Regular Paper

A Study on Effectiveness of SNS Data in Flood Estimation

Kei Hiroi†, Akihito Kohiga‡, and Yoichi Shinoda‡

†Disaster Prevention Research Institute, Kyoto University, Japan
‡Japan Advanced Institute of Science and Technology, Japan
hiroi@dimsis.dpri.kyoto-u.ac.jp

Abstract - Reports of damage posted to social networking services (SNSs) by residents of disaster-stricken areas at the time of a disaster are expected to be of great use. They may be a valuable source of information in areas where it is difficult to install, operate, and maintain observation devices or where devices are missing. However, their effective use for damage assessment has not yet been determined. Therefore, a study on the complementary use of SNS data for flood analysis using data assimilation to improve damage assessment is urgently needed. In this paper, we report the evaluation results of data assimilation assuming that SNS data can be collected stably, and we discuss how useful SNS data are for flood damage assessments.

Keywords: flood estimation, state-space model, temporal-spatial analysis, data assimilation

1 INTRODUCTION

There are concerns that the risk of floods will intensify on a global scale. The Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) stated that global warming is gradually progressing, and it is likely that the frequency and intensity of rainfall will change accordingly [1]. There are many areas worldwide where the frequency and intensity of heavy rain and flooding are increasing [2], [3]. Among the measures against flood damage, observing rainfall, rivers, and flooding and understanding the changing situations of rainfall and rivers as well as their influences enable people to determine what actions should be taken and to take effective steps to prevent or mitigate damage. Many previous studies have attempted to estimate flood risks using area vulnerability. For example, in [4], the flood risk in the city was estimated with a detailed spatial resolution of approximately 2 meters. [5], [6] conducted research to estimate index-based flood risk using a theoretical hydraulic engineering model. Furthermore, a Chinese case in [7] examined recognition of risks during the 1997 Red River flood situation. Studies are actively conducted to correctly analyze risks by presenting risks to people in affected areas and raising awareness of individual flood risks, which can lead to mitigation behavior [8], [9].

However, these previous studies are not temporal estimation methods; rather, they are static estimation approaches used to calculate maximum water level. A static estimation result is a risk estimation in which the risk value might change due to rainfall fluctuations. Considering evacuation behavior, dynamic risk estimation is required because flood situations change very rapidly with flooding phenomena over streets due to water overflowing from small rivers and waterways spreading throughout the city in a complicated manner and due to rainwater that cannot be completely drained. Therefore, it is necessary to calculate the high temporal-spatial flood level, which fluctuates according to the rainfall situation, to understand risk with a high temporal resolution for guiding evacuation behavior. Our research goal is to detect flooding as time-series data with only a limited number of observation devices.

This paper investigates whether SNS data can be used to assess flooding. SNS data are effective for determining flood levels even in places where it is difficult to install, operate, and manage observation devices. Although there have been many studies on flood damage detection using SNSs, their effectiveness has not been clarified, and the amount and content of data collected are not fixed depending on the flood damage case. This paper validates the SNS data using the following procedure based on our state-space model (SSM), which was used in our previous research. The purpose of SNS data validation is to investigate whether SNS data can contribute to the accuracy of flood assessment and under what conditions SNS data can improve accuracy.

Then, we also suggest a system for an appropriate SNS use case that could further improve the accuracy of flooding assessment by adding SNS data as well as observation data and calculating flooding conditions for the entire affected area. To realize this system, it is necessary to validate the effectiveness of the SNS data.

We generate quantified SNS data from flood simulations in this paper. To generate the SNS data, we use time-series data collected from observation devices at multiple locations. The results of flood analysis simulation were assimilated using the time-series data to simulate the SNS locations, the timing of postings, and numerical flood levels. Then, we determine the errors in the simulated SNS data. Afterward, we regenerate a flood level at the observation location based on the simulated SNS data with errors and examine the error accuracy against the data assimilation accuracy.

2 RELATED WORKS

2.1 Flood Monitoring

Traditional river sensors [10], [11] have succeeded in detecting disaster signs in large-scale rivers, which have the advantage of stable monitoring and the disadvantage of installation limitation (i.e., very large equipment, high installation cost of several million dollars, and complicated preconfiguration). Improvements in installation limitations enable the possibility of a large number of sensor installations and reliable detection to improve monitoring sensors with higher resolution. Flood prediction, hydrological techniques [12] or artificial neural networks [13], [14] are proposed as high prediction methods. Predicting rising river levels has resulted in highly precise river inflow in the view of large-scale river analyses. However, these previous methods cannot predict the flooding of smaller rivers and waterways. This is because complex water flow prediction requires analyzing complicated relationships among a plurality of confluent rivers and factoring in the impact of rainfall dynamics.

2.2 Risk Estimation with Higher Spatial Resolution

Various studies have already attempted to generate information about places that are dangerous. In case studies, such as [15], [16], their research aimed to present risks on maps. Sinnakaudan et al. [15] developed an ArcView GIS extension as an efficient and interactive spatial decision support tool for flood risk analysis. Their extension is capable of analyzing computed water surface profiles and producing a related flood map for the Pari River in ArcView GIS. In another GIS-based flood risk assessment, Lyu et al. [16] studied the Guangzhou metro system's vulnerability. Their results showed the vulnerability of several metro stations using the flood event that occurred in Guangzhou on May 10, 2016.

Some studies have proposed modeling methods that collect data strictly as input data [4], [17], [18]. Ernst et al. [4] presented a microscale flood risk analysis procedure as a 2-meter grid, relying on detailed 2D inundation modeling and on a high-resolution topographic and land-use database. However, detailed risk estimation requires detailed data measurements, such as laser altimeter data, and it is not realistic to measure these data in all areas.

2.3 Flood Detection through Social Networking Services

Another way to learn about flooded areas is through social networking services (SNSs). Kim et al. [19] stated that social networking is the fourth most popular information source for accessing emergency information. Then they applied social network analysis to convert emergency social network data into knowledge for the 2016 flood in Louisiana. Their objective was to support emergency agencies in developing their social media operation strategies for a disaster mitigation plan. This study explored patterns of interaction between online users and disaster responses.

Sufi et al. [20] designed a disaster monitoring system on social media feeds related to disasters through AI- and NLP-based sentiment analysis. Their system has a mean accuracy of 0.05. They report that their system shows potential disaster locations with an average accuracy of 0.93. Teodorescu [21] designed a method to analyze SNSs for forecasting and relief and mitigation measures. His method analyzes SNS-related time series with the aim of establishing correlations between the disaster characteristics and the SNS response. Although studies using SNS have been applied in many flood damage cases, SNS data are not always posted as expected, and the accuracy may not be achieved as reported in these studies.

2.4 Issues and Approaches

To find safe evacuation routes, it is important to determine the situation regarding the roads in urban areas. Currently, flood damage assessment is based on two methods: numerical simulation (e.g., flood analysis) and monitoring using low-resolution ground observation data (precipitation and river water levels).

Numerical simulations are based on differential equations for flood flow in urban areas for a given amount of precipitation, and the maximum flood level in a detailed given area (e.g., a 10-m grid) is calculated. Based on the calculated results, areas that are anticipated to be hazardous during heavy rainfall are published. However, the analysis uses an ideal model that assumes fixed parameters, such as the amount of precipitation, its runoff coefficient, and the outflow conditions of drainage channels. Therefore, in urban areas with complex rainfall distributions and land uses, the analytical results and actual flood levels will differ. As a result, flooding of roads occurs prior to the announcement of warnings and evacuation information, leading to damage.

