regions to ensure the performance.

V. CONCLUSION AND FUTURE LINES OF INQUIRY

In this article, we proposed a framework that enables rapid estimation of the maximum water level and topographic deformation after simultaneous floods and debris flows through the use of remote sensing imagery and topographic data, a calculation that was not possible by using remote sensing image analysis and simulation alone. Our framework generates synthetic data of target variables and corresponding binary change maps based on simulation. It trains CNN-based regression models that take a binary change map and DEM as input and produce the maximum water level and topographic deformation as output. The CNN-based regression model can compensate for the missing part of the input detection map, which simplifies change detection and makes the whole process automatic and fast. Experiments based on two disaster events demonstrated the effectiveness of our framework both quantitatively and qualitatively.

In our future research, we intend to develop techniques that allow us to directly use remote sensing images instead of change maps as input for regression as well as techniques to scale up the system so that it can function over a much larger area.

ACKNOWLEDGMENT

The authors would like to thank Hiroshima Prefecture, Japan, for providing the LiDAR data used in this work.

REFERENCES

[1] M. J. Froude and D. N. Petley, “Global fatal landslide occurrence from 2004 to 2016,” Natural Hazards Earth Syst. Sci., vol. 18, no. 8, pp. 2161–2181, Aug. 2018. [Online]. Available: https://www.nat-hazards-earth-syst-sci.net/18/2161/2018/

[2] S. Chignell, R. Anderson, P. Evangelista, M. Laituri, and D. Merritt, “Multi-temporal independent component analysis and Landsat 8 for delineating maximum extent of the 2013 colorado front range flood,” Remote Sens., vol. 7, no. 8, pp. 9822–9843, Jul. 2015. [Online]. Available: http://www.mdpi.com/2072-4292/7/8/9822

[3] S. Schlaffer, P. Matgen, M. Hollaus, and W. Wagner, “Flood detection from multi-temporal SAR data using harmonic analysis and change detection,” Int. J. Appl. Earth Observ. Geosci., vol. 38, pp. 15–24, Jun. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0303243414002645
[4] Y. Li, S. Martinis, S. Plank, and R. Ludwig, “An automatic change detection approach for rapid flood mapping in Sentinel-1 SAR data,” Int. J. Appl. Earth Observ. Geoinf., vol. 73, pp. 123–135, Dec. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0303243418302782

[5] X. Tong et al., “An approach for flood monitoring by the combined use of Landsat 8 optical imagery and COSMO-SkyMed radar imagery,” ISPRS J. Photogramm. Remote Sens., vol. 182, pp. 178–191, Jan. 2019.

[6] M. Wieland and S. Martinis, “A modular processing chain for automated flood monitoring from multi-spectral satellite data,” Remote Sens., vol. 11, no. 19, p. 2330, Oct. 2019.

[7] Y. Li, S. Martinis, and M. Wieland, “Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence,” ISPRS J. Photogramm. Remote Sens., vol. 152, pp. 205–213, Aug. 2019.

[8] S. Martinis, A. Tewe, and S. Voigt, “Towards operational near-real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data,” Natural Hazards Earth Syst. Sci., vol. 9, no. 2, pp. 303–314, Mar. 2009.

[9] S. Martinis, A. Tewe, and S. Voigt, “Unsupervised extraction of flood-induced backscatter changes in SAR data using Markov image modeling on irregular graphs,” IEEE Trans. Geosci. Remote Sens., vol. 49, no. 1, pp. 251–263, Jan. 2011.

[10] T. Sakamoto, N. Van Nguyen, A. Kotera, H. Ohno, N. Ishitsuuka, and Y. Yokozawa, “Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong delta from MODIS time-series imagery,” Remote Sens. Environ., vol. 109, no. 3, pp. 205–213, Aug. 2007.

[11] N. Mueller et al., “Water observations from space: Mapping surface water from 25 years of landsat imagery across Australia,” Remote Sens. Environ., vol. 174, pp. 341–352, Mar. 2016.

[12] N. Wrobly, S. Martinis, and M. Wieland, “A modular processing chain for automated flood monitoring from multi-spectral satellite data,” Remote Sens., vol. 11, no. 19, p. 2330, Oct. 2019.

