Master Thesis

TVVR 21/5010

Diagnosis of Synoptic Weather Patterns Causing Heavy Rainfall Occurrence Using Self-Organizing Maps

Aoi Kasai

Division of Water Resources Engineering

Department of Building and Environmental Technology

Lund University

Diagnosis of Synoptic Weather Patterns Causing Heavy Rainfall Occurrence Using Self-Organizing Maps

By:
Aoi Kasai

Master Thesis

Division of Water Resources Engineering
Department of Building & Environmental Technology
Lund University
Box 118
221 00 Lund, Sweden

Water Resources Engineering
TVVR-21/5010
ISSN 1101-9824

Lund 2021
www.tvrl.lth.se

Master Thesis

Division of Water Resources Engineering

Department of Building & Environmental Technology

Lund University

Swedish title:	Diagnosticering av storskaliga vädersystem orsakande skyfall med hjälp av SOM
English title:	Diagnosis of Synoptic Weather Patterns Causing Heavy Rainfall Occurrence Using Self-Organizing Maps
Author(s):	Aoi Kasai
Supervisor:	Koji Nishiyama, Cintia Bertacchi Uvo
Examiner:	Magnus Persson
Language	English
Year:	2021
Keywords:	Self-Organizing Map; heavy rainfall frequency; classification; meteorological fields: prediction: heavy rainfall risk

Acknowledgement

This paper is the product of the degree project (VVRM01) throughout my master course at both of the department of Urban and Environment System Engineering, Kyushu University and the department of Faculty of Engineering LTH, Water Resources Engineering, Lund University.

Due to the uncertainty of COVID-19 pandemic, my schedule of double degree program was modified, and it would be the hardest situation for double degree student to handle this problem. However, thanks to the international coordinators both in Lund and Japan, I can continue the thesis work from Japan, and I have the opportunity to present my findings to professors in Lund.

I would like to express my gratitude to my supervisor, Assistant Professor Koji Nishiyama. Mr. Nishiyama gave me significant support such as teaching the background of this study and discussing what should be revealed in this research. Moreover, I would also want to show my gratitude to Associate Professor Yoshinari Hiroshiro. As is mentioned above, due to the uncertain situation, I almost lost my motivation to finalize the double degree program before starting my thesis work, however, Mr. Hiroshiro encouraged me, and his word made me positive during the work.

I would like to express my gratitude to my supervisor, Professor Cintia Bertacchi Uvo and my examiner, Professor Magnus Persson. They gave me positive feedback about my thesis work and told me what should be improved in the future study.

I would like to express my gratitude to people who met in Lund including professors and classmates of water resources engineering. When I started my degree program in Lund, it was pretty hard to catch up the classes due to the lack of my academic knowledge and English skills, however, they helped and gave me right advices. After adapting to the academic life in Lund, I could discuss the ongoing problem such as climate change with professors and they told me an interesting consideration, which can deepen my academic interest.

I would like to appreciate my colleague Tatsuya Nagino and Shunsuke Takeuchi for supporting each other during master program and gave me effective advice about the thesis work.

I would like to appreciate my family and friends for helping everything in Kyushu University and Lund university.

Finally, the JRA-55 dataset that is employed in this study is provided by the Japan Meteorological Agency (JMA).

Abstract

The annual number of events with precipitation over 50 mm/h have increased from 1970s and this trend is likely to be continued due to climate change in Kyushu, Japan. Therefore, it is important to recognize what kind of meteorological fields have contributed to the occurrence of heavy rainfall events and to reveal whether it is possible to diagnose heavy rainfall risk. Firstly, this study combines the Self-Organizing Map (SOM) and Radar/Rain gauge analyzed precipitation to provide the distribution of heavy rainfall frequency on the two-dimensional map. As a result, 19520 meteorological fields observed for 40 years are classified into 40 synoptic weather groups and the top 10 groups are characterized by four reasons, the existence of 1) strong southwest wind and large amounts of precipitable water (PW), 2) counterclockwise circulation with large PW, 3) tropical cyclone and 4) stationary front. Secondly, the Global Spectral Model is combined with the structured SOM for diagnosing the probability of heavy rainfall occurrence within a range of few days. Following the case studies, the probability of heavy rainfall occurrence can be increased and stabilized around 36 h before heavy rainfall events. Furthermore, the SOM can relate the diagnosed patterns to historical rainfall events.

Table of Contents

Chapter 1: Introduction

1.1 Background

The fifth synthesis report (IPCC, 2014) of the International Panel on Climate Change suggested the frequency and intensity of heavy rainfall events has likely increased over the second half of the 20th century due to the anthropogenic forcing. Human influence on the climate system is thought to be clear, and recent anthropogenic emissions of greenhouse gases are the highest in history, which has caused widespread impacts on human and natural systems.

In Japan, JMA(Japan Meteorological Agency) investigated the long-term trends in precipitation recorded from 51 observation stations since 1901, which suggested the annual number of days with precipitation of ≧100 mm, ≧200 mm, respectively, are virtually certain to have increased over 100 years, as shown by Figure 1.1. Moreover, JMA introduced the Automated Meteorological Data Acquisition System abbreviated as "AMeDAS" in the latter part of the 1970s. The period is relatively shorter than above-mentioned dataset, however, there are around twenty-five times as many AMeDAS station as them. Therefore, AMeDAS is better equipped to capture heavy rainfall events that take place on a limited spatial scale. Figure 1.2 describes that the annual numbers of events with precipitation of ≧50 mm and ≧ 80 mm per hour in AMeDAS stations, respectively, have increased (statistically significant at a confidence level of 99 %).

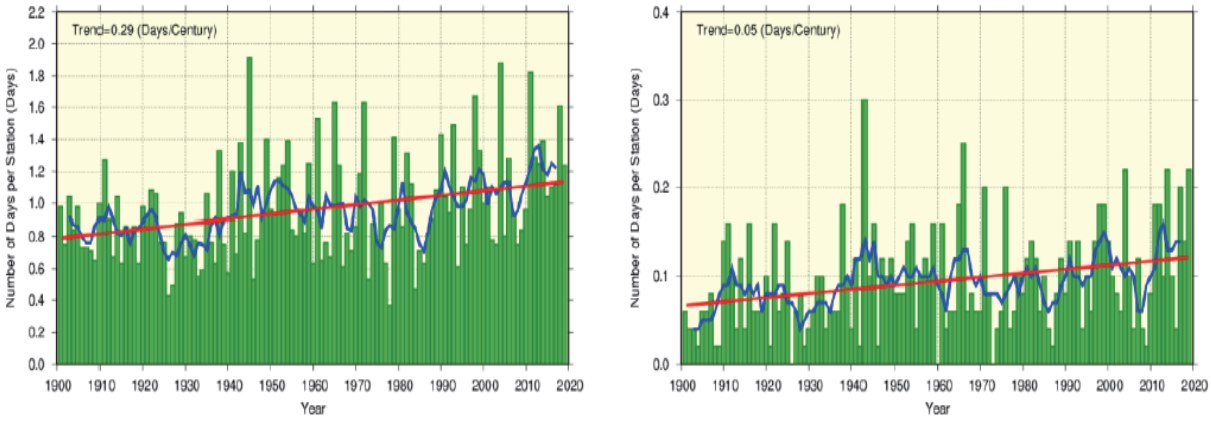


Figure 1.1: Annual number of days with precipitation ≧100 mm (left figure) and ≧200 mm (right figure) from 1901 to 2019. The green bars indicate the annual number of days per station for each year. The blue line indicates the five-year running mean, and the straight red line indicates the long-term liner trend (JMA, 2020).

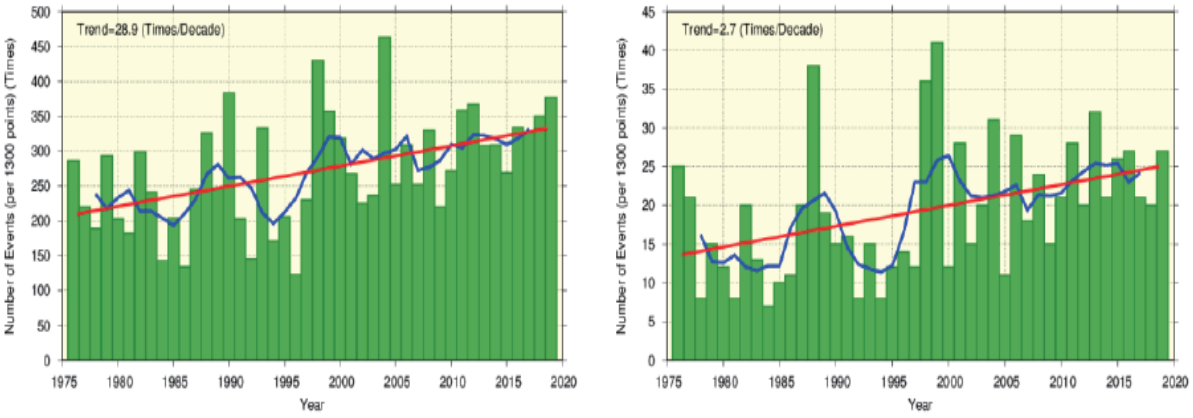


Figure 1.2: Annual number of events with precipitation of ≧50 mm and ≧80 mm per hour from 1976 to 2019. The green bar indicates the annual number of events per 1,300 AMeDAS stations for each year, and the straight red line indicates the long-term liner trend (JMA, 2020).