On the other hand, monitoring establishes thresholds for dangerous water levels at specific locations where there is concern about road underpasses and river breaches. This method involves situation monitoring to detect the occurrence of flooding based on observation data. This method easily assesses the actual damage but has limited observation points.

SNS data are expected to solve these monitoring limitations. As indicated in the previous section, the importance of SNSs in flood damage detection has long been known and has been applied in many flood damage cases. However, there is a fundamental problem with water damage detection using SNS. That is, SNS data are not necessarily posted in every case. While it may work effectively in floods with a high number of postings, it is highly likely that it will not be as accurate as reported in floods with a low number of postings. In particular, it may be difficult to post while ensuring safety in heavily damaged areas, and communication problems may prevent posting. We are convinced that these problems are obstacles to the effective use of SNSs for flood damage detection. Therefore, in this paper, we investigate how much SNS data regarding the number of postings, their contents, and the timing of postings would be effective for damage assessment.

We intend to develop a system for improving the accuracy of flood level estimation through data assimilation using heterogeneous data for investigation in this paper. We have previously proposed a method for estimating the expansion process of flooding by applying data assimilation using heterogeneous observation time-series data to simulate flood analyses. Our estimation method showed a significant improvement, with an error of less than 9 cm. We are planning to add SNS data to this estimation method to further improve its accuracy and to develop a system to present the flood disaster situation in the entire affected area. However, although there have been many studies on flood damage detection using SNSs, the amount and content of the data collected are not well defined for each flood damage case. The effectiveness of using SNS for flood level estimation has not yet been clarified. Therefore, with the aim of developing a system to accurately estimate and present flooding situations for entire affected areas, this paper investigates the relationship between SNS data and the accuracy of flooding estimation through simulations.

3 PROPOSAL OF A DATA ASSIMILATION METHOD FOR IMPROVING THE ACCURACY OF FLOOD ESTIMATION

3.1 Overview

This research proposes a system for estimating floods through data assimilation to identify safe evacuation routes and timing in the event of an urban flood. We propose a flood estimation system based on data assimilation. This system aims to improve the accuracy of flood estimation by assimilating data using observation time-series data and SNS data. We intend to further improve accuracy by encouraging system users to submit messages that compensate for the lack of observation locations. The system will leverage the advantages of both observations, which can provide accurate time-series data, and SNS data, which can easily provide a large amount of data based on the number of posted messages.

The procedure in our system consists of three phases(Fig. 1). In Phase 1: SNS Data Validation, where we must validate the effectiveness of SNS data for improving the accuracy of flooding assessments. Although some previous studies have shown that SNS data are effective for disaster damage assessment, they are empirical validations based on data collected during that particular disaster, and we cannot eliminate the possibility that the highly effective data were simply collected coincidentally. There have been few verifications in terms of what kind of SNS data are effective or ineffective among the SNS data posted during disasters. Consequently, we need a scientific validation of the relationship between SNS data and analysis.

Phase 2: Flood Assessment Promotion Requirement Validation, which is a more detailed survey compared with the validation conducted in Phase 1. We investigate the flooding conditions of various past flood events and the social networking data posted at the time of the event. Simultaneously, we simulate SNS data based on the relationship between the Past SNS data and the physical measurement of events from learning process. Furthermore, SNS Generator simulates numerical data of flooding under the scenario of disaster occurrence and generates SNS data based on the measured data. Through phase 2, the effectiveness of SNS data in flood assessment will be validated for as many flood disasters as possible. Moreover, the system creates dissemination requirements to determine what kind of SNS data would improve the accuracy of the assessment.

Phase 3: Flood assessment, as shown on the right side of Fig. 1. In Phase 2, we will store the SNS data knowledge that

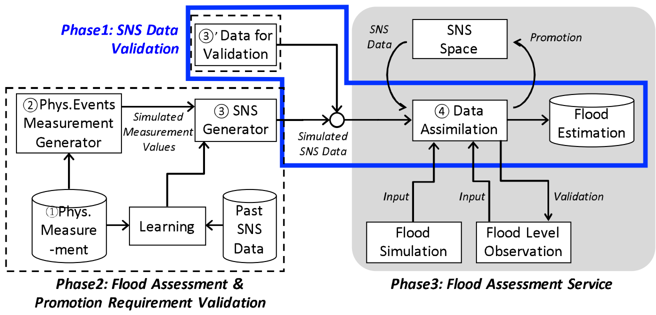


Figure 1: Proposed System for Flood Assessment Approach

is useful for flood assessment. Furthermore, our system will use this knowledge to disseminate the desired data to the SNS space. Phase 3 involves performing a highly accurate flood assessment using the collected data. This assessment would involve an estimation of flood level using data assimilation. In data assimilation, the estimated water level analyzed by the flood simulation is corrected with observed data and SNS data using a state-space model. The results of data assimilation are compared with the observed data and SNS data. As a result of the comparison, the SNS data are obtained through commercial services and other means in areas where flooding has been detected with low accuracy. The results can also be used to improve the data assimilation process. We plan to develop a dissemination function that encourages SNS space to post SNS data in areas with insufficient data. By providing SNS data from SNS space, the data assimilation results can be updated to improve accuracy.

3.2 Flood Assessment Approach

Our proposed simulation is based on the use of observations and SNS data. Previous studies [22] have shown that observation data contribute significantly to the accuracy of flood water level correction, but there is an upper limit to the number of observation devices because they need to be installed. Therefore, the effective use of SNS data when devising a system is considered to be critical to improving accuracy. As mentioned in 2.4, there are many studies on flood damage detection using SNS data, most of which are analyses of flood damage, where a large amount of data is available. However, whether or not we can collect a large amount of SNS data depends on the characteristics of the flooded area. Thus, it is not always possible to collect SNS data that can be used for flood analysis in all flood disasters. In addition, SNS data collected at the time of actual flooding also vary in the expression of information and the timing of postings. Furthermore, the accuracy of many of the data cannot be verified. Therefore, to sustainably utilize SNS data for disaster management, it is necessary to verify its effectiveness in various cases.

We investigate the conditions under which SNS can contribute to flood estimation and how SNS data can improve the accuracy of flood analysis. Phase 2 in Fig. 1 shows our expected use of SNSs. The SNS data collected from past floods are compared with actual flood damage (①physical measurement), and the results are used by a ②physical events measurement generator to simulate the damage that occurred at the time of the flood as simulated measurement values. Based on simulated measurement values, ③an SNS generator generates simulated SNS data. Using simulated SNS data collected in this phase, we perform ④data assimilation that combines flood analysis simulations and observed data, and calculate under what conditions the SNS data should be collected to improve flood estimation accuracy. The process of generating simulated SNS data here involves data analysis in a cyber-physical system. By feeding the analytical results back to the data assimilation process, that is, the physical space, we expect to improve the accuracy of data assimilation and disseminate contributions by the SNS Promotor. The purpose of this research is to verify the accuracy of flood estimation using simulated SNS data and to determine the accuracy, precision, and timing of postings on SNSs that can be used in times of disaster.

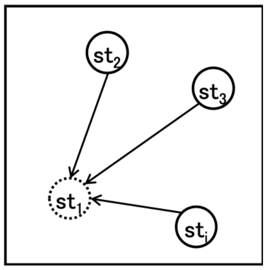
3.3 Main Objective of This Paper

Of the three phases required to develop our system, this paper focuses on Phase 1: SNS Data Validation. This phase is shown in the blue box in Fig. 1. The objective is to investigate whether SNS data containing various errors (sometimes the errors are considered to have a significant impact) are effective in improving the accuracy of data assimilation and flood estimation. The ③′validation data are simulated as expected data that would be posted based on the characteristics of the SNS data. We use Twitter as the SNS data source. These SNS data contain various errors in location, timing of postings, and flood water levels. However, these data are generated by a simple simulation and are not fully replicated in the SNS data.