[13] B. Adriano, N. Yokoya, H. Miura, M. Matsuoka, and S. Koshimura, “A semiautomatic pixel-object method for detecting landslides using multitemporal ALOS-2 intensity images,” Remote Sens., vol. 12, no. 3, p. 561, Feb. 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/3/561

[14] Z. Zhao, Z. Lu, Q. Zhang, and J. de la Fuente, “Large-area landslide detection and monitoring with ALOS/PALSAR imagery data over Northern California and Southern Oregon, USA,” Remote Sens. Environ., vol. 124, pp. 348–359, Sep. 2012.

[15] D. Tien Bui et al., “Landslide detection and susceptibility mapping by AIRSAR data using support vector machine and index of entropy models in Cameron highlands, Malaysia,” Remote Sens., vol. 10, no. 10, p. 1527, Sep. 2018.

[16] S.-J. Park, C.-W. Lee, S. Lee, and M.-J. Lee, “Landslide susceptibility mapping and comparison using decision tree models: A case study of Junumjin area, South Korea,” Remote Sens., vol. 10, no. 10, p. 1545, Sep. 2018.

[17] K. Burrows, R. J. Walters, D. Milledge, K. Spans, and A. L. Densmore, “A new method for large-scale landslide classification from satellite SAR,” Remote Sens., vol. 10, no. 7, pp. 1–21, Jan. 2019.

[18] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. Meena, D. Tiede, and J. Aryal, “Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection,” Remote Sens., vol. 11, no. 2, p. 196, Jan. 2019.

[19] Y. Wang, Z. Fang, and H. Hong, “Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China,” Sci. Total Environ., vol. 666, pp. 795–993, May 2019, doi: 10.1016/j.scitotenv.2019.02.263.

[20] H. Miura, “Fusion analysis of optical satellite images and digital elevation model for quantifying volume in debris flow disaster,” Remote Sens., vol. 11, no. 9, p. 1096, May 2019.

[21] D. Yanazaki, S. Kanae, H. Kim, and T. Oki, “A physically based description of floodplain inundation dynamics in a global river routing model,” Water Resour. Res., vol. 47, no. 4, p. 1–21, Apr. 2011. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009726

[22] T. Sayama, G. Ozawa, T. Kawakami, S. Nabetaka, and K. Fukushima, “Rainfall–runoff–inundation analysis of the 2010 Pakistan flood in the Kabul river basin,” Hydrolog. Sci. J., vol. 57, no. 2, pp. 298–312, Feb. 2012, doi: 10.1080/02626667.2011.644245.

[23] D. Rickenmann, D. Laigle, B. W. McArdell, and J. Hübli, “Comparison of 2D debris-flow simulation models with field events,” Comput. Geosci., vol. 10, no. 2, pp. 241–264, May 2006.

[24] D. Rickenmann, D. Laigle, B. W. McArdell, and J. Hübli, “Comparison of 2D debris-flow simulation models with field events,” Comput. Geosci., vol. 10, no. 2, pp. 241–264, May 2006.

[25] Y. Xu, B. Li, and D. Li, “A new method for large-scale landslide classification from satellite SAR,” Remote Sens., vol. 10, no. 11, p. 1781, Nov. 2018.

[26] D. Revellino, O. Hungr, F. M. Guadagno, and S. G. Evans, “Velocity and runout simulation of destructive debris flows and debris avalanches using saturation density and entrainment on debris flow runout modeling: Examples from the Swiss Alps,” Natural Hazards Earth Syst. Sci., vol. 17, no. 17, pp. 2569–2583, Nov. 2017.

[27] N. Wrobly, S. Martinis, and M. Wieland, “Remote sensing image caption generation,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 4, pp. 2183–2195, Apr. 2018.

[28] A. Venkataraman, S. Venkataraman, and V. Venkatesan, “Remote sensing-based landslide mapping and prediction for natural disaster management,” Remote Sens., vol. 11, no. 7, p. 760, Mar. 2019. [Online]. Available: https://www.mdpi.com/2072-4292/11/7/760
Wei He (Member, IEEE) received the B.S. degree from the School of Mathematics and Statistics, Wuhan University, Wuhan, China, in 2012, and the Ph.D. degree from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, in 2017. He was a Researcher with the RIKEN Center for Advanced Intelligence Project, Tokyo, Japan, from 2018 to 2020, where he is a Research Scientist with the Geoinformatics Unit. His research interests include image quality improvement, remote sensing image processing, matrix/tensor analysis, and deep learning.