In Kyushu area, where it is in the western part of Japan, torrential rainfall events have frequently occurred for 39 years. Figure 1.3 shows the annual variation in heavy rainfall frequency (≧70 mm per hour) in Kyushu area and it is based on the data from AMeDAS stations which have observed continuously from 1979 to 2017. When dividing 39 years into two periods (period A and B), annual heavy rainfall frequency in each period was calculated. Then, it is found that the increasing rate of annual average is around 80%. In other words, there is an increasing trend in the occurrence of heavy rainfall events in Kyushu area (Ohashi, 2018).

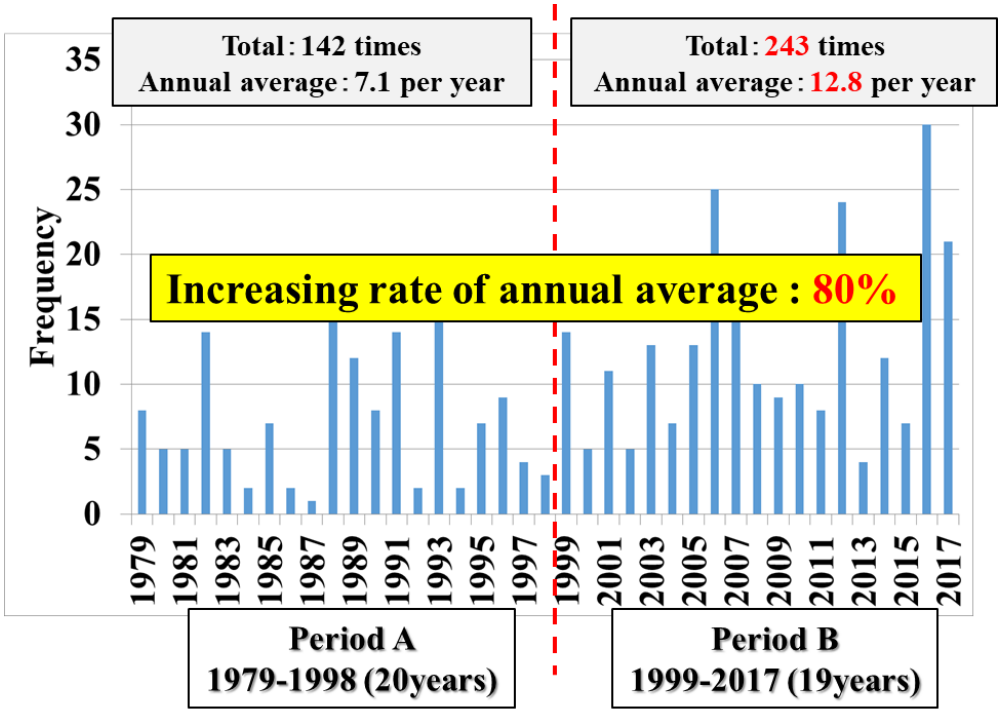


Figure 1.3: Annual variation in heavy rainfall frequency (≧70 mm/h) during a warm season over the last 39 years in Kyushu Island (Ohashi, 2018).

It is well known that the rainy season in the far east Asia occurs in early summer. The band-like maximum rainfall zone lying from east to west is located over South China in late spring. This maximum rainfall zone advances northwards gradually toward over the Japan Islands in the end of June and the beginning of July (Matsumoto, 1971). This west-to-east elongated rain zone formed in the rain season is called by the Meiyu-Baiu front (MBF). The strong meridional convergence of the moisture flux sustains the large precipitation in the MBF. The differential advection of the equivalent potential temperature generates convective instability against the stabilizing effect of the cumulus convection and sustains moist neutral stratification and intense precipitation (Ninomiya, 2007).

Under these circumstances in Japan, natural disasters such as floods and landslides have often happened during rainy season, which caused severe damages to infrastructures and human beings. In July 2017, there was a heavy rainfall event referred as Northern Kyushu Island rainfall, which caused landslide and debris flow over wide areas directly. A massive volume of sediment flowed to downstream areas, nearly covering the rivers. This raised the riverbed, causing flooding on a massive scale (Ministry of the Environment et al., 2018). Moreover, in August 2020, there was a heavy rainfall event in overall Kyushu area with maximum intensity of approximately 500 mm per day and 110 mm per hour in the particular area (JMA, 2020), which has inundated over 12,000 ha and caused flooding in approximately 200 rivers.

Focusing on the climate change in the world, the fifth synthesis report (IPCC, 2014) projected that extreme precipitation events will become more intense and frequent in many regions as global mean surface temperature increases. IPCC introduces the Representative Concentration Pathways (RCPs) which describes four different 21st century pathways of greenhouse gas to project the consequences for the climate system including the mean temperature change. They include a stringent mitigation scenario (RCP2.6), two intermediate scenarios (RCP4.5 and RCP6.0), and one scenario with very high GHG emissions (RCP8.5). Table 1.1 shows the projected change in global mean surface temperature from four different emission scenarios.

Table 1.1: Projected change in global mean surface temperature, relative to the 1986-2005 period (IPCC,2014).

		2046–2065		2081–2100	
	Scenario	Mean	*Likely* range [c]	Mean	*Likely* range [c]
Global Mean Surface Temperature Change (°C) [a]	RCP2.6	1.0	0.4 to 1.6	1.0	0.3 to 1.7
	RCP4.5	1.4	0.9 to 2.0	1.8	1.1 to 2.6
	RCP6.0	1.3	0.8 to 1.8	2.2	1.4 to 3.1
	RCP8.5	2.0	1.4 to 2.6	3.7	2.6 to 4.8

Despite some recent progress on bending the emission curve, RCP8.5 which represents the most aggressive scenario in assumed fossil fuel use for global climate models is the best match out to midcentury

under current and stated policies with still highly plausible levels of CO2 emissions in 2100 (Schwalm et al., 2020). Under the RCP8.5 scenario, Figure1.4 describes the projected change in the annual number of events with precipitation of over 50 mm per hour and it indicates a statistically significant increase of events with extreme precipitation for the end of the 21st century for all areas and seasons. Western Japan and Okinawa and Amami area are projected to have intensive rainfall events compared to other areas (Ministry of the Environment et al., 2018).

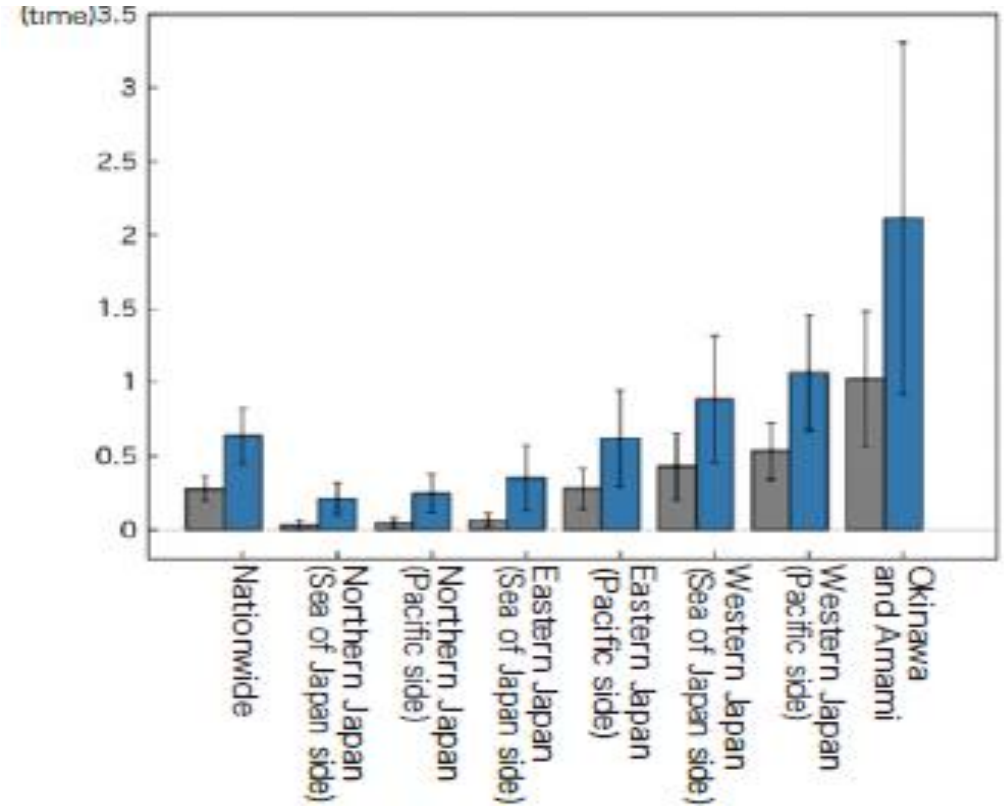


Figure 1.4: Projected change in the annual number of events with precipitation of 50 mm/hour or more. The bars indicate the number of occurrences per site (gray: present; blue: future). Thin black lines indicate the standard deviation of interannual variability (Ministry of the Environment et al., 2018).

1.2 Application of the SOM to synoptic meteorology (Study Review)

The Self-Organizing Map (SOM) is a kind of artificial neural network (ANN) that uses unsupervised learning to build a two-dimensional map of a problem space, which can generate a visual representation of data on a hexagonal or rectangular grid. Several articles focus on providing a demonstration of the methodology and output about SOM technique, with the application to a particular research question as a secondary goal for several decades (Sheridan, 2011). The literature review was conducted by Sheridan and synoptic climatological application has seen increased attention. For instance, the SOM was used to identify the primary features of the synoptic-scale circulation over the region (Hewitson and Crane, 2002; Cassano et al., 2006).

Furthermore, when applying the SOM technique, meteorological variables such as temperature, wind speed, atmospheric pressure, rain and humidity can be utilized and they are classified so well in the clustering process (Lopez et al., 2011). In these variables, temperature, wind and precipitation data has been used to assess extreme events which have the potential to cause torrential natural disaster (Cavazos, 1999; Nishiyama et al., 2007; Ohba et al., 2015, 2016; Yamada, 2018).