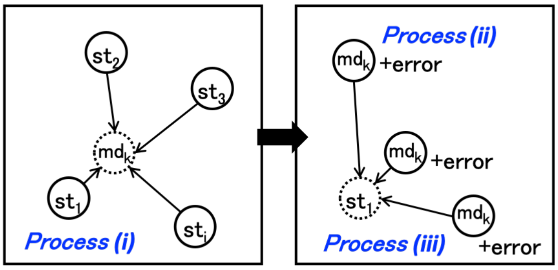
When using SNS data for analysis, we need a process to convert text- and image-based posts into numerical values. The quantification of SNS data does not ensure an accurate calculation of flood levels, since the data representation of text-based SNSs is ambiguous; thus, it affects the quantification of flood levels. SNS data reporting water levels during flooding could include measurement of water levels (e.g., 30 cm) or be expressed in comparison to a body (e.g., up to his or her knee). Not only water levels but also location and timing contain errors in quantification, depending on the type of data representation. We investigate the SNS data impact considering this error on statistical flood analysis, such as data assimilation.

3.4 SNS Data Validation for Data Assimilation

The basic idea of Phase 1 is described as follows: SNS Data Validation for Data Assimilation in this paper is illustrated in Fig. 2(a), state-space model (SSM) using SNS Data. The purpose of SNS data validation is to investigate whether SNS data can contribute to the precise estimation of flood assessment and under what conditions SNS data can improve accuracy. Since we are unable to determine the error rate contained in the SNS data collected at the time of flooding, we generate alternative quantified SNS data from simulations. To generate the SNS data, we use time-series data collected from observation devices at multiple locations. The results of flood



(a) SSM: Observation Data



(b) SSM: SNS Data

Figure 2: SNS Data Validation for Data Assimilation ($st_{1,...,i}$:observation location, md_k:SNS data posting location)

analysis simulation are assimilated with the time-series data to simulate the SNS locations, the timing of postings, and numerical flood levels (Fig. 2(b)(i)). For this data assimilation, we used the spatial-temporal SSM proposed in our previous study [22](Fig. 2(a) SSM using Observation Data). Note that, a spatial-temporal state-space model in this paper, is applied without using the waterway and sewer data used in the previous study [22], because these data are generally limited in availability.

Then, we determined the errors in the simulated SNS data (Fig. 2(b)(ii)). This process assumes the errors in location, time, and water level value that are present in the textual data of the actual SNS data. Afterward, we regenerate a flood level on the observation location based on the simulated SNS data with errors and examine the accuracy of the errors on the data assimilation accuracy (Fig. 2(b)(iii)). This regeneration uses a state-space model that applies the state-space model of the previous study [22] toward the spatial direction. Here, if there is a small difference between the time-series data and the data assimilation results at the observation location, the SNS data are applicable to flood assessment by data assimilation. In contrast, if the data assimilation accuracy is low despite the small error appended to the simulated SNS data, then there are problems using the SNS data.

3.4.1 Process(i) Simulated SNS data for Flood Analysis Simulation

The process flow is shown in Fig. 3. In process(i), the results of the flood analysis simulation are assimilated with time-series data collected from observation locations to simulate the locations, timing of posting, and flood water level values. This section outlines the spatial-temporal state-space model used in our data assimilation. The basic flood analysis is

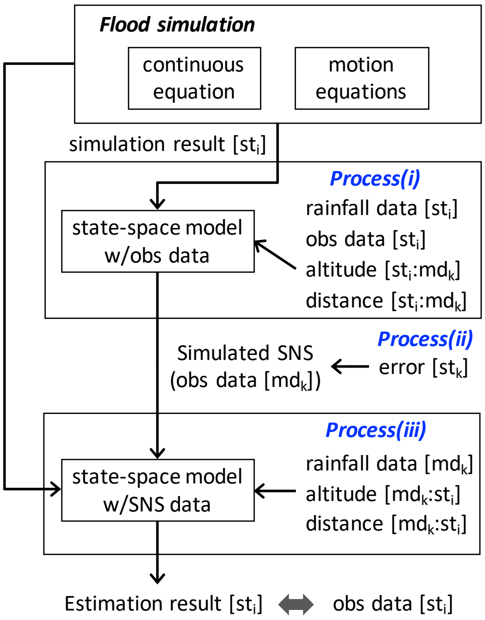


Figure 3: SNS Data Process Flow ($st_{1,...,i}$:observation location, md_k:SNS data posting location)

based on a conventional simulation that uses a surface flooding model. This method calculates the amount of runoff at each grid location by expressing the flooding flow as a continuous equation and motion equations.

A continuous equation is defined as follows.

$$\frac{\partial h}{\partial t} + \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} = 0 \quad (1)$$

The motion equations are given as follows:

$$\frac{\partial M}{\partial t} + \frac{\partial UM}{\partial x} + \frac{\partial VM}{\partial y} + gh\frac{\partial H}{\partial x} + \frac{1}{\rho}\tau_x(b) = 0 \quad (2)$$

$$\frac{\partial N}{\partial t} + \frac{\partial UN}{\partial x} + \frac{\partial VN}{\partial y} + gh\frac{\partial H}{\partial y} + \frac{1}{\rho}\tau_y(b) = 0 \quad (3)$$

Each parameter is defined as t: time, H: water level, h: flood level, U: flow velocity (X direction), V: flow velocity (Y direction), g: gravity acceleration, ρ: water density, M: flux (X direction), and N: flux (Y direction) ($M = uh$, $N = vh$).

Here, the shear force in the x direction $\tau_x(b)$ and the shear force in the y direction $\tau_y(b)$ are defined as follows.

$$\tau_x(b) = \frac{\rho g n^2 \overline{U}\sqrt{U^2+V^2}}{h^{\frac{1}{3}}} \quad (4)$$

$$\tau_y(b) = \frac{\rho g n^2 \overline{V}\sqrt{U^2+V^2}}{h^{\frac{1}{3}}} \quad (5)$$

The roughness coefficient n (the resistance value of river water to touch obstacles) can be expressed as follows, considering the influence of a building.

$$n^2 = n_0^2 + 0.020 \times \frac{\theta}{100-\theta} \times h^{\frac{4}{3}} \quad (6)$$

(n:bottom roughness coefficient, n_o:composition equivalent roughness coefficient,and θ:building occupancy rate)

Equation (1)-(3) calculates flood level h for each grid, accounting for the runoff from the inside of the sewer line to the ground surface and the flooding to the ground surface due to rainwater. For the above equations, the inflow into each grid represents the flux into each grid from adjacent grids and the effect of buildings on the inflow in each grid.

We define D as the two-dimensional space corresponding to the region of interest and divide D into m grids of d meters each. Let $s_i \in D$ denote the location coordinates of each grid (s_i is denoted by i). Using equation (1)-(3), we calculate $h_t(i)$ for the flood level of each grid at time t.

