# Visualization of relationships between precipitation and river water levels

Yuki TANAKA
Ochanimizu University
Tokyo, Japan
Email: g1920522@is.ocha.ac.jp

Grammatikaki Angeliki
Vienna University of Technology
Vienna, Austria
Email:angeliki.grammatikaki@tuwien.ac.at

Ehlers Henry
Vienna University of Technology
Vienna, Austria
Email:henry.ehlers@tuwien.ac.at

Renata Raidou
Vienna University of Technology
Vienna, Austria
Email:rraidou@cg.tuwien.ac.at

Eduard Gröller
Vienna University of Technology
Vienna, Austria
Email: groeller@cg.tuwien.ac.at

Takayuki Itoh
Ochanomizu University
Tokyo, Japan
Email: itot@is.ocha.ac.jp

***Abstract*—Observation of precipitation changes is important for a variety of purposes such as predicting river levels. Previous studies for data visualization of precipitation and river water levels plotted graphs and color bars were many stations on a map. Instead of such visualizations on a map, we construct a graph to imitate a connected structure such as a tributary of a river in this study. Our method displays two pseudo-coloring sparklines at nodes of the graph as the stations. The method can visualize the time difference between the increase in precipitation upstream and the increase in river water level downstream. Users can observe precipitation and river water levels at different observation points. Our method uses a Delaunay diagram connecting gauging positions to interpolate and calculate precipitation at river level observation points. This avoids the discrepancy between observation points.In addition, we adjust the amount of visualized information by skipping the display of several observation points based on the similarity of the time-series data at each station, which is calculated by applying the dynamic time-stretching method. The visualization results show that downstream, once the water level rises, it tends to take longer for the water level to drop. In addition, the results show that a time lag occurs between the increase in precipitation and the rise in river levels in the mainstream, while tributaries have little time lag. In addition, data on rainfall and river levels at the same station over multiple periods and their relationship are plotted as scatter plots. The scatter plots make it easier to compare data from multiple periods at the same time than two-tone pseudo coloring sparklines.**

***Index Terms*—Meteorological Information,River Water level,Geographic Information**

## I. Introduction

River levels are closely related to changes in precipitation. Therefore, observation of precipitation changes is an important factor in predicting river levels. In the existing methods, the placement of meteorological data on a map enabled the simultaneous visualization of numerical and geographical information. However, it is difficult to visually capture information along the river flow while using such visualization methods especially if the geometry of the river is complex. The objective of our research is to visualize water level rise in downstream rivers after rainfall upstream.

To address this problem, we propose a method to visualize the time difference between the increase in precipitation upstream of a river and the increase in water level downstream. This method constructs and displays a graph that imitates the flow of a river. In our visualization, the vertical direction of the graph depicts the flow of a river from upstream to downstream, while the horizontal direction depicts the connections between tributaries and the main river. The method places two-tone pseudo coloring sparklines plotting precipitation and river level data as the nodes of the graph to make it easier to compare the numerical changes between them. In this paper, as an extension of our previous method [1], we propose a method for adjusting the amount of visualizing information that calculates the similarity of multiple time-series data using the Dynamic Time Warping (DTW) method, and combines the data of observation points that have high similarity into one. This process makes the visualization simpler and therefore enables us easier to discover highly important parts of the data to the user at many observation points.

Here, it is often difficult to compare data of many periods at the same time by using two-tone pseudo coloring sparklines. Therefore, we applied a scatterplot in addition to the sparklines. Plotting not only the data values of rainfall and river transition themselves, but also data showing their relationship (e.g., how many hours after a heavy rainfall the water level returns to normal) as a scatterplot allows users to easily read the relationship between precipitation and river water level.

## II. Related Work

This section introduces the details of three types of related work as follows.