From these results, the SOM can be used to assess future meteorological changes in extremes (Paul et al.,2017) and the predictability of this SOM based approach is significantly improved than that from the conventional method using only the predicted rainfall intensity (Nguyen-Le et al., 2017).

Therefore, the SOM is available for synoptic meteorology which investigates the relationship between meteorological variables (e.g., precipitable water and wind) in a targeted area and features of meteorological fields detected in a synoptic scale (Tamaki, 2018; Ohashi ,2018) and risk prediction caused by heavy rainfall in advance(Nguyen-Le et al., 2019).

1.3 Objective

Based on the above-mentioned background, it can be very important to analyze current synoptic weather patterns in Kyushu area during rainy season and provide an early prediction and warning for heavy rainfall events in Kyushu area (Nguyen-Le et al., 2017).

The first objective is to reveal what kinds of meteorological fields that have highly contributed to extreme precipitation events, and natural disaster events related to heavy rainfall in Kyushu area so far. In this study, the Self-Organizing Map (SOM) (Kohonen, 1995) is employed to classify patterns of input meteorological data. Then, the Radar/Rain gauge analyzed precipitation (R/A) are combined to reveal synoptic weather patterns which have contributed to heavy rainfall events. Therefore, the SOM with heavy rainfall data can analyze torrential rainfall events happened in Kyushu area comprehensively and systematically.

Chapter 2: Clustering techniques

This chapter explains the background of the SOM and employed clustering techniques to decide the SOM structure. Among these techniques, Ward-method, K-means-method and Davies-Boulding Index (DBI) are commonly used in related researches and detailed formula are described in the previous work by Nishiyama et al. (2007). Then, the training process of the SOM is precisely discussed in this chapter. Moreover, the evaluation method of Quantization error (QE) and Topographic error (TE), and principle component analysis are also described in this chapter.

2.1 SOM

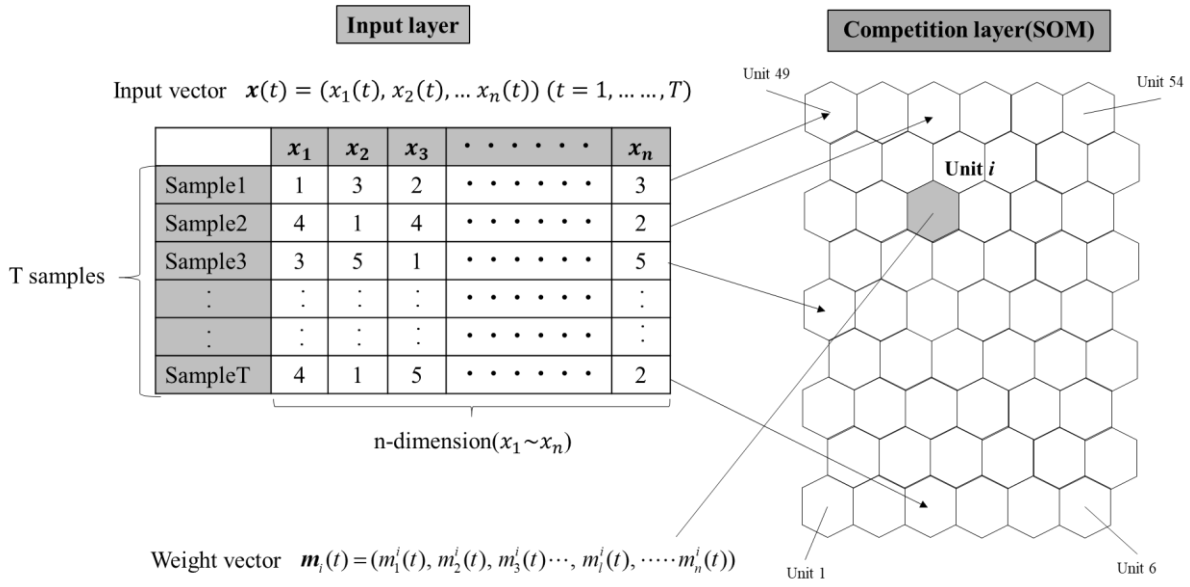
2.1.1 Basic information about the SOM

The Self-Organizing Map (SOM) was developed by Kohonen (1995) as an unsupervised learning algorithm to discover patterns of input data. It is currently used as a powerful clustering algorithm and it considers a neighborhood structure among the clusters, which provides a good visualization of complicated multidimensional data (Cottrell, 2018). In the unsupervised training process of the SOM, unlabeled multidimensional data is classified into two-dimensional map.

2.1.2 The SOM network structure

Figure 2.1 describes the conceptual model of the SOM network. This network is composed of an input layer and a competition layer. These input vectors are classified into the competition layer through the training process of SOM, which is explained in the latter chapter by using mathematical formula. Contrarily, the competition layer is composed of two-dimensional units, as is shown in Figure 2.1. Similar input vectors are classified into a specific unit characterized by a weight vector in the competition layer. This weight vector represents the average characteristics of input vectors allocated to the unit. In this process, the weight vector has the multi-dimension which is same as the input vector.

There are a few options about the shape of the units in a related research. This research mainly follows the method developed by Nishiyama et al. (2007) and it employs the regular hexagonal shape of units because the Euclidian distance between a unit and the nearest six neighboring unit is completely same on the whole map.



	x_1	x_2	x_3	$\cdots\cdots$	x_n
Sample1	1	3	2	$\cdots\cdots$	3
Sample2	4	1	4	$\cdots\cdots$	2
Sample3	3	5	1	$\cdots\cdots$	5
⋮	⋮	⋮	⋮	$\cdots\cdots$	⋮
⋮	⋮	⋮	⋮	$\cdots\cdots$	⋮
SampleT	4	1	5	$\cdots\cdots$	2

Figure 2.1: Conceptual model of the SOM

2.1.3 The SOM training process

A: Data preparation before the SOM training

Before the SOM training process, the normalization of input vectors and the initialization of weight vectors are executed as the first process. The mathematical formula of the input vector $\tilde{x}(t)$ and the normalized input vector $x(t)$ are shown in Eq. (2.1) and Eq. (2.2), respectively.

$$\tilde{\boldsymbol{x}}(t) = (\tilde{x}_1(t),\ \tilde{x}_2(t),\ \tilde{x}_3(t)\cdots,\ \tilde{x}_l(t),\ \cdots\cdots \tilde{x}_n(t)) \tag{2.1}$$

$$\boldsymbol{x}(t) = (x_1(t),\ x_2(t),\ x_3(t)\cdots,\ x_l(t),\ \cdots\cdots x_n(t)) \tag{2.2}$$

where subscript n is the number of elements(dimensions) in each input vector. *t* is the input vector's number from t=1 to t=T. T is the number of input vectors used for the training process as is shown in Figure 2.1. This normalized $x(t)$ is employed for this study after normalizing process in Eq. (2.3).

$$x_l(t) = \frac{\tilde{x}_l(t) - min\{\tilde{x}_l(t)\}}{max\{\tilde{x}_l(t)\} - min\{\tilde{x}_l(t)\}} \tag{2.3}$$

Subsequently, to determine the initial value of $m_i(0)$, random numbers between 0 and 1 are given to all elements of $m_i(0)$. The weight vectors $m_i(t)$ shown in Eq. (2.4) are updated unless the training process is continued. Then, the final weight vectors M_i are given in Eq. (2.5).

$$\boldsymbol{m}_i(t) = (m_1^i(t),\ m_2^i(t),\ m_3^i(t)\cdots,\ m_l^i(t),\ \cdots\cdots m_n^i(t)) \tag{2.4}$$

$$\boldsymbol{M}_i = (M_1^i,\ M_2^i,\ M_3^i\cdots, M_l^i,\ \cdots\cdots M_n^i) \tag{2.5}$$

where superscript i is the unit's number on the SOM.

B: Training process

Figure 2.2 shows the flow chart of the data preparation and training process of the SOM. After completing the preparation process, the first step is to search for the Best Matching Unit (BMU). The BMU is the weight vector which is closest to the input vector. After calculating the Euclidean distance between the input vector $x(t)$ and all weight vectors $m_i(t)$ on the SOM, the BMU is decided by following Eq. (2.6). (i.e.) the weight vector showing the minimum Euclidean distance value of $\|x(t) - m_i(t)\|$ is computed and it is selected as BMU.