Then, using the state-space model, we estimate the flood level of grid s_k from the observations $y_t^{(i)}$ collected at the observation location at time t. Grid $s_k(k = 1, 2, 3...., m)$ is the location indicated by the SNS data. The state-space model is represented by two types of observation equations; the flood analysis simulation result $h_t(i)$ at grid s_i, and the difference between the flood analysis simulation and the observed value at the observation location. This state-space model is defined by the equations (7)(8)(9).

$$\boldsymbol{y}_t^{(i)} = S_t \boldsymbol{r}_t^{(i)} + \boldsymbol{G}_t^{(i)} \boldsymbol{x}_t^{(i)} + \boldsymbol{e}_t^{(i)} \quad (7)$$

$$\boldsymbol{r}_t^{(i)} = \boldsymbol{r}_{t-1}^{(i)} + \boldsymbol{v}_t^{(i)} \quad (8)$$

$$\boldsymbol{x}_t^{(i)} = \boldsymbol{x}_{t-1}^{(i)} + \boldsymbol{u}_t^{(i)} \quad (9)$$

The $\boldsymbol{r}_t^{(i)}$ denotes the state at time t and $\boldsymbol{v}_t^{(i)}$ denotes noise. The term $\boldsymbol{G}_t^{(i)}\boldsymbol{x}_t^{(i)}$ represents the total inflow/outflow, and $\boldsymbol{x}_t^{(i)}$ is the difference between the flood analysis simulation results and the observed values. The $\boldsymbol{u}_t^{(i)}$ denotes the noise at time t. $\boldsymbol{G}_t^{(i)}$ is the adjacency matrix indicating the spatial component.

3.4.2 Process(ii) Appending Errors to Simulated SNS Data

Now, in process(ii), we append the error component to the simulated SNS data. Actual SNS data show a variety of representations of water levels. For example, "It is flooded up to my knees," "The car is flooded," or pictures are posted with comments such as "It is raining so hard." This paper considers SNS data that express water levels in words or show flood condition. The flood level $h_t(k)$ indicated by the SNS data is assumed to contain an error component e. The represen-tation type of the SNS data is considered to be a factor that causes errors due to the quantification of the SNS data (Table 1). We assume that the type of data representation occurs for each of the posted location, time, and water level values. For water level values, SNS data can be expressed in the form of measurement, comparison with an object, or description of the situation. In the case of measurement, it is considered to be measured by a guess, which results in a difference from the actual water level.

When compared with objects, water levels are explained based on objects such as knee height or up to the ankles, but the sizes of these objects vary from each user, so even if quantified, they differ from the actual water level. Although this is only an assumption, the average length below the knee for Japanese people is 46.7 cm for males and 42.9 cm for females, a difference of approximately 4 cm even in the average value. In the describing the situation, the data mostly describe the flooding aspect, with little mention of water levels; consequently, it is expected that quantification itself is often difficult. In cases of pictures showing flood conditions, errors could be contained during the estimation process of converting the images to numerical data.

For information representation of location, the following information can be considered: GPS, address, road/river, landmark, and city/town name. When GPS data are attached to SNS data, the exact location at the time of posting can be reflected in the quantified SNS data. However, if the location indicated by the posted message differs from the location at the time of posting, there is an error compared with the posted location. The same error occurs for other types of data representations. In some cases, addresses of flooded areas are posted for rescue in flooding situations. Although there may be an error of a few meters, the location information would be approximately correct. If a road/river is described as a location, it is considered difficult to determine the exact location from the text content itself. In the case of landmarks, the data may indicate the location in front of the landmark, whereas it is also possible that the data indicate the location where the landmark is visible, in which case a large error of approximately 100 meters or more would occur. For city/town names, we consider a significant error of several kilometers when identifying the location due to the wide range of areas indicated by the data.

For the time information, the following three forms are considered: timestamp, comparison, and date/time range. The timestamp shows the exact date and time in the simulated SNS data. For comparison, it is considered to be a popular form of time; nevertheless, representations, such as "just now" are likely to include an error of several tens of minutes, as the sense of time differs among individuals. Additionally, it is assumed that there are many cases describing a range of dates/times, such as "approximately 21:00" or "this evening". While errors are expected to be small for numerical time representation, in the case of "night" and other representations, errors are likely to be on the order of several hours. Furthermore, as with location, there are cases in which the time of posting also differs from the time of flooding. This difference may result in a significant error in the time representation.

We define the error components at location s_k, resulting from the quantification, as the error in the representation type $e_{\Delta v,\Delta l,\Delta t}(k)$, the error relative to the posting location/time $\zeta_{\Delta l'}(k)$, and the error from the time of posting $\zeta_{\Delta t'}(k)$.

3.4.3 Process(iii) Flood Level Estimation and its Validation

Process(iii) regenerates the time-series data for the observation location to investigate the error impact on the data assimilation accuracy based on the simulated SNS data with error $h_t(k) + e_{\Delta v,\Delta l,\Delta t}(k) + \zeta_{\Delta l',\Delta t'}(k)$. Here, we employee a

Table 1: SNS Data Representation Types

Data	Types	Example
Level	measurement	30 cm
	comparison	knee height
	description	looks like river
Location	GPS	34.9104, 135.8002
	address	1-1 Gokasho,Uji-city,
	road/river	Route 24
	landmark	In fromt of Kyoto Station
	city/town name	Uji city
Time	timestamp	2022/7/8/21:00
	comparison	just now
	range	approximately 21:00

state-space model that utilizes the model detailed in 3.4.1 in the spatial direction. For equation (7)(8)(9), at time t, the simulated SNS data $h_t(k) + e_{\Delta v,\Delta l,\Delta t}(k) + \zeta_{\Delta l',\Delta t'}(k)$ is substituted into $y_t^{(i)}$ to estimate the flood level $h'(k')$ of the target location $s_{k'}$. Simulated SNS data on a particular location is not continuous time-series data. Thus, there is only one t in the simulated SNS data, and the state-space model in process(iii) is applied only applied spatially.

The difference between the restored water level $h'(k')$ and the actual water level $h(k')$ is shown as the effect of the error component $e_{\Delta v,\Delta l,\Delta t}(k) + \zeta_{\Delta l',\Delta t'}(k)$ on the data assimilation method. This paper validates the SNS data effectiveness by comparing flood level $h'(i)$ regenerated from the simulated SNS data at location k with the actual observed water level $h(i)$.

3.5 Evaluation and Discussion

3.5.1 Evaluation Purpose and Flood Case

This evaluation investigates the data representation effect that text-based SNSs have on flood level quantification. SNS data reporting water levels during floods often include the measurement of water levels or are expressed in comparison to a body of water. In addition to water levels, location and time also contain errors in quantification, depending on the type of representation. We investigate the SNS data impact with these errors on statistical flood analysis, including data assimilation. Based on the results, we will discuss what SNS data form would contribute to flood assessments.

Our evaluation is based on SNS data simulated by a state-space model with observed flood data. We append errors to the simulated SNS data and observe the effects of the errors. Then, we determine errors that could occur due to the quantification of the SNS data, as shown in Fig. 4. This paper uses flood observations collected in Tsushima city, Aichi Prefecture, Japan, from October 22 to 23, 2017, due to rainfall caused by Typhoon 23. A rainfall amount of 32 mm/h was observed at 23:00 at the precipitation gauge nearest to our target area (Aisai Observatory, Aichi Prefecture). Using these rainfall data as input values, we calculated a flood analysis simulation with the flood analysis simulation NILIM 2.0. With the results of the simulated flood analysis, we apply the flood estimation method using the state-space model. To generate

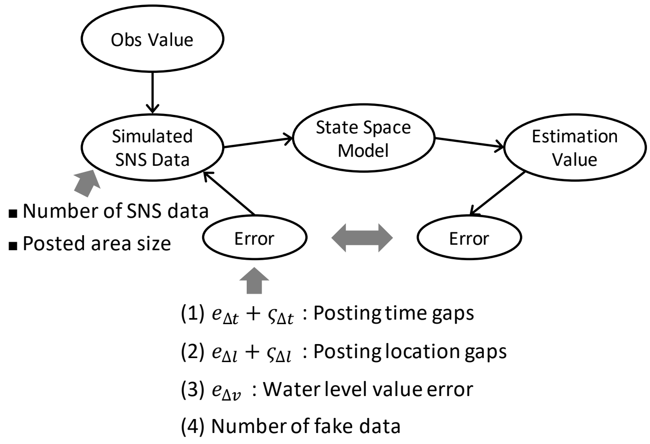


Figure 4: Appending Errors to Simulated SNS Data

simulated SNS data, the state-spatial model with observation data uses water level observation data at waterways collected every 5 minutes from pressure-type sensors installed at four locations in the target area.