The first is a method for visualizing precipitation and river level data on a map. Yagi et al. [2] and Ibrahim et al. [3] presented that the proper arrangement of detailed information at each observation point on a map allows numerical information from many stations to be displayed while preserving

2375-0138/24/$31.00 ©2024 IEEE
DOI 10.1109/IV64223.2024.00020

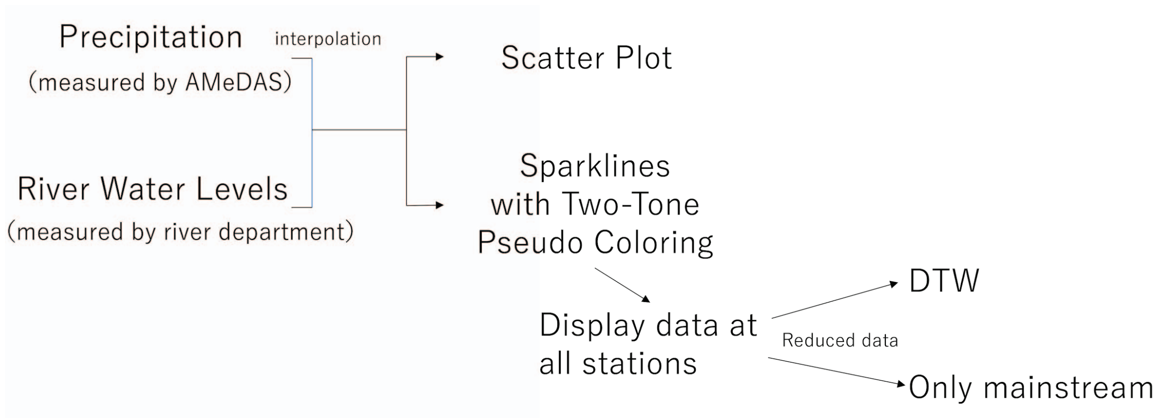



Fig. 1. Overview of the proposed method.

geographical features. However, although this method visualizes the location of the stations, it is often difficult to visually capture the information along the river flow especially if the geometry of the river is complex.

The second is two-tone pseudo coloring sparklines. Saito et al. [4] presented that two-tone pseudo-color improves the accuracy of reading changes in data in a smaller space than with conventional line graphs. The method uses two discrete colors to separate the top and bottom parts of the graph. Two types of pseudo-color scales are used for different purposes: linear scale display, in which the range of values indicated by each color is constant and the reading accuracy is also constant, and log scale, in which a wide range of values can be handled with a small number of colors.

The third is a study on dynamic time-warping (DTW). Müller [5] defined the DTW distance as the sum of the distances obtained by mapping the two time series data in such a way that the distance is minimized at each point of the two time series data.

## III. Proposed Technique

### A. Interpolation of rainfall data

Precipitation data and river water level data are usually measured at different locations. For example, precipitation data in Japan is generally observed by AMeDAS of the Japan Meteorological Agency(JMA), while river water level data is measured by water level gauges managed by the river department of each municipal government under the jurisdiction of the Ministry of Land, Infrastructure, Transport and Tourism. Therefore, we obtained the precipitation at the river level observation points by interpolating them.

This section defines the known precipitation values at precipitation observation points $p_0$, $p_1$ and $p_2$, as $C_0$, $C_1$ and $C_2$. Specifically, our implementation calculates the position coordinates $u_p$, $v_p$ in the local coordinate system of a point $p$ surrounded by three points $p_0$, $p_1$ and $p_2$, by the following formula.

$$p = p_0 + u_p(p_1 - p_0) + v_p(p_2 - p_0) \tag{1}$$

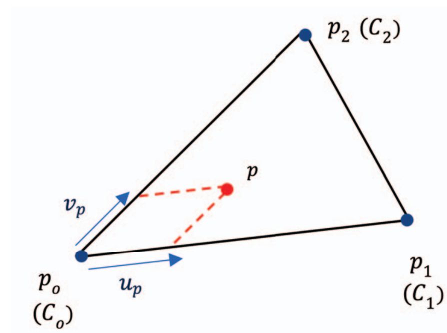



Fig. 2. Linear interpolation inside a triangular element

The amount of precipitation $C$ at the point $p$ is linearly interpolated inside a triangular element.

$$C = C_0 + u_p(C_1 - C_0) + v_p(C_2 - C_0) \tag{2}$$

### B. Scatterplot

As an extension of our own previous study [1], we draw the data as several types of scatterplots in order to compare trends in the data over multiple time periods. The implementation of scatterplots aims to observe the relationship between precipitation and river water levels.

Therefore, we use data on the relationship between rainfall and water level at the same station over multiple periods. Specifically, our implementation draws scatterplots by applying two of the following five variables.

- **Water level difference**
  Maximum water level minus minimum water level (meters).
- **Time difference until water level rises**
  Time from maximum water level to maximum rainfall in three hours (minutes).
- **Time difference until water level recedes**

The value obtained by subtracting the time of maximum rainfall for three hours from the time when the water level returns to its original level (minutes).