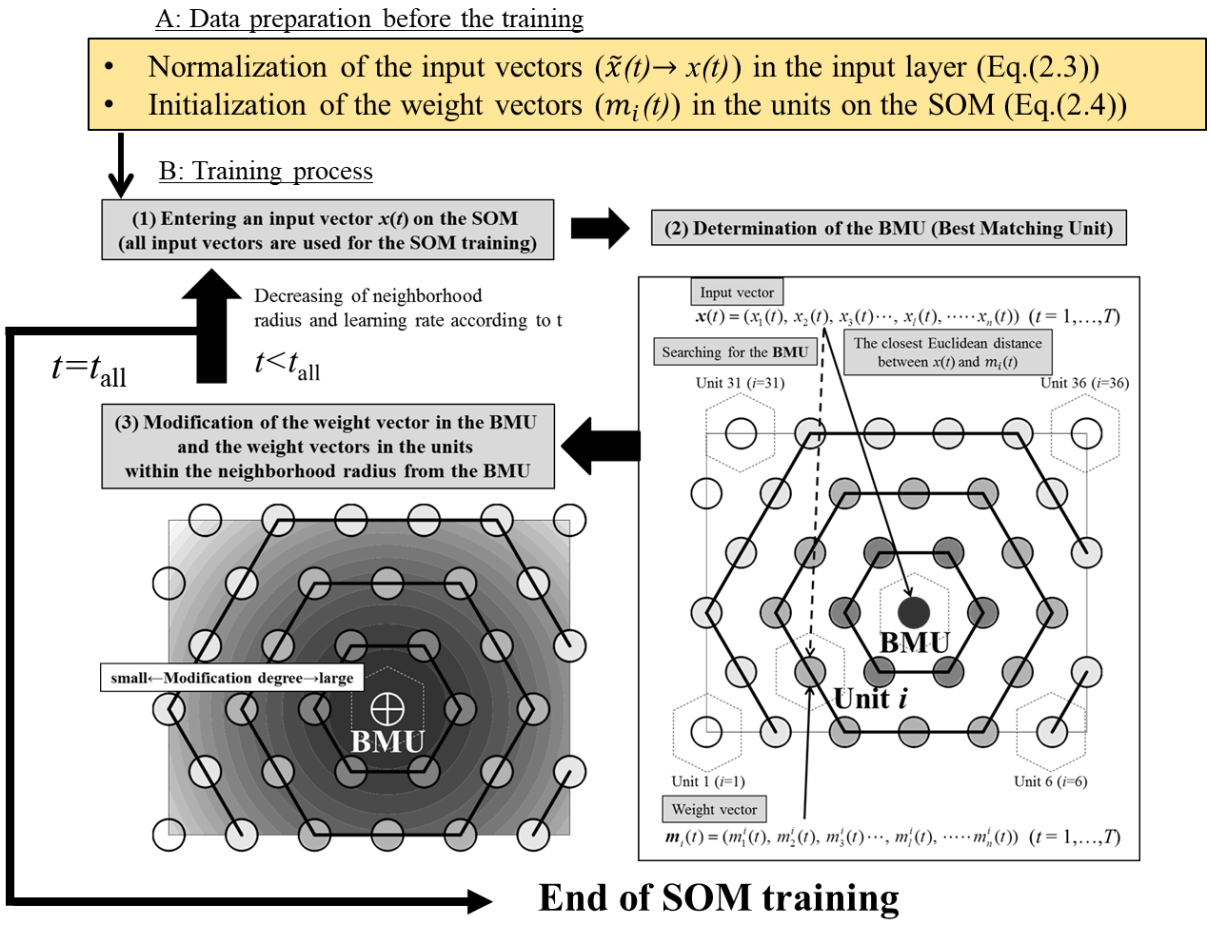


Figure 2.2: The flow of the data preparation and training process of the SOM

$$c = arg\ \min_i \left\{ \left\| \boldsymbol{x}(t) - \boldsymbol{m}_i(t) \right\| \right\} \tag{2.6}$$

where c is the BMU. In the next step, the weight vectors in each unit are modified and respectively updated by repeating the training process as is shown in Eq. (2.7).

$$\boldsymbol{m_i}(t+1) = \boldsymbol{m_i}(t) + h_{ci}\left(t_e, \left\| \boldsymbol{r_c} - \boldsymbol{r_i} \right\|\right)\left[\boldsymbol{x}(t) - \boldsymbol{m_i}(t)\right] \tag{2.7}$$

where t is the learning step. t_e is the epoch number and T_e is the total epoch number. 1 epoch is defined as the one learning of all input vectors from t=1 to t=T. The neighboring function expressed as $h_{ci}(t_e)$ shown in Eq. (2.8) is fixed until the training during one epoch has been completed. This means the degree of modification of the weight vectors is constant through each epoch. For instance, if 1000 input vectors are trained by the SOM with the maximum number of $T_{e=}$ 10, $t_{e=}$1 corresponds to the learning from t=1 to t=1000, $t_{e=}$2 corresponds to the learning from t=1001 to t=2000. The learning process will be end after finishing $T_{e=}$ 10 (from t=9001 to t=10000). In other words, the total number of the learning process is 10000 with the maximum number of $T_{e=}$ 10.

$$\begin{aligned} h_{ci}\left(t_e, \left\| \boldsymbol{r_c} - \boldsymbol{r_i} \right\|\right) &= \alpha(t_e) \cdot exp\left(- \frac{\left\| \boldsymbol{r_c} - \boldsymbol{r_i} \right\|^2}{2\sigma^2(t_e)} \right) \quad \left(\left\| \boldsymbol{r_c} - \boldsymbol{r_i} \right\| \le \sigma(t_e) \right) \\ h_{ci}\left(t_e, \left\| \boldsymbol{r_c} - \boldsymbol{r_i} \right\|\right) &= 0.0 \qquad\qquad\qquad\qquad \left(\left\| \boldsymbol{r_c} - \boldsymbol{r_i} \right\| > \sigma(t_e) \right) \end{aligned} \tag{2.8}$$

where $r_{c\ and}$ r_i means the location vectors of units c and i, respectively. $\alpha(t_e)$ is the learning rate and $\sigma(t_e)$ is the radius of neighborhood. The weight vectors located in the neighborhoods of the BMU within the radius of $\sigma(t_e)$ are modified by following the Eq. (2.8), whereas, the weight vectors located outside of that of $\sigma(t_e)$ are not modified. $\alpha(t_e)$ and $\sigma(t_e)$ in this study follows the method employed by Nishiyama et al. (2007) and defined as Eq. (2.9).

$$\begin{aligned} \alpha(t_e) &= max\left\{ \alpha(0)\frac{T_e - t_e + 1}{T_e},\ \ 0.005 \right\} \\ \sigma(t_e) &= max\left\{ \sigma(0)\frac{T_e - t_e + 1}{T_e},\ \ 1.1 \right\} \end{aligned} \tag{2.9}$$

As is seen in Eq. (2.9), when the learning rate and neighboring radius fall below 0.005 and 1.1, these values can be settled as 0.005 and 1.1, respectively. In other words, the degree of the modification of the weight vectors gradually decreased as the learning step continues and it becomes fixed in the particular point. Subsequently, the range of the modification becomes smaller same as the weight vectors do. After these values fixed, only the weight vector in the BMU and the weight vectors in the closest neighbor units of

the BMU are modified. The initial learning rate, initial neighboring radius and the maximum epoch number are set as 0.2, 5 and 50, respectively.

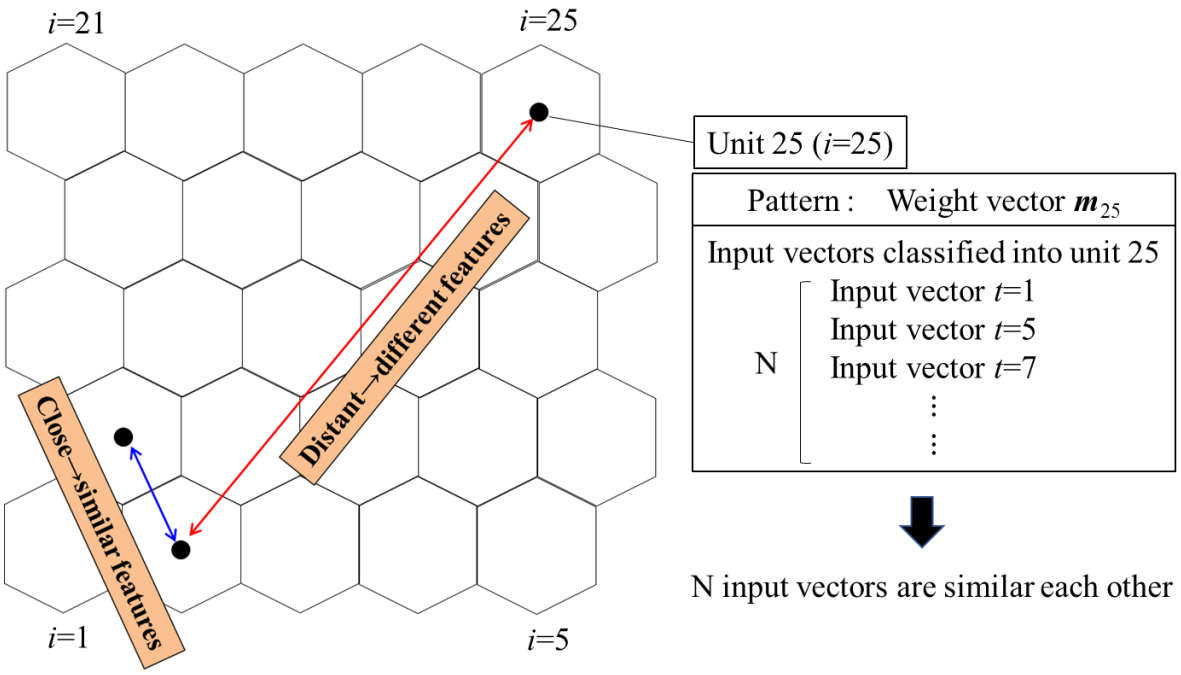


Figure 2.3: The structure of the SOM after finishing the training process

Figure 2.3 shows how to interpret the result of SOM after finishing the training process. Regular hexagonal shape of unit can be interpreted as a pattern characterized by a weight vector. Similar input vectors are classified in each unit and the weight vector shows typical characteristic of these similar input vectors. Moreover, the neighboring units are interpreted as having similar features. On the contrary, distant units are interpreted as having different features. Therefore, the similarity and dissimilarity among the input data or classified patters can be easily and visually understood on the two-dimensional map.

2.2 QE (Quantization error) and TE (Topographic error)

To ensure the Self-Organizing Map (SOM) to be an accurate model, it should be preserved the topology and neighborhoods of the input data while also fitting the data (Kivilnoto, 1996). The quality measures chosen for this study are quantization error (QE) and Topographic error (TE).