Detailed information about the observation locations is presented as follows. Observation locations 1 and 2, and 3 and 4 are on the same waterway. The distance between observation locations 1 and 2 is approximately 500 meters, and the distance between locations 3 and 4 is approximately 600 meters. Observation location 4 is connected to the sewer. Two waterways are approximately 500 meters apart. There are no floodgates between the observation locations. The difference in elevation in this area is within 0.30 meters, and the elevation values (elevation model by the Geospatial Information Authority of Japan) are equal at all four observation locations. The heights from the bottom of the waterway to the road are 1.01, 1.14, 0.72, and 1.28 meters, and the usual water levels are 0.06, 0.07, 0.16, and 0.31 meters. On the day of the flood, the installed sensor devices showed, water overflowing the waterways and flood levels of up to 0.26, 0.25, 0.63, and 0.48 meters above the road.

3.5.2 Evaluation Procedure

The evaluation randomly extracts simulated SNS data according to the number of SNS data from the posted area size in Table 2. Errors appending to the simulated SNS data are parameterized to indicate fluctuations. The parameters that indicate the fluctuation of each error are shown in Table 2. We begin by using the number of SNS data posted and the area size where SNS data were posted as two common parameters. Subsequently, we adopt four different parameters to generate the errors as in (1)-(4). We repeat the above procedure five times and compute the mean value of the estimation. A total number of 489,637 simulated SNSs appended with the following errors are applied to the state-space model to calculate the estimated water levels for the four observation sites with time-series data.

Evaluation (1) There are two types of time information errors originating from the representation of time information: fluctuations in the estimation of time information from posted

Table 2: Parameters related to Error Fluctuations

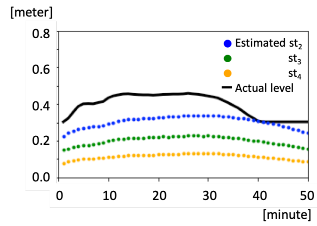
Parameters	Fluctuations
Number of SNS data	8,24,48,80,120,435
Posted area size (Radius)	10,20,30,40,50, 100[meters]
(1)Posting time lags	10,20,30,40[minites]
(2)Posting location (Radius)	10,20,30,40,50, 100[meters]
(3)Water level value error	0,10,20,30[cm]
(4)Number of fake data (Ratio)	10,20,30,40,50, 60,70,80,90[%]

SNS data, and fluctuations due to the gap between the posted time and the flooding conditions indicated by the posting messages. We append these two fluctuations of time information, $e_{\Delta t}(k)$ and $\zeta_{\Delta t'}(k)$, to the simulated SNS data as posting time lags behind flood conditions. The value of the parameter: posting time lags in Table 2 adds a delay to the time of the simulated SNS data.

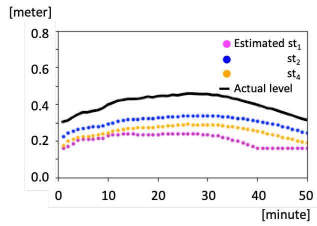
Evaluation (2) Location errors originating from the representation of location information can be considered as fluctuations in location information when location information is estimated from posted SNS data, and fluctuations due to the gap between the location indicated by the posted messages and posting location. These two location fluctuations $e_{\Delta l}(k)$ and $\zeta_{\Delta l'}(k)$ are appended to the simulated SNS data as posting location gaps. We randomly swap the location information of the simulated SNS data within the radius area indicated by the parameter: posting location gaps in Table 2.

Evaluation (3) The error, resulting from the information representation of the water level value, could be the fluctuation of the value when the water level is estimated from the posted SNS data. Although this fluctuation can have a range of different values, this paper assumes a fixed parameter for the error in order to verify the effect of the SNS data. We add the value indicated by the parameter: water level value error in Table 2 to the water level in the simulated SNS data.

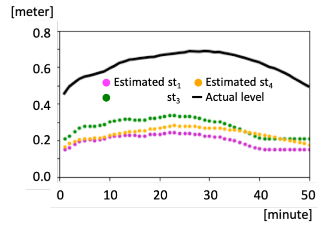
Evaluation (4) Furthermore, SNS data can be considered to include the posting of fake data. To evaluate the effect of the fake data on the assessment of flooding, we substitute the simulated SNS data with the fake data. We prepare two types of fake data; fake data that posts under flooding conditions ”no flooding is occurring (water level: 0 m)”, and fake data that post regardless of the current water level ”flooding is occurring to the extent of the first floor of a building (water level: 1.5 m)”. We parameterize the ratio of fake data among the simulated SNS data to calculate the estimated water level. Then we substitute 0 or 1.5 meters of simulated SNS data at the rate indicated by the parameter number of fake data in Table 2.



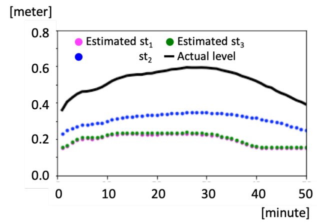
(a) st_1 Estimation Value



(b) st_2: Estimation Value



(c) st_3: Estimation Value



(d) st_4: Estimation Value

Figure 5: Preliminary result: SSM with Observation Data

Table 3: RMSE: SSM with Observation Data

Location	mean	minimum	maximun
st_1	0.19	0.12	0.24
st_2	0.20	0.15	0.23
st_3	0.38	0.28	0.43
st_4	0.28	0.18	0.33

3.5.3 Results

Preliminary result: SSM with Observation Data To compare accuracies, this section shows the estimated flood level $l^{(i)}_{t_T,k}$ using the state-space model with observation data (Fig. 2 SSM using observation data). Of the four observation locations, we apply the observation data from three locations (from time $t = 0$ to $t = 50$) to "SSM using Observation Data" to estimate the flood water level at the remaining one location (the water level at st_1 is estimated using the water level time-series data at st_2, st_3, and st_4). The estimation results are shown in Fig. 5. The black lines indicate the actual observed flood water levels, and the magenta, green, blue, and orange dots indicate the estimated water levels using the state-space model. Root Mean Squared Error: RMSE between the mean of the estimation results and the actual observed values is shown in Table 3.