- **Maximum rainfall**
  Maximum rainfall for three hours (millimeters).
- **Rainfall duration**
  The longest duration of rainfall (minutes).

The scatterplots do not directly apply the rainfall data itself, which is observed every ten minutes by the JMA in our dataset, but the accumulated precipitation over the previous three hours. Here, if we use the rainfall data observed every ten minutes, the maximum water level would be reached later in many cases than the time when the rainfall was momentary. In general, the water level rises after a certain period of rainfall.

### C. Time-varying representation by sparklines with two-tone pseudo coloring

This method uses two-tone pseudo-color display (Figure 3) to represent both river water level data at river level observation points and precipitation data obtained by interpolating them. The use of pseudo-colors with linear colors facilitates the reading of data changes over time. Here, our implementation applies five colors for sparklines.

We applied the log scale for precipitation data because this implementation can handle a wide range of values with a small number of colors. The log scale avoids the difficulty in reading normally appearing small values. This problem can be caused by large values of data that appear only rarely and take up more space than necessary.

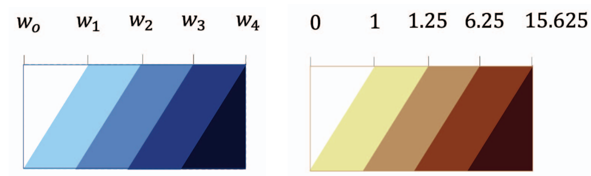
In contrast, we applied a linear scale to the river level data. The ranges of values indicated by each color is the same, and therefore the reading accuracy is also constant. Our implementation defines the linear scale of river level data as follows. Normal river water level data differs depending on the observation points. Therefore, the minimum and maximum values of river water level data for a certain location over a seven-day period are set as $w_{min}$ and $w_{max}$, respectively. At the same time, the boundary $w_{0 \sim} w_4$ of the numerical interval is set for each observation location using the following equation, where $t = (w_{max} - w_{min})/4$.

$$\begin{aligned} w_0 = w_{min}, w_4 = w_{max}, \\ w_{i+1} = w_i + t(i = 1, 2, 3) \end{aligned} \tag{3}$$

The color assigned to the range $[w_m, w_{m+1}]$ is defined as $c_m$. The color assigned to the range of values $w_4$ or greater is defined as $c_5$. When the value $w$ at any given time is in the range of $[w_m, w_{m+1}]$, the top $(w_{m+1} - w)/t$ is colored $c_{m+1}$ and the bottom $(w - w_m)/t$ is colored $c_m$.

### D. Adjustment of the number of sparklines to be displayed by dynamic time warping

As an extension of our own previous study [1], we employ the dynamic time-warping (DTW) method to calculate the similarity of time-series values at multiple observation points. The DTW distance is the sum of the distances of the two time-series data points, which is the result of a brute-force search for the path with the shortest distance between the two points. Waveforms with close DTW distances have high similarity, while waveforms with far distances have low similarity. Our implementation reduces the amount of displaying information by applying DTW and skipping the display of station informaion that have highly similar values with another station.



Water Level. Precipitation.

Fig. 3. Color scale.

Two time series data for calculating similarity are defined as $S = \{s_1, s_2, ....s_m\}$ and $T = \{t_1, t_2, ..., t_n\}$. The path connecting each point in the two time series data are defined as $W = \{w_1, w_2, ..., w_k\}$. The distance of each point is described as $(i, j) = |s_i - t_j|$. The DTW distance is obtained by the following formula.

$$DTW(S, T) = \min_{W} \sum_{k=1}^{P} \delta(w_k) \tag{4}$$

## IV. Results

### A. Dataset

This section presents visualization results using precipitation data observed every ten minutes at a total of 89 locations (43 in Niigata Prefecture and 46 in Nagano Prefecture in Japan) by JMA and hourly river water level data observed in the Shinano River system by the river bureaus in Niigata and Nagano prefectures. In this study, among the weather cases that caused disasters in Niigata Prefecture, cases in which the cause was heavy rainfall were extracted, and data for the following seven periods were used as examples. On the scatterplot, the year and month of the data are indicated in following parentheses.