Quantization error (QE) is a measure of the average distance between the data points and the map units (BMU) to which they are mapped, as shown in Eq. (2.10). The smaller value indicates a better fit to the map. The value reported is on the same scale as the input data.

$$\mathrm{QE} = \frac{1}{N}\sum_{i=1}^{N}\left\|x_i - m_{BMU(i)}\right\| \tag{2.10}$$

where N is the total number of input vectors used for training. x_i is the input vector (i). $m_{BMU(i)}$ is the weight vector of the BMU for the input vector x_i.

Topographic error (TE) is a measure of how well the structure of the input space is modeled by the map. TE is calculated by finding the best-matching and second best-matching unit on the map for each input and evaluating the positions. If the units are adjacent to each other, the topology is preserved and $\mathrm{u}(x_i)$ is 0, whereas, units are not adjacent to each other, $\mathrm{u}(x_i)$ is 1, as shown in Eq. (2.11). This measurement only evaluates how well individual data points are mapped to the units and dose not account for larger distortions.

$$\mathrm{TE} = \frac{1}{N}\sum_{I=1}^{N} u(x_i) \tag{2.11}$$

Generally, the value of QE and TE decreases as the SOM training proceeds and it shows plausible performance in terms of the classification. Moreover, larger map size leads to higher TE value, whereas, QE becomes lower. Therefore, it should be needed to choose the optimal size by evaluating the combination of QE and TE.

2.3 Principle Component Analysis

The number of components to retain is a potentially thorny issue. Principle Component Analysis (PCA) has been utilized for studies of variability in meteorology and climatology for decades. Standard PCA creates a set of new uncorrelated variables that most closely and efficiently represent the variable in a dataset. The new variables provide a simpler and more easily interpretable version of the original data. (i.e.) It reduces the dimensionality of the data and only few components are significant in general (Reusch, 2005).

Some techniques have been suggested for making the important decision to extract insignificant components such as the scree test of Cattell (1966), the eigenvalue-one method of Rummel (1970), and the eigenvalue degeneracy test of North et al. (1982).

Chapter 3: Methodology

3.1 Overview of methodology in this study

Figure 3.1 illustrates an overview of the methodology employed in this study. This flowchart shows how to make the SOM structure and combine the frequency of heavy rainfall obtained by Radar/Rain gauge analyzed precipitation (R/A) in chapter 4, and how to relate the predicted meteorological data obtained by Global Spectral Model (GSM) to the synoptic weather patterns in chapter 5.

In Chapter 4, many meteorological data observed for four decades by JRA-55 project are classified into preliminary synoptic weather patterns in accordance with the similarity of the meteorological fields by following the heuristic rule. Then, the structure of the SOM is obtained after employing two clustering techniques such as Ward method and K-means method to make several groups. Then, the Davies-Bouldin Index (DBI) is employed to evaluate the clustering performance, which decides the best sorted SOM. Finally, the SOM is combined with the precipitation data given by R/A data, which shows the heavy rainfall frequency and possibility leading to heavy rainfall in each unit. The methodology of these processes is precisely discussed in Chapter 3.4.

In Chapter 5, predicted meteorological data by GSM is modified in accordance with the horizontal resolution of JRA-55 and this modified data is diagnosed by the structured SOM. As a result, specific number of unit and group is allocated to the predicted data and they show the probability leading to heavy rainfall, which can be used for decision makers for early warning of heavy rainfall events.

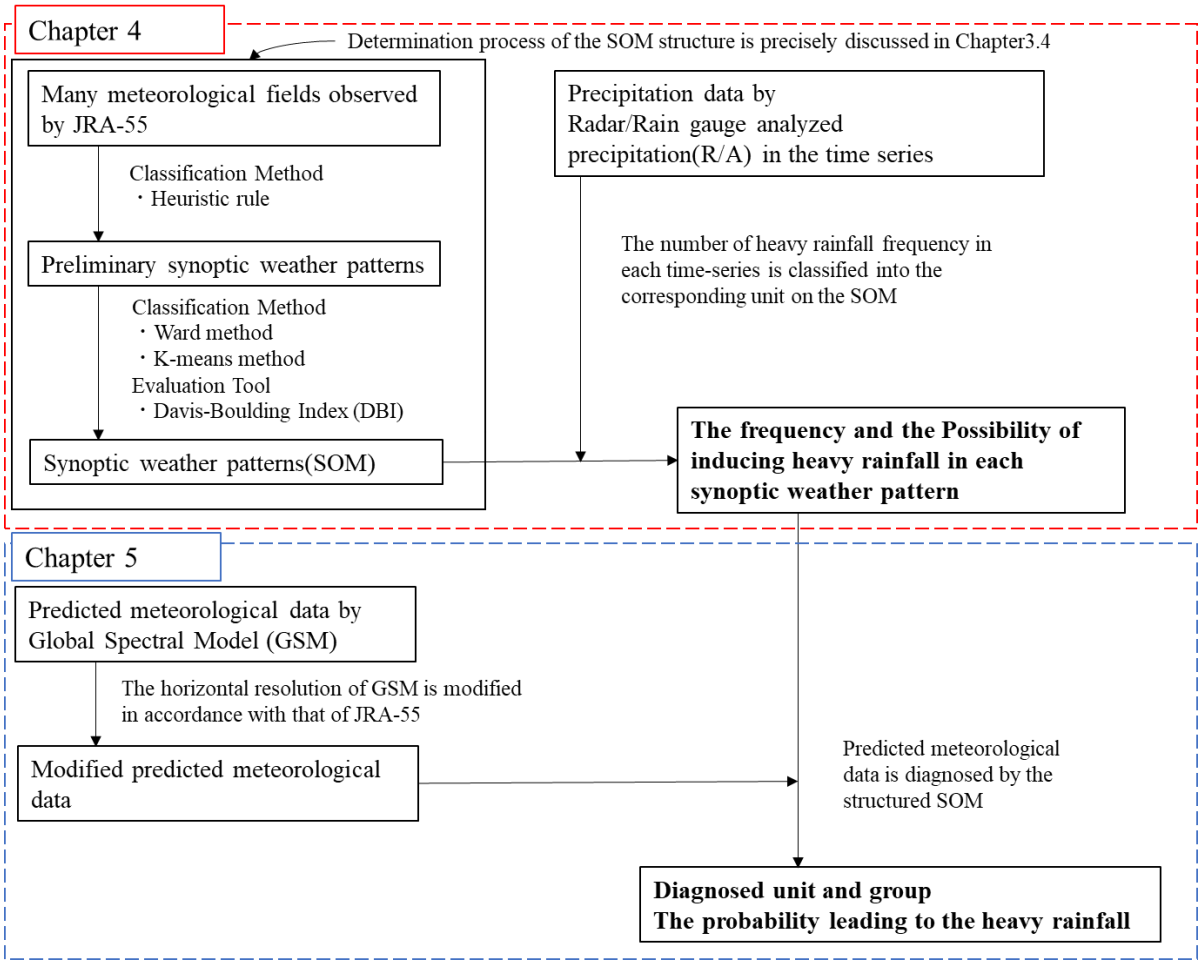


Figure 3.1: Overview of this study

3.2 Specification of metrological fields

Before the SOM training process, data pre-processing of meteorological fields is required to make an input vector for the SOM. This study follows the method designed by Nishiyama et al. (2007) and Ohashi (2018), which interprets the main factors causing heavy rainfall such as intrusion of moisture, lower level jet and rotation of a lower pressure system by employing precipitable water (PW) and wind components at 850 hPa. Then, according to these researches, meteorological fields used for input vectors are represented by three components; (1) PW (unit: mm), (2) U wind (East-West wind at the 850 hPa level)(unit: m/s) and (3) V wind (South-North wind at the 850 hPa level) (unit: m/s).

Meteorological data used for this study is obtained from the Japanese 55-year Reanalysis (JRA-55) (Kobayashi et al., 2015). The accuracy of meteorological data by JRA-55 is better than the NCEP/NCAR data employed by above-mentioned researchers (Wang et al., 2020). JRA-55 is the project involved comprehensive global atmospheric reanalysis based on four-dimensional variation analysis for the last half of the 20th century. (1958 onward) Moreover, this data is subject to produce a comprehensive atmospheric dataset suitable for studying multidecadal variability and climate change. The horizontal resolution of

meteorological fields is 1.25 ° and the data is recorded four times per day (3, 9, 15 and 21 JST: JST is UTC+9 h), every 6 h.