In Fig. 5 (a), the estimated water level at st_1 was the most accurate using the data from st_3, with an average RMSE of 0.19 meters. The estimated water level for st_2 in Fig. 5 (b) was nearly the most accurate, with an average RMSE of 0.20 meters, using data from st_3. Both the estimation for st_3 in Fig. 5 (c) and st_4 in and Fig. 5 (d) using data from other observation locations were less accurate, with average RMSEs of 0.38 meters and 0.28 meters, respectively. The results of either estimation resulted in a large difference from the water level at the flood peak. These results are due to the failure of the "SSM using Observation Data" to follow the rising and falling water levels. In our previous study [22], the maximum

Table 4: RMSE: Result(1)：Posting Time Lags

Time Lags [minutes]		10	20	30	40
st_1	mean	0.09	0.10	0.12	0.14
	minimun	0.00	0.00	0.00	0.06
	maximun	0.42	0.45	0.29	0.24
st_2	mean	0.06	0.06	0.07	0.09
	minimun	0.00	0.00	0.00	0.00
	maximun	0.51	0.42	0.26	0.23
st_3	mean	0.16	0.16	0.16	0.12
	minimun	0.00	0.00	0.00	0.00
	maximun	0.31	0.31	0.30	0.24
st_4	mean	0.09	0.09	0.08	0.05
	minimun	0.00	0.00	0.00	0.00
	maximun	0.52	0.46	0.22	0.31

error was 9 cm because we included data of waterways and sewers. However, this paper does not use those data, in order to apply our system in areas where it is difficult to collect the data. For all observation locations, we found that estimation using data from distant observation locations resulted in large errors.

Result (1)：Posting Time Lags This section describes the results of the estimation when time lags are appended to the simulated SNS data. The values of the time lags are generated as delays of 10, 20, 30, and 40 minutes. The RMSEs between estimated flood levels and actual observations applied to the state-space model are shown in Table 4. For st_1 and st_2, the larger the time delay is, the larger the mean value of the RMSE is. However, for st_3 and st_4, the larger the time lags are, the smaller the RMSE is. When the time delay was 10 minutes, the estimated water level difference was 0.09 meters, an improvement by 10 cm from the estimated value in 3.5.3. The minimum value of RMSE was 0.00 meters for all st_i, equal to the observed value. On the other hand, the maximum value of RMSE was 0.52 meters, resulting in a large error.

We found that the large time lag did not result in a very significant effect on the estimated water level. One reason for this may be that the actual flood levels were not large, with a maximum of 0.6 meters, and there was no sudden water level rise at any of the observation locations. However, it was found that the time delay in the SNS data was allowable for a flood of this scale.

Result (2)：Posting Location Gaps As in the Posting Location Gaps we rewrote the location information of the simulated SNS data within a specific area. We randomly changed the values of the appended location gaps to different location information from the true location of the area with a radius of 10, 20, 30, 40, 50, and 100 meters. Table 5 shows the RMSEs of the flood estimation results using simulated SNS data with location gaps. The minimum value was 0.00 meters, equal to the actual observed value in all observation locations. For st_1 and st_2, the RMSE increases with the location gap. The mean RMSE for st_3 and st_4 is approximately the same regardless of the location gap, whereas the maximum RMSE

Table 5: RMSE: Result(2)：Posting Location Gaps

	Location Gaps [meter]	10	20	30	40	50	100
st_1	mean	0.08	0.08	0.08	0.08	0.09	0.09
	minimun	0.00	0.00	0.00	0.00	0.00	0.00
	maximun	0.23	0.25	0.24	0.25	0.26	0.29
st_2	mean	0.06	0.06	0.06	0.06	0.06	0.07
	minimun	0.00	0.00	0.00	0.00	0.00	0.00
	maximun	0.19	0.19	0.20	0.19	0.19	0.33
st_3	mean	0.15	0.15	0.14	0.14	0.14	0.16
	minimun	0.00	0.00	0.00	0.00	0.00	0.00
	maximun	0.29	0.30	0.29	0.29	0.29	0.30
st_4	mean	0.09	0.09	0.08	0.08	0.08	0.08
	minimun	0.00	0.00	0.00	0.00	0.00	0.00
	maximun	0.22	0.22	0.34	0.30	0.37	0.58

results in a larger RMSE as the location gap increases. Comparing only the mean values, the RMSE values are almost the same for any st_i. This finding can be explained by the fact that our location gap area radius is at most 100 meters, which is a small area. Thus, the estimation results indicate the possibility of estimating the flooding situation for flooding of this scale, even if there is a gap in the posting location within these area sizes.

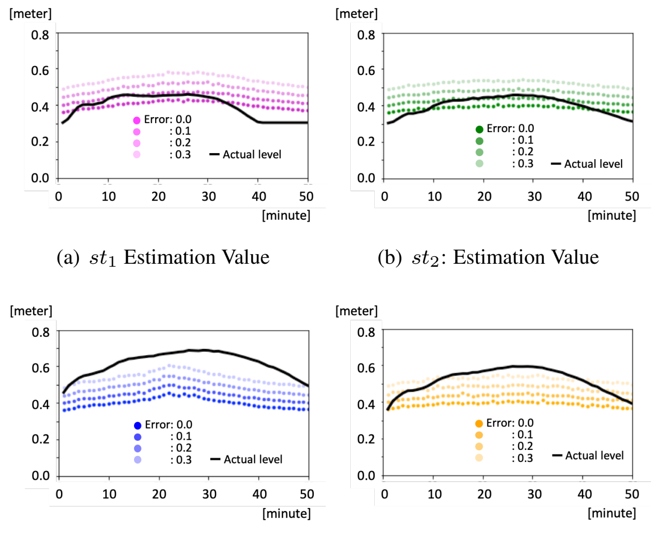
Result (3)：Water Level Value Error This section shows the results of applying the state-space model with error values appended to the simulated SNS data as Water Level Value Error in Table 6. When no error values were added, the mean values of st_1 and st_2 showed the smallest RMSE. The larger error value resulted in a larger RMSE. On the other hand, st_3 and st_4 showed a large RMSE for mean value.

Figure 6 compares the time-series data of the actual observations with the mean value of the estimated values. For st_1 and st_2, as the water level changes, the water level estimated from the SNS data also changes and shows little difference from the actual observation. In addition, larger error values tend to provide larger estimates of the results. Moreover, ST_3 and ST_4 result in a difference by approximately 0.20 meters between the estimated result (error value: 0) and the actual water level at the peak of the flooding. We consider that the estimated values were closer to the actual observed value when the error value was increased since there was such a large difference at the error value of 0. One possible reason for this is that the values of st_3 and st_4 were estimated to be lower due to the use of data from other observation locations when generating the simulated SNS data.

Table 6 shows that the minimum value is almost 0.00 meters, even when large values are added as errors. However, the maximum value results in an extremely large value. In st_1 and st_2, the mean values of RMSE do not change significantly after adding the error value, contrary to this the error value: 0.30 meters in st_4 shows an RMSE of 0.52 meters, which is a large error. Accordingly, this would require a data collection method and analysis process that reduces the error in values, and a process to validate the reliability when large water levels are estimated would be required.

Table 6: RMSE: Result(3)：Water Level Value Error

	Value Error [meter]	0.00	0.10	0.20	0.30
st_1	mean	0.04	0.05	0.10	0.15
	minimun	0.00	0.00	0.01	0.06
	maximun	0.13	0.17	0.27	0.30
st_2	mean	0.04	0.03	0.07	0.12
	minimun	0.00	0.00	0.00	0.03
	maximun	0.26	0.15	0.25	0.39
st_3	mean	0.21	0.17	0.13	0.08
	minimun	0.00	0.03	0.00	0.00
	maximun	0.30	0.28	0.25	0.22
st_4	mean	0.13	0.09	0.06	0.05
	minimun	0.00	0.00	0.00	0.00
	maximun	0.22	0.18	0.37	0.52



(a) st_1 Estimation Value (b) st_2: Estimation Value

(c) st_3: Estimation Value (d) st_4: Estimation Value

Figure 6: Result(3)：Water Level Value Error (Mean Value)

Result (4)：Number of Fake Data We describe the estimation results of substituting simulated SNS data for fake data. The fake data were either 0 or 1.5 meters values at 10, 20, 30, 40, 50, 60, 70, 80, and 90% of the simulated SNS data used in the state-space model. This evaluation did not add error values to the water level values.