- July 10-18, 2004 (Jul.,2004)
- July 24-August 5, 2011 (Jul.,2011)
- June 29-July 7, 2017 (Jun.,2017)
- August 27-September 3, 2018 (Aug.,2018)
- September 3-6, 2018 (Sep.,2018)
- August 22-26, 2021 (Aug.,2021)
- August 1-8, 2022 (Aug.,2022)

Our implementation applied Delaunay triangulation (Figure 4) to a total of 89 locations where precipitation were observed every hour, with the latitude and longitude at each location set as coordinates. Then, it calculated the precipitation at river level observation point $p$, which is surrounded by precipitation observation points $p_0$, $p_1$, and $p_2$ that make up a triangle, by linear interpolation.

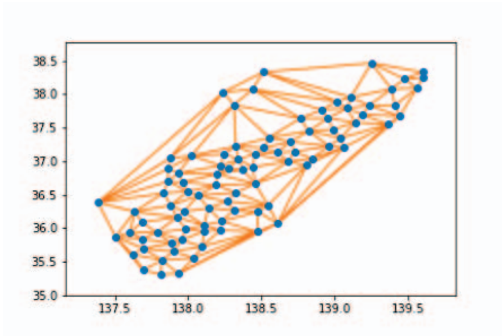



Fig. 4. Delaunay triangulation with observation points as vertices.

### B. Scatterplot

Our implementation draws scatterplots to represent the relationship between water level and precipitation over multiple time periods. The scatterplots apply two of the five different variables defined in Section III-B at the same station (Usui-bashi in this case as an example) over the seven periods described in Section IV-A. We draw ten scatterplots, applying two of the five variables, as shown in Figure 5.

Among these, this section focuses on a scatterplot, shown in Figure 6, that represents characteristic trends well. This scatterplot is drawn with Water level difference assigned to the vertical axis and Time difference until water level recades assigned to the horizontal axis. The colors in this scatterplot indicate the Rainfall duration by using in a colormap shown in Figure 7. The blue line depicts the regression line on the scatterplot. The values in June 2017, September 2018, and July 2004 are plotted away from the regression line in this scatterplot. The values in July 2017 and July 2004 are from the period when the duration of rain is extremely long. June 2017 is a rainy season front, September 2018 is an autumn rain front, and July 2004 is a heavy rainfall caused by a rainy season front [7]. Other periods were caused by other weather phenomena such as typhoons. We could observe that the stagnation of the front and the long duration of rainfall produce differences in the relationship between "Water level difference" and "Time difference until water level recade" compared to other meteorological conditions.

Thus, by utilizing the scatterplots, we can observe that differences in the way the water level changes are occurred depending on the rainfall trend.

### C. Sparklines with Two-tone pseudo coloring

After observing scatterplots of multiple time periods, users can use the sparkline visualization to see the details of the water level and rainfall data for a given time period. As an example, Figure 8 shows the data observed at the Usui-bashi in the Shinano River system in Niigata Prefecture from August 1 to 7, 2022. The top part of the figure represents precipitation data and the bottom part represents river water level data. The darker colored areas in the sparklines have larger values. Two pseudo-coloring sparklines illustrate a time lag between the