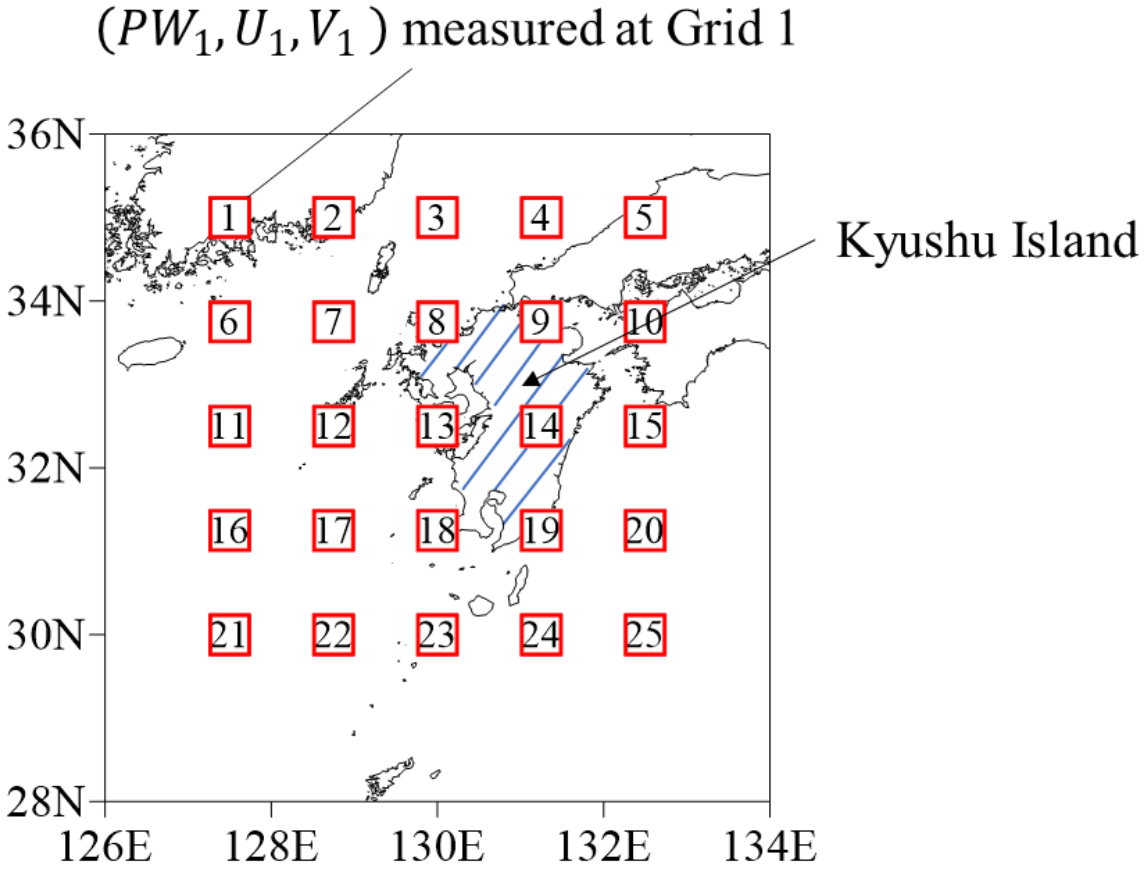


Figure 3.2: Meteorological field observed at 25 grid points

In this study, meteorological fields (PW, U and V) observed at 25 grid points with the extent of longitudes (127.5° to 132.5°) east and latitudes (30° to 35°) north during warm seasons (June to September) from 1979 to 2018, as shown in Figure 3.2. Then, an input vector x for representing meteorological fields is composed of 75 dimensions (25 grid points, 3 dimensions), as expressed by Eq. (3.1). Therefore, 19520 meteorological fields for 40 years are used for the SOM training process.

$$x = (PW_1, \dots, PW_{25}, U_1, \dots, U_{25}, V_1, \dots, V_{25}) \tag{3.1}$$

3.3 Specification of Radar/Rain gauge analyzed precipitation (R/A)

To acquire the actual meteorological data, JMA started to operate Automated Meteorological Data Acquisition System (AMeDAS) from 1976. This data is composed of rainfall, wind, temperature, and sunshine duration and so on, with a mean resolution of about 17 km. This data has been widely used to understand the frequency of heavy rainfall (Nishiyama et al., 2007). However, it has the limitation of analyzing the heavy rainfall frequency quantitatively due to the coarse horizontal resolution.

The Radar/Rain gauge analyzed precipitation (R/A) data represents accumulated rainfall amounts per an hour estimated from observations of radars and rain gauges. The data covers all of Japan and its surroundings. At first, the spatial resolution is approximately 5 km × 5 km (Oki et al, 1997). However, its resolution is downscaled from 5 km to 1 km in 2006. Then, the R/A data can be the highly downscaled dataset of precipitation and useful to understand the total frequency of heavy rainfall such as over 50mm per hour quantitatively. Therefore, this research applies the R/A data instead of the AMeDAS data employed in the previous study. To ensure the consistency of R/A data, input data is taken from 2006 to 2018. Moreover, the period is limited to warm seasons (June to September) when the heavy rainfall events are mainly observed.

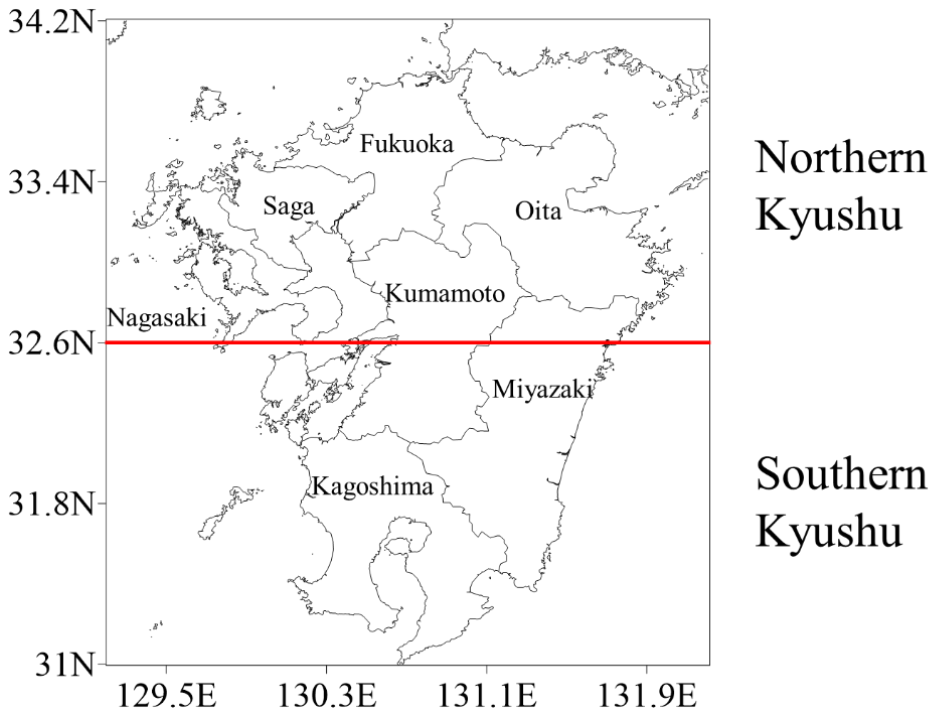


Figure 3.3: Targeted area for this study which shows each prefecture name

The target area is between 31° and 34. 2° north latitude, and 129.2° and 132.2° east longitude to cover almost all of Kyushu area. As shown in Figure 3.3, the area is divided into the northern (32.6° to 34.2° north latitude) and the southern (31° to 32.6° north latitude) area depicted in red line to understand local characteristics of heavy rainfall in terms of the frequency distribution and the synoptic pattern.

3.4 Determination of the SOM structure

Recent studies have been carried out in several different ways when deciding the SOM structure, however, the suitable way of producing process is not established. Therefore, this study employs some kinds of methods and extracts the result of them. After analyzing all of them, it is tried to choose the plausible way for making the SOM structure because the number of map nodes determines the accuracy and generalization capability. The overview of this determination process is shown in Figure 3.4. Generally, the bigger map

gives the higher resolution for pattern recognition; however, the topographical adjacency is further among the clusters. Therefore, it is important to choose the appropriate size of the SOM.

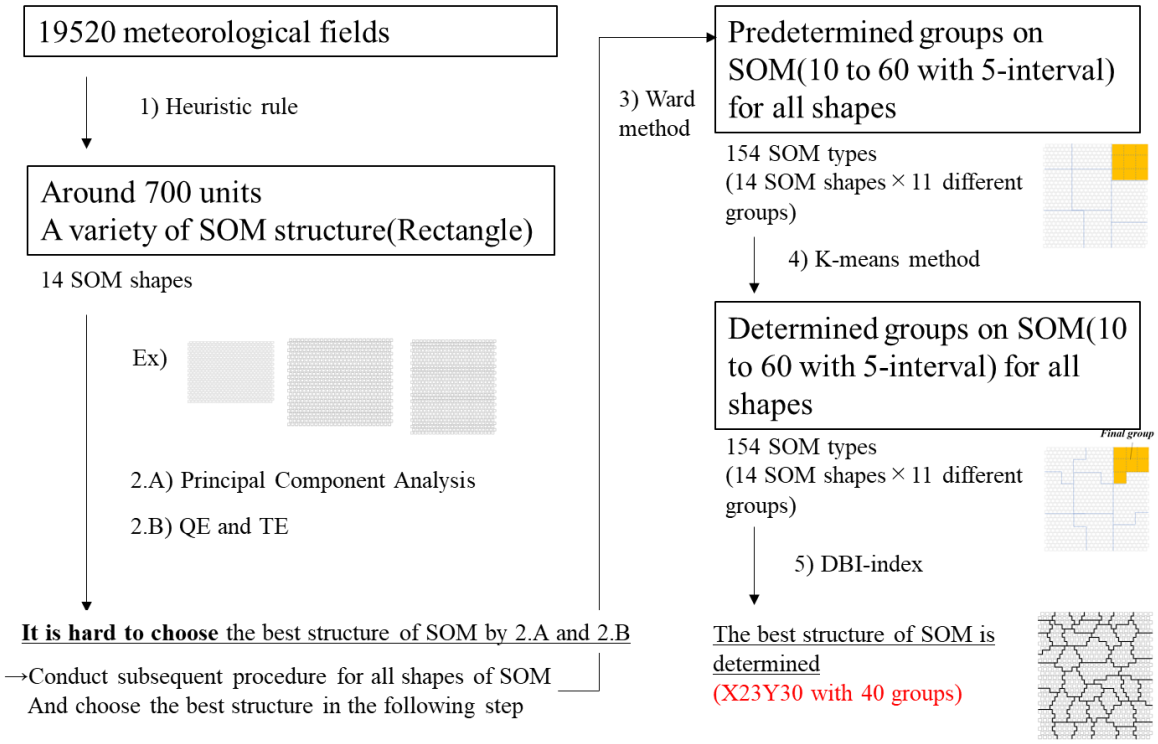


Figure 3.4: Overview of methodology to define the structure of SOM

1) Heuristic rule

In the first process, the number of SOM units are determined by the heuristic rule as presented in Eq. (3.2). This initial method has been used in these recent researches (Nguyen et al., 2015). Then, the optimal number of map unit is provided as 699 given by the 19520 input meteorological data.

$$M \approx 5\sqrt{n} \tag{3.2}$$

with M denoting the number of the SOM units and n representing the number of input data

There are several formulas to determine the number of rows, and the ratio of width(x) and length(y) of the SOM such as employing eigenvalues of the transformed data which is one of the principle component analysis (Hilario and Ivan, 2004). Moreover, the quantization error (QE), topographical error (TE), and the Davies-Bouldin index (DBI) (Nishiyama et al., 2007) are precisely described in Chapter 2. Then, these above-mentioned methods are employed to determine the structure of the SOM.