Table 7 shows an abstract of the RMSE results. Compared to Table 6, the RMSE was larger and less accurate for the results with the addition of the fake data. The RMSE becomes larger as the ratio of fake data increases. When the ratio of fake data 0 is 90%, st_3 has an error of 0.69 meters. When the fake data were set to 1.5 meters, a larger error resulted, with a maximum error of 1.20 meters. In st_2, with 0 fake data, the maximum error was 1.5 meters. The process of averaging offsets this error value, as the amount of fake data is as small as 10%, resulting in a mean value of 0.15 meters.

Table 7 indicates that if the ratio of fake data is as small as 10 or 20%, the error would not be significant. Moreover, st_3, where the ratio of fake data is even 20%, shows a large error in the mean RMSE, This arises from low estimation accuracy even without error value st_3. This evaluation did not add error

Table 7: RMSE: Result(4)：Number of Fake Data

Ratio [%]			10	30	50	70	90
st_1	0.0	mean	0.07	0.13	0.21	0.29	0.36
		minimun	0.00	0.00	0.00	0.00	0.00
		maximun	0.46	0.46	0.46	0.46	0.46
	1.5	mean	0.15	0.36	0.57	0.81	1.02
		minimun	0.00	0.00	0.00	0.00	0.00
		maximun	1.20	1.20	1.20	1.20	1.20
st_2	0.0	mean	0.07	0.14	0.21	0.30	0.37
		minimun	0.00	0.00	0.00	0.00	0.00
		maximun	1.50	0.46	0.46	0.46	0.46
	1.5	mean	0.14	0.34	0.56	0.80	1.01
		minimun	0.00	0.00	0.00	0.00	0.00
		maximun	1.50	1.20	1.20	1.20	1.20
st_3	0.0	mean	0.26	0.34	0.42	0.51	0.58
		minimun	0.01	0.01	0.01	0.07	0.07
		maximun	0.69	0.69	0.69	0.69	0.69
	1.5	mean	0.22	0.20	0.44	0.63	0.81
		minimun	0.00	0.00	0.00	0.00	0.00
		maximun	1.01	1.04	1.04	1.04	1.04
st_4	0.0	mean	0.17	0.25	0.33	0.41	0.49
		minimun	0.00	0.00	0.00	0.00	0.00
		maximun	0.59	0.59	0.59	0.59	0.59
	1.5	mean	0.18	0.32	0.50	0.71	0.90
		minimun	0.00	0.00	0.00	0.00	0.00
		maximun	1.14	1.14	1.14	1.14	1.27

values to the simulated SNS data other than the fake data. We also consider that adding error value, time lags, and location gaps to the SNS data will further increase the error.

3.5.4 Discussion

The evaluation applied simulated SNS data with error data appended to the data as fluctuations in four aspects (1)-(4), to compare the estimated results to actual water levels. (1) time lag and (2) location gap showed that a narrow range of fewer than 100 meters does not significantly affect the RMSEs. For (3) water level error and (4) fake data, we found that as long as the error is small and the ratio of fake data is small, the estimation accuracy does not reduce significantly. All evaluations showed significant improvements compared to the state-space model with observations. This finding shows that the observation locations are approximately 500 meters away from the estimated locations, whereas the simulated SNS data are within 100 meters, allowing for more accurate estimation. Unlike time-series data, SNS data are sparse in the time direction, although it is effective for estimation if collected at locations that are near the estimated location.

In this evaluation, we also treated the number of SNS data and posted area size as parameters. These two parameters did not significantly affect the estimation results. Hence, we found that even if the number of SNS data is small, locations near the estimation location can be estimated with sufficient accuracy. For st_1 and st_2, the smaller the parameters (1)-(4) are, the smaller the error is. These results indicate that time lag and location gap are allowable within the range of values of this evaluation and that a small number of errors in water level values and fake data prevents a large error. For st_3 and st_4, the errors are large even without adding the water level error, requiring investigation as to the cause. Since this paper covers only four observation locations, we assume that this is due to the effect of low water levels observed at the other locations. The process in 3.4.1 calculated a low-accuracy estimation of the observed location. As a result, the simulated SNS data based on the estimation are also to be lower than the actual water level. Therefore, we need to improve the state-space model itself. One idea is to apply observations and SNS data together to the state-space model and implement a Kalman filter for locations where time-series data are available.

The purpose of our research is to investigate the conditions under which SNSs can contribute to flood estimation: how SNS data, if available, can improve the accuracy of flood analysis. To achieve the system in Fig. 1 Phase 1 validated the effectiveness of SNS data in improving the accuracy of flooding assessments as SNS data validation. The results of the assessment using simulated SNS data showed that a certain degree of error was allowable and that the accuracy was better than the estimation using observation data collected over a longer distance. We conclude that further processing improvements are needed, such as removing larger errors by estimation using a combination of observed and SNS data.

4 CONCLUSION

This study investigated whether SNS data can be used to assess flooding. We believe that SNS data can be effective in determining flood levels even in places where it is difficult to install, operate, and manage observation devices. Although there have been many studies on flood damage detection using SNSs, their effectiveness has not been clarified, and the amount and content of data collected are not fixed depending on the case of flood damage. This estimation method could further improve the accuracy of flooding assessment by adding SNS data as well as observation data and calculating flooding conditions for the entire affected area. Thus, we considered it necessary to validate the effectiveness of the SNS data. For the estimation, we utilized the method we have developed for estimating the expansion process of flooding using observation time-series data. This is a state-space model that uses observation data to compensate for flooding analysis simulations.

This paper validates the SNS data using the following procedure based on our state-space model. The purpose of SNS data validation is to investigate whether SNS data can contribute to the accuracy of flood assessment and under what conditions SNS data can improve the accuracy. Since we were unable to determine the error rate contained in the SNS data collected at the time of flooding, we generate quantified SNS data from simulations in this paper. To generate the SNS data, we used time-series data collected from observation devices at multiple locations. The results of flood analysis simulation were assimilated using the time-series data to simulate the SNS locations, the timing of postings, and numerical flood levels (Fig. 2(b)(i)). Then, we appended errors to the simulated SNS data (Fig. 2(b)(ii)). Afterward, we regenerated a flood level at the observation location based on the simulated SNS data with errors and examined the accuracy of the errors on the data assimilation accuracy (Fig. 2(b)(iii)).

The estimation results show that even if the posted SNS data have a time lag and location gap, the error is small, averaging approximately 0.10 meters for a small area within 100 meters, except for a certain observation location. In cases where water levels contained errors and fake data were posted, we found that the errors were small and that if the ratio of fake data was small, the estimation accuracy would not be significantly reduced. All evaluations showed significant improvements compared to the state-space model with observations. Unlike time-series data, SNS data are sparse in the time direction, although it is effective for estimation if collected at locations that are near the estimated location. However, some of the sensors show large errors in the process of calculating simulated SNS data. These errors appeared because waterway and sewer data were not used in the state-space model. Since waterway and sewer data are difficult to collect, we plan to resolve this issue statistically using a Kalman filter. In addition, while this paper evaluates a flood case from a waterway, it is necessary to improve the accuracy by integrating observed time series data provided by water level observation devices so that the method can adapt to rapid water level rising, such as floods caused by large outflows from rivers. After a further improvement in accuracy, we will develop and implement Phase 2: Flood Assessment Promotion Requirement and Phase 3: Flood Assessment.