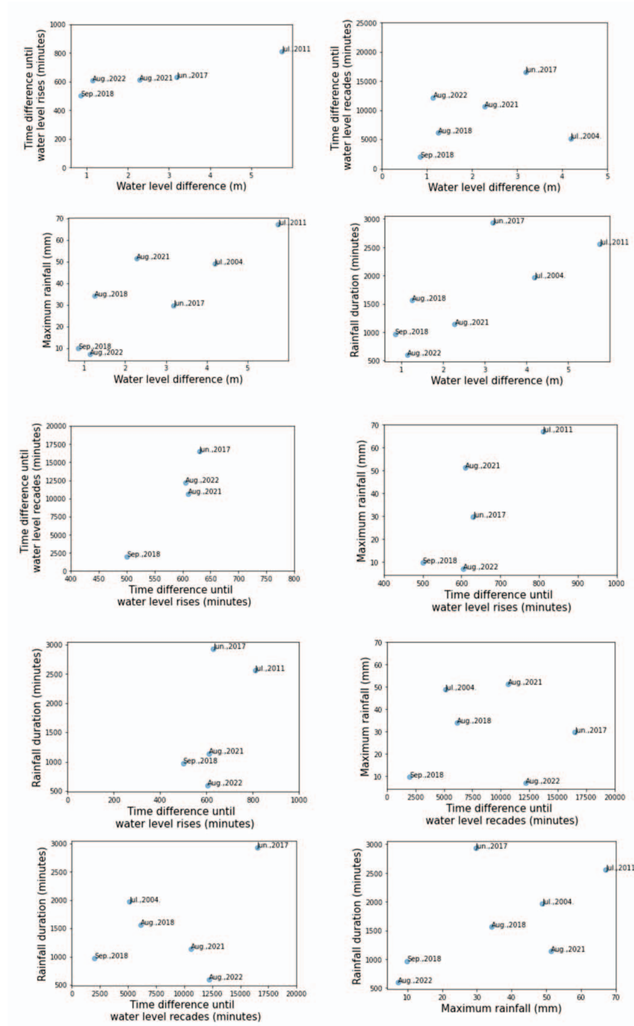



Fig. 5. Scatter plots over seven periods in Usui-bashi

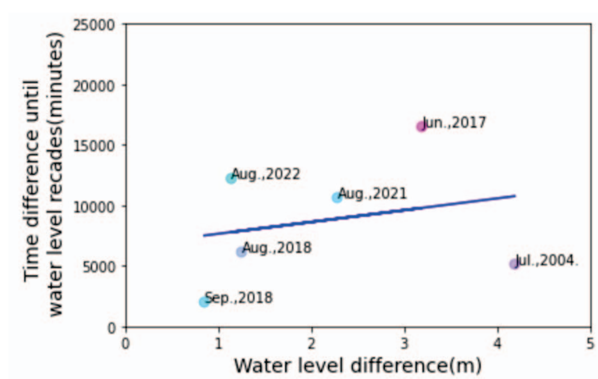



Fig. 6. The scatterplot that represents characteristic trends at Usui-bashi.

Fig. 7. Color bar used in the scatterplots.

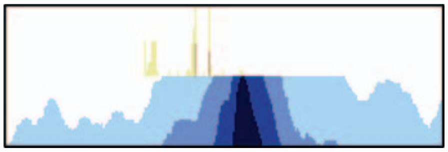


Fig. 8. two-tone sparklines at the Usui-bashi. (top) precipitation data. (bottom) water level data.

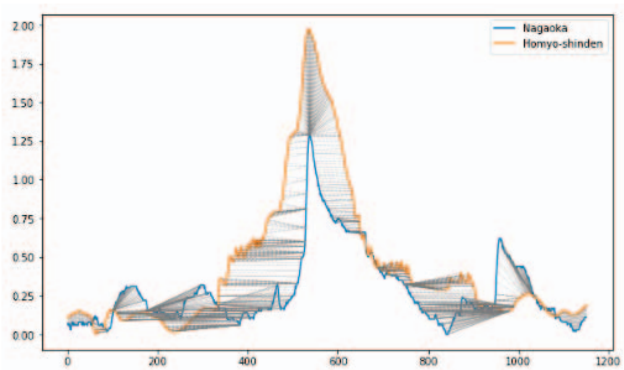



Fig. 9. Application of dynamic time warping to the water level values at Nagaoka and Homyo-Shinden.

increase in precipitation and the rise in water level at the Usui-bashi.

### D. Dynamic time warping

While generating the dataset, we obtained the time-series values of water levels at each station by subtracting the values of the normal water level defined by the river authority from the time-series data of water levels. Figure 9 shows an example of the dynamic time-warping method to the river level values observed at two water level stations in the Shinano River system, Nagaoka and Homyo-Shinden, Niigata Prefecture, from August 1 to 7, 2022. This process finds the path with the shortest total distance between paths at each point of the two time series data. This line chart illustrates that characteristic waveforms such as the peak of the values correspond to each other in two time series data.

### E. Arrangement of sparklines

Our implementation arranged two pseudo-coloring sparklines depicting precipitation data and river water level data at 28 locations in a graph structure that imitates the flow of a river. Figures 10 to 12 shows the examples as follows.

- Figure 10 displays only the data from stations located in the main stem of the river.
- Figure 11 displays the data at all stations in the Shinano River system.
- Figure 12 selectively displays the data of two adjacent locations. If they are highly similar (the sum of the DTW distances for the two water level data is less than the threshold), just one of the two locations is displayed.