2.A) Principal Component Analysis

In this analysis, the square root of the ratio between the biggest two eigenvalues of the transformed data was calculated to determine the structure of SOM. Table 3.1 shows the result which indicates the

contribution rate of the two main components. Moreover, the ratio of width (x) and length (y) of SOM can be obtained as presented in Eq. (3.3). Then, the obtained ratio between the width and length is 1.33, which means the width(x) and length(y) would be 23 and 30, respectively, by considering the number of total units (699) as is given in heuristic rule 1).

$$\frac{y}{x} = \sqrt{\frac{e_1}{e_2}} \tag{3.3}$$

with $e_{1,}e_2$ representing the eigenvalue of the first and second main components, respectively as shown in Table 3.1

Table 3.1: The statistic value of first and second main components among 75 components

	Cumulative Proportion	Standard deviation	Eigenvalue
The first main component	0.3847	5.3715	28.85301
The second main component	0.6015	4.0326	16.26186

However, these two main components explain only 60% of all phenomena and the other 40% of that can be ignored. It would be difficult to conclude the structure with the width(x) of 23 and the length(y) of 30 as the best structure among all of the SOM structures. Then, the other methods are employed to determine the best structure of SOM.

2.B) QE and TE

Before calculating quantization error (QE) and topographical error (TE), several SOM shapes which have different width(x) and length(y) are prepared based on the total unit number of 699. In this case, fourteen cases are selected as the candidates for the optimal SOM structure, as is shown in Table 3.2. Table 3.2 also shows that total number of units for each case is close to the optimal number of 699. For instance, the map size X20Y35 means the length of the width(x) and length(y) on the map is 20 and 35, respectively.

Table 3.2: The fifteen cases selected as the candidates for the SOM structure

Map size	QE	TE
X20Y35(700units)	0.0511	0.0461
X21Y33(693units)	0.0513	0.046
X22Y32(704units)	0.0513	0.0488
X23Y30(690units)	0.0514	0.0507
X24Y29(696units)	0.0513	0.0461
X25Y28(700units)	0.0513	0.0488
X26Y27(702units)	0.0512	0.0464
X27Y26(702units)	0.0512	0.0473
X28Y25(700units)	0.0511	0.0491
X29Y24(696units)	0.0512	0.0457
X30Y23(690units)	0.0512	0.0431
X32Y22(704units)	0.0512	0.0419
X33Y21(693units)	0.0513	0.0459
X35Y20(700units)	0.0512	0.0449

As a result, the difference of the QE and TE after completing the SOM training process between these candidates is negligible small, which indicates any case is plausible for the SOM structure. Therefore, a subsequent process for all candidates are conducted and tried to find the best structure of the SOM.

3,4,5) Ward method, K-means method and Davies-Bouldin Index (DBI)

Following the previous result, Ward method and subsequent K-means method are applied for these fourteen candidates in the clustering phase. Then, around 700 units are classified into several groups in proportion to the degree of similarity. The procedure of these methods is precisely shown in Figure 3.5. Finally, David-Bouldin Index (DBI) is utilized for evaluating the better sorted SOM in the evaluation phase.

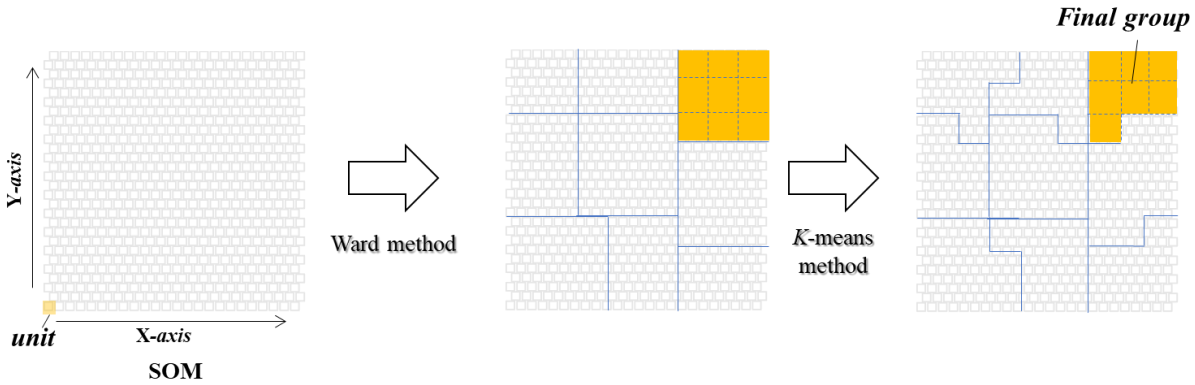


Figure 3.5: The procedure of Ward method and K-means method

At first, by employing Ward method, the clustering starts from the bottom-left unit and continues until all units on the SOM are classified into predetermined final groups (from 10 to 60 groups at 5-group interval). Then, K-means method is utilized for the rearrangement of the units located close to the boundary depicted by the blue line in Figure 3.5.

Finally, the DBI index is obtained after completing K-means method per a determined group (from 10 to 60 groups at 5-group interval) for fourteen different map sizes. Figure 3.6 shows the score of DBI for the number of groups on the SOM in the all fourteen cases as is mentioned in Table 3.2.

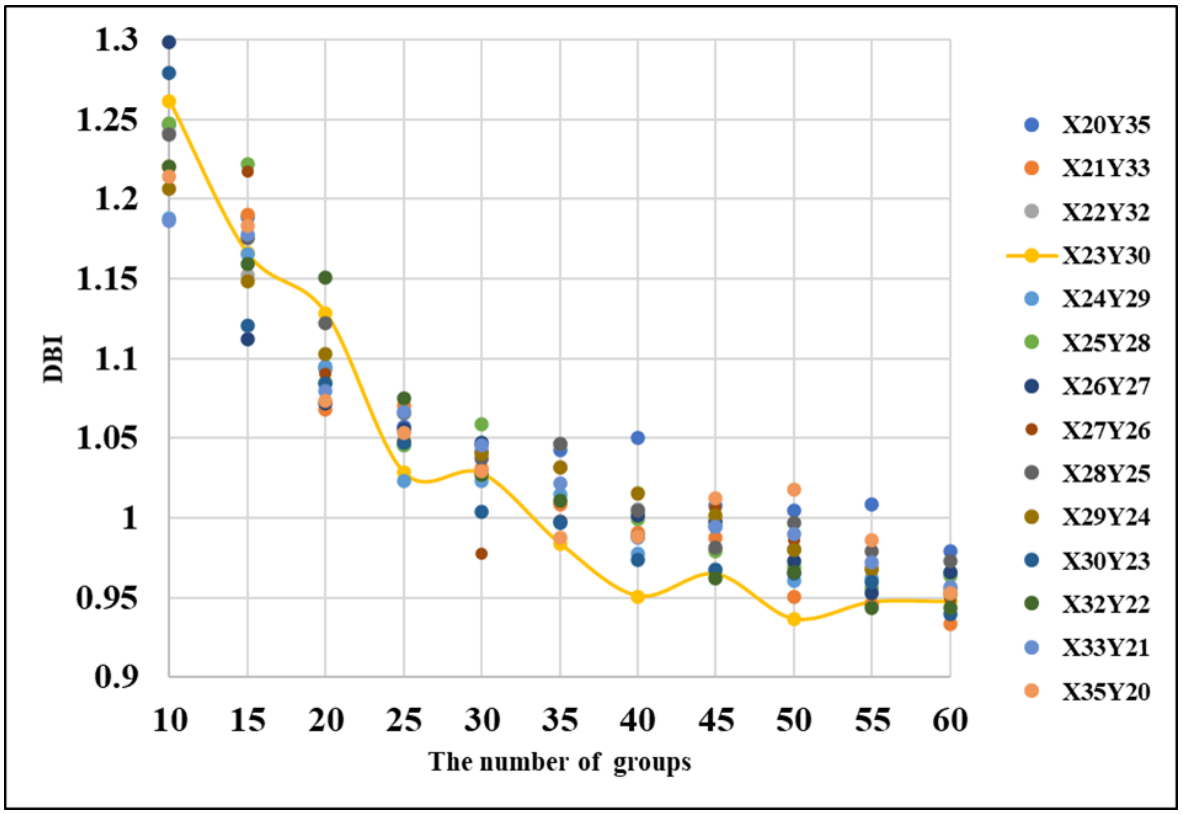


Figure 3.6: The score of David-Bouldin Index (DBI) on the number of groups in all SOM shape

According to Figure 3.6, the DBI decreases as the number of group increases, which means the performance of clustering is better in much more groups. Moreover, the score of DBI decreases sharply from 1.3 to 1.0 with an increase of groups between 10 and 40. On the contrary, the DBI after 40 groups shows a steady decrease (i.e.) the variability after 40 groups is relatively small. Then, it is plausible to choose the less classified groups with lowest DBI because it can be difficult to distinguish the characteristics of each group in 50 or more groups.