Gerald Baier (Member, IEEE) received the M.Sc. degree in electrical engineering from the Université catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium, and the Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2012, and the Ph.D. degree in synthetic aperture radar interferometry from the Technical University of Munich, Munich, Germany, in 2018. From 2014 to 2018, he was with the Department of SAR Signal Processing, Remote Sensing Technology Institute, German Aerospace Center (DLR), Munich. Since 2018, he has been a Post-Doctoral Researcher with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Tokyo, Japan. His research interests include signal processing, machine learning, synthetic aperture radar, and high-performance computing.

Bruno Adriano (Member, IEEE) received the M.Eng. degree in disaster management from the National Graduate Institute for Policy Studies, Tokyo, Japan, and the International Institute of Seismology and Earthquake Engineering, Building Research Institute, Tsukuba, Japan, in 2010, and the Ph.D. degree from the Department of Civil and Environmental Engineering, Graduate School of Engineering, Tohoku University, Sendai, Japan, in 2016. From 2016 to 2018, he was a Research Fellow with the Japan Society for the Promotion of Science and the International Research Institute of Disaster Science, Tohoku University. Since 2018, he has been a Post-Doctoral Researcher with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Tokyo. His research interests include integrating machine learning and high-performance computing for disaster management using remote sensing technologies.

Hirokuni Miura (Member, IEEE) received the M.Eng. and Ph.D. degrees from the Department of Built Environment, Tokyo Institute of Technology, Tokyo, Japan, in 2001 and 2004, respectively. He was a Post-Doctoral Research Fellow and an Assistant Professor with the Tokyo Institute of Technology from 2004 to 2012. He is an Associate Professor with the Graduate School of Advanced Science and Engineering, Hiroshima University, Hiroshima, Japan. He is also a member of Resilience Research Center, Hiroshima University, established after the July 2018 torrential rain disaster in western Japan. He has conducted disaster-related researches not only for spatial data analysis using remote sensing and topographic data but also for earthquake engineering. His research interest includes deep learning-based identification of disaster-induced damage from remote sensing data.

Satoru Oishi (Member, IEEE) received the M.Eng. and Ph.D. degrees from the Department of Civil Engineering, Kyoto University, Kyoto, Japan, in 1993 and 1998, respectively. He was a former Director of the Research Center for Urban Safety and Security (RCUSS), Kobe University, Kobe, Japan, which was established just after the 1995 earthquake to accumulate related research. From 1993 to 1999, he was a Research Associate with Disaster Prevention Research Institute, Kyoto University. From 2000 to 2009, he was an Associate Professor with the University of Yamanashi, Kofu, Japan. Since 2009, he has been a Professor with Kobe University. He is a Professor with RCUSS and the Department of Civil Engineering, Graduate School of Engineering, Kobe University. He is also the Team Leader of the Computational Disaster Mitigation and Reduction Research Team, RIKEN Center for Computational Science (R-CCS), Kobe. As the Team Leader of RIKEN R-CCS, he is also working for promoting high-performance computing simulation of disaster, including earthquake, tsunami, and weather-related disaster. He conducted disaster-related research for more than 25 years. He also conducted many observation experimental projects like BIWAKO Project and X-BAIU to measure atmospheric data and compare the measured data with numerically simulated results. He studied river engineering and sabo engineering at Yamanashi, Japan, where many target rivers and tributaries could be found. He collected some samples from rivers to detect the sources of sediments. He also conducted many field studies abroad under the project of CREST. This activity helped him to expand his international activity. He developed a mini X-MP radar with a private company at Hyogo, Japan. He used the radars for sewage water pipe operation, increasing citizens’ awareness against flood and inundation in Japan’s local area, and volcanic debris-flow warnings. He mainly focused on cumulus cloud generation and developing by using numerical simulation of the cloud-resolving model. His main research interests include the application of remote sensing techniques for disaster prevention.