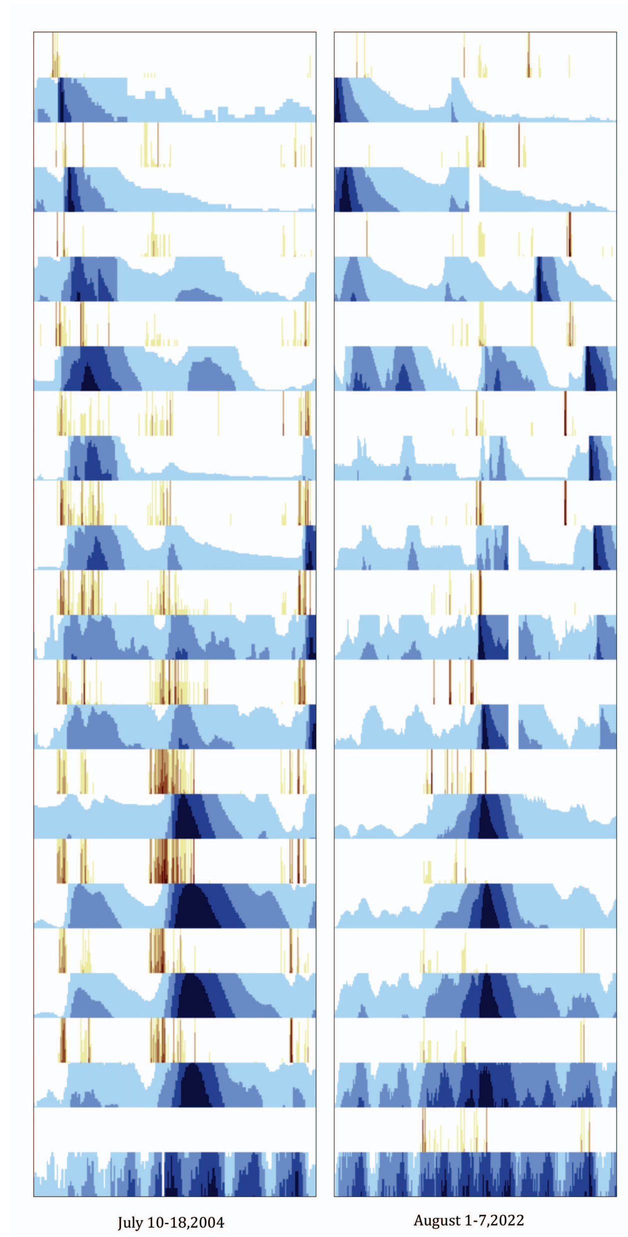



Fig. 10. Data in the main stream of the Shinano River

Users can select one of the above three modes. Readability of the visualization can be improved by selecting the above modes and consequently adjusting the amount of information to be displayed.

As a result, we could find the following trend by visualizing the relationship between precipitation and river level in the Shinano River and displaying the data from all the stations.

- Figure 10 shows that the river water lever tends to take longer to decrease after it rises downstream.
- Figure 11 shows that a time difference tends to occur

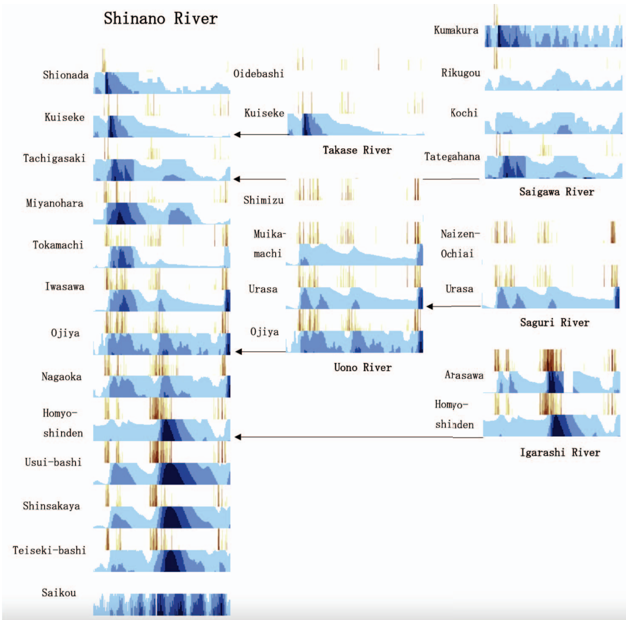



Fig. 11. Data From all stations in the Shinano River system.