ACKNOWLEDGEMENT

This work was supported by JST, PRESTO Grant Number JPMJPR2036, and the commissioned research(No. 05401) by National Institute of Information and Communications Technology (NICT), Japan.

REFERENCES

[1] "Intergovernmental Panel on Climate Change Fifth Assessment Report (AR5)", Retrieved April 17, 2018, from https://www.ipcc.ch/report/ar5/.

[2] P. Milly, D. Christopher, T. R. Wetherald, A. K. Dunne, T. L. Delworth, "Increasing Risk of Great Floods in a Changing Climate, Nature, Nature Publishing Group", Vol.415, No.6871, page514, (2002).

[3] Y. Hirabayashi, R. Mahendran, S. Koirala, L. Konoshima, D. Yamazaki, S. Watanabe, H. Kim, S. Kanae, "Global Flood Risk under Climate Change", Nature Climate Change, Nature Publishing Group, Vol.3, No.9, page816, (2013).

[4] J. Ernst, J. B. Dewals, S. Detrembleur, P. Archambeau, S. Erpicum, M. Pirotton, "Micro-scale Flood Risk Analysis based on Detailed 2D Hydraulic Modelling and High Resolution Geographic Data", Natural Hazards, Springer, Vol.55, No.2, pp.181–209, (2010).

[5] A. S. M. Saudi, I. S. D. Ridzuan, A. Balakrishnan, A. Azid, D. M. A. Shukor, Z. I. Rizman, "New Flood Risk Index in Tropical Area Generated by using SPC Technique", Journal of Fundamental and Applied Sciences, Vol.9, No.4S, pp.828–850, (2017).

[6] S. F. Silva, M. Martinho, R. Capitão, T. Reis, C. J. Fortes, J. C. Ferreira, "An Index-based Method for Coastal-flood Risk Assessment in Low-lying Areas (Costa de Caparica, Portugal)", Ocean & Coastal Management, Elsevier, Vol.144, pp.90–104, (2017).

[7] D. H. Burn, "Perceptions of Flood Risk: A Case Study of the Red River Flood of 1997", Water Resources Research, Wiley Online Library, Vol.35, No.11, pp.3451–3458, (1999).

[8] J. W. Hall, I. C. Meadowcroft, P. B. Sayers, M. E. Bramley, "Integrated Flood Risk Management in England and Wales", Natural Hazards Review, American Society of Civil Engineers, Vol.4, No.3, pp.126–135, (2003).

[9] R. Miceli, I. Sotgiu, M. Settanni, "Disaster Preparedness and Perception of Flood Risk: A Study in an Alpine Valley in Italy", Journal of Environmental Psychology, Elsevier, Vol.28, No.2, pp.164–173, (2008).

[10] L. Pulvirentia, M. Chinib, N. Pierdiccaa, L. Guerrieroc, P. Ferrazzolic, "Flood Monitoring using Multi-temporal COSMO-SkyMed Data: Image Segmentation and Signature Interpretation", Remote Sensing of Environment, Vol.115, No.4, pp.990-1002, (2011).

[11] E. A. Basha, S. Ravela, D. Rus, "Model-based Monitoring for Early Warning Flood Detection", In Proceedings of the 6th ACM conference on Embedded network sensor systems (SenSys), pp.295-308, (2008).

[12] A. Elshorbagy, G. Corzo, S. Srinivasulu, D. Solomatine, "Experimental Investigation of the Predictive Capabilities of Data Driven Modeling Techniques in Hydrology —part 2: Application", Hydrology and Earth System Sciences, Vol.14, pp.1943-1961, (2010).

[13] A. Rafieeinasab, A. Norouzi, S. Kim, H. Habibi, B. Nazari, D. Seo, H. Lee, B. Cosgrove, Z. Cui, "Toward High-resolution Flash Flood Prediction in Large Urban Areas – Analysis of Sensitivity to Spatiotemporal Resolution of Rainfall Input and Hydrologic Modeling", Journal of Hydrology, Vol.531, part 2, pp.370-388, (2015).

[14] F. A. Ruslan, A. M. Samad, Z. M. Zain, R. Adnan, "Flood Prediction using NARX Neural Network and EKF Prediction Technique: A Comparative Study", In Proceeding of the IEEE 3rd International Conference on System Engineering and Technology (ICSET), pp.203-208, (2013).

[15] S. K. Sinnakaudan, A. A. Ghani, M. S. S. Ahmad, N. A. Zakaria, "Flood Risk Mapping for Pari River Incorporating Sediment Transport", Environmental Modelling & Software, Elsevier, Vol.18, No.2, pp.119–130, (2003).

[16] H. M. Lyu, W. J. Sun, S. L. Shen, A. Arulrajah, "Flood Risk Assessment in Metro Systems of Mega-cities using a GIS-based Modeling Approach", Science of the Total Environment, Elsevier, Vol.626, pp.1012–1025, (2018).

[17] N. M. Hunter, P. D. Bates, M. S. Horritt, M. D. Wilson, "Simple Spatially-distributed Models for Predicting Flood Inundation: A Review", Geomorphology, Elsevier, Vol.90, No.3-4, pp.208–225, (2007).

[18] M. S. Horritt, P. D. Bates, “Evaluation of 1D and 2D Numerical Models for Predicting River Flood Inundation”, Journal of Hydrology, Elsevier, Vol.268, No.1-4, pp.87–99, (2002).

[19] J. Kim, M. Hastak, “Social Network Analysis: Characteristics of Online Social Networks after a Disaster”, International Journal of Information Management, Vol.38, No.1, pp.86–96, (2018).

[20] F. K. Sufi, I. Khalil, “Automated Disaster Monitoring From Social Media Posts Using AI-Based Location Intelligence and Sentiment Analysis”, IEEE Transactions on Computational Social Systems, pp.1–11, (2022).

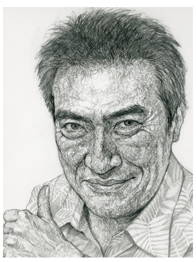
[21] H. N. Teodorescu, “Emergency-Related, Social Network Time Series: Description and Analysis”, Time Series Analysis and Forecasting, Springer International Publishing, pp.205–215, (2016).

[22] K. Hiroi, D. Murakami, K. Kurata, T. Tashiro, Y. Shinoda, “A Proposal of Data Assimilation Approach for Flood Level Estimation and Evaluation with Urban Flood Disasters”, Journal of Information Processing: Consumer Devices and Systems, Vol.10, No.3, pp.55–64, (2020) (In Japanese).

(Received: January 16, 2023)
(Accepted: August 5, 2023)

Kei Hiroi received her Master of Media Design and Ph.D. in Media Design in 2011 and 2014, respectively from Keio University. She has been an assistant professor in the department of Information and Communication Engineering, Graduate School of Engineering, Nagoya University. She is currently an associate professor in Disaster Prevention Research Institute, Kyoto University. Her research interests include disaster simulation, and crisis computing.

Akihito Kohiga is a Project Researcher of Information Sciences at Japan Advanced Institute of Science and Technology. He received Ph.D (Information Science) in 2020. His research interests lie in cloud computing, massive distributed simulation, modeling and architecture, especially in flood and evacuation fields.



Yoichi Shinoda is currently a professor of Japan Advanced Institute of Science and Technology. His research interests include Impact of digital technologies on human activities, Parallel and Distributed Systems, Networking Protocols and Systems, and Information Handlinf Systems.