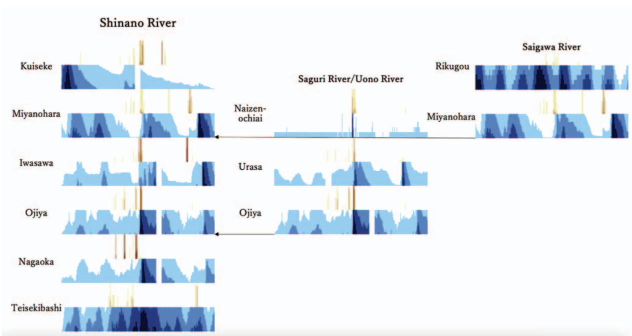



Fig. 12. Data after applying the dynamic time warping method.

between an increase in precipitation and a rise in river water level in the main stem of the Shinano River; on the other hand, this time difference is hardly observed in the tributaries of the Uono River (Nobori River) and the Mikuni River.

## V. Discussion

Pseudo-coloring sparklines are suitable for visualization of changes in river water level values because they are continuous. On the other hand, when visualizing changes in precipitation values, there remains the problem that the width of painted regions of pseudo-coloring sparklines tend to narrow during periods of temporary very heavy rainfall because precipitation values are often discontinuous. It may make it difficult to read the changes in precipitation values. In order to solve such shortcomings, we would like to consider the use of difference values instead of raw precipitation values.

The ultimate goal of our research is to develop a visualization system that can assist in predicting the likelihood of flooding by, for example, indicating locations where water levels take time to drop even after rainfall subsides, based on observed historical data. For this purpose, it is necessary to classify the trends of rainfall and water level changes in past data. So, by applying scatterplots such as the one proposed in this paper to a larger period of data, we will investigate what type of data is appropriate to be used for classification.

## VI. Conclusion

This paper presented a method to visualize the relationship between river water level and precipitation applying two visualization components: scatterplots and sparklines with two-tone pseudo coloring. Precipitation at river level observation points is calculated by interpolation, and visualized using the data of river level and precipitation at those points. Two pseudo-coloring sparklines can display time-series values in a smaller space than the conventional line charts, thus enabling the visualization of data from a large number of stations. By arranging two pseudo-coloring sparklines along the flow of the river from upstream to downstream and from tributary to main stream, we could compare the time difference between rainfall upstream and downstream, and the difference in the rise of water levels in tributaries and main streams. In addition, as an extension of our own previous study, we calculated the similarity of multiple time-series values by applying the dynamic time warping method, and displayed only one of them if the numerical values are highly similar among neighboring stations. This enables the user to adjust the amount of time-series values to be visualized at a time. We also developed a new scatterplot using the relationship between water level and precipitation. It enabled the visualization of the relationship between two types of time-series values during the same period over multiple time periods.

## References

[1] Yuki Tanaka, Takayuki Itoh, Visualization of relationships between precipitation and river water levels (in Japanese), Proceedings of the Symposium on Visualization, OS12-5, 2023.

[2] Sayaka Yagi, Takayuki Itoh, Mayumi Kurokawa, Yuichi Izu, Takahisa Yoneyama, Takashi Kohara, Data Layout and Level-of-Detail Control for Flood Data Visualization upon Maps, IIEEJ Transactions on Image Electronics and Visual Computing, Vol. 3, No. 2, pp. 101-113, 2015.

[3] Ibrahim Demir, Witold F. Krajewski, Towards an integrated Flood Information System: Centralized data access, analysis, and visualization, Environmental Modelling & Software, Vol. 50, pp. 77-84, 2013.

[4] Takafumi Saito, et al., Two-Tone Pseudo Coloring: Compact Visualization for One-Dimensional Data, IEEE Symposium on Information Visualization, 2005.

[5] Meinard Müller, Dynamic time warping, Information Retrieval for Music and Motion, pp. 69-84, 2007.

[6] Naofumi Shibata, Naoki Kita, Takafumi Saito, A Study on Color Scheme of two-tone Pseudo-color Display in Visualization of Meteorological Data (in Japanese), The 84th National Convention of IPSJ, 5ZF-03, 2022.

[7] Weather Cases that Brought Disasters to Niigata Prefecture (Local Meteorological Observatory in Niigata-prefecture) https://www.data.jma.go.jp/niigata/menu/saigai_NI.html(accessed on June 14, 2024)