Following above-mentioned discussion, X23Y30 with 40 groups shows the lowest DBI compared to other SOM with same groups depicted in yellow dot in Figure 3.6. Hence, the SOM of X23Y30 with 40 groups is selected for this study.

3.5 The frequency and possibility of heavy rainfall corresponding to meteorological field

A meteorological field data used as an input vector of the SOM training is provided 4 times per day (3, 9, 15, 21 JST: JST is plus 9 h to GMT) (Kobayashi et al., 2015). However, hourly precipitation data is obtained every hour by the R/A system. To define the one event between meteorological data (JRA-55) and precipitation data (R/A), one day is divided into 4 time periods (0-6 JST, 6-12 JST, 12-18 JST, 18-24 JST). Then, the frequency of heavy rainfall over 30, 50, 70 mm per hour in Kyushu area are counted every 6 h to correspond to the defined time periods, respectively. The data of total heavy rainfall frequency in 1 event is stored on the SOM and summed up in each unit for 40 years (1979~2018).

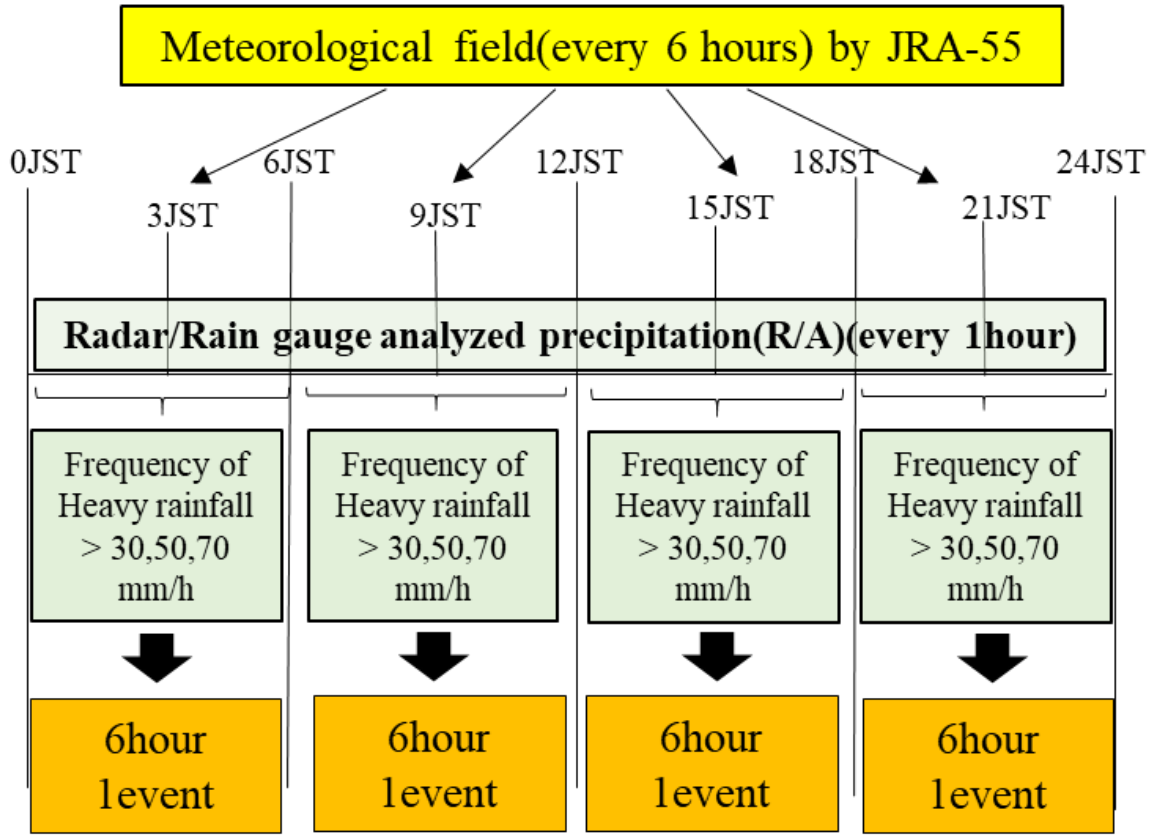


Figure 3.7: Relation between the meteorological fields obtained by JRA-55 and the total frequency of heavy rainfall in each time series

Moreover, the heavy rainfall probability of each unit can be calculated by counting the number of the past events (heavy rainfall events, non-heavy rainfall events during 2006-2018) in each unit classified by the SOM, as expressed by Eq. (3.4). If the intensity of heavy rainfall over 30, 50, 70 mm per hour is observed during each event at least one time, it is regarded as heavy rainfall event, respectively. In the result, the intensity over 50 mm/h is employed for analyzing the heavy rainfall events and other intensity are shown in the Appendix.

$$p = \frac{frequency(heavy\ rainfall)}{frequency(heavy\ rainfall+non\ heavy\ rainfall)\}} \quad (3.4)$$

3.6 Application of Global Spectral Model to SOM

In the diagnosis process discussed in chapter 5, the Global Spectral Model (GSM) is employed for obtaining the predicted meteorological fields in Kyushu area. The GSM is a high-resolution global atmospheric general circulation model with 20-km-mesh (Mizuta et al., 2006). This model predicts the meteorological data such as U, V wind and PW every 3 h for the next 84 h.

Owing to the difference of the horizontal resolution between GSM (20km each) and JRA-55 (1.25° each), the GSM data should be averaged to match the resolution of JRA-55, as shown in Figure 3.8. The meteorological factors at the observation points by GSM depicted in blue dots are averaged by following each JRA-55 grid depicted by a red frame.

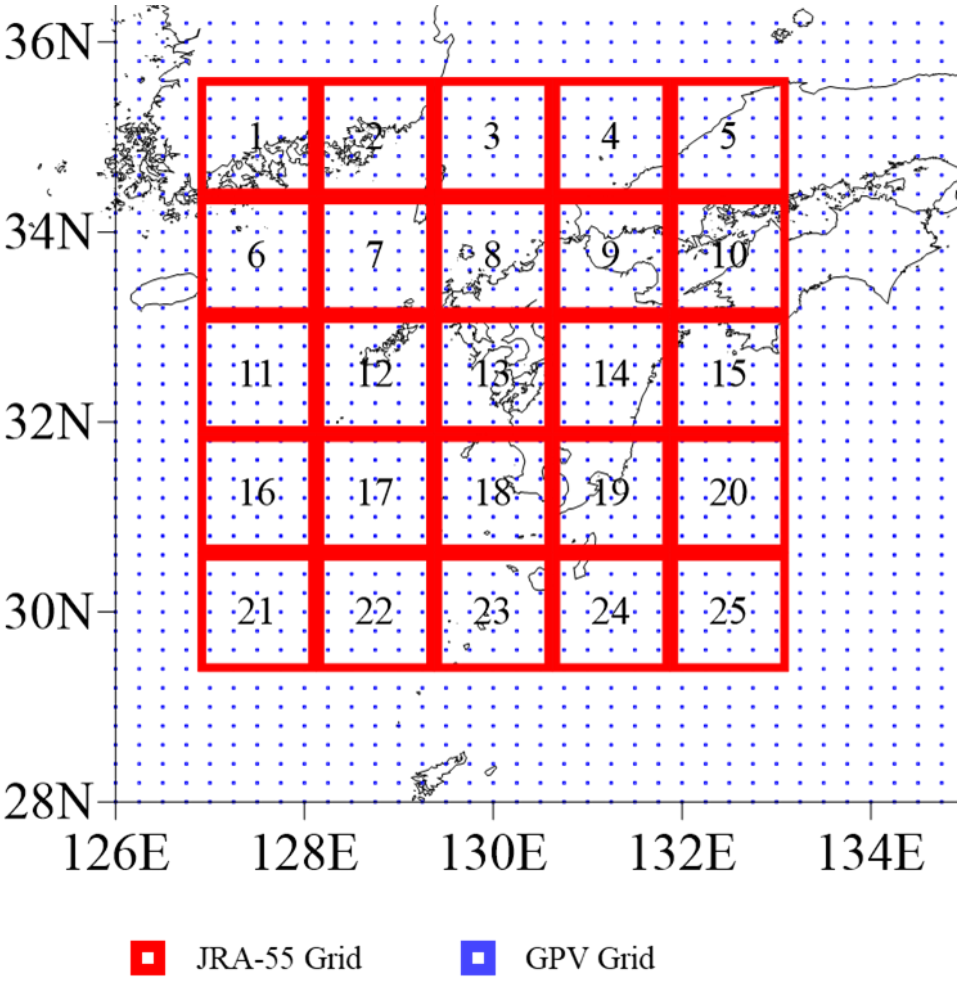


Figure 3.8: The illustration of Grid (JRA-55 and GPV)

The GSM data is diagnosed by the SOM structured in Chapter 4 and the specific unit number is allocated in accordance with the similarity of meteorological field on the SOM, as shown in Figure 3.9. It

should be noted that the weight vectors of the structured SOM do not change and predicted meteorological does not influence the structure of the SOM.

The scope of this research is to verify whether it is possible to diagnose heavy rainfall risk in advance. Then, recent heavy rainfall cases which caused torrential rainfall in Kyushu area after 2019 are employed for heavy rainfall diagnosis, which predicts the probability leading to heavy rainfall in each event, and extracts features of past heavy rainfall cases between 2006 and 2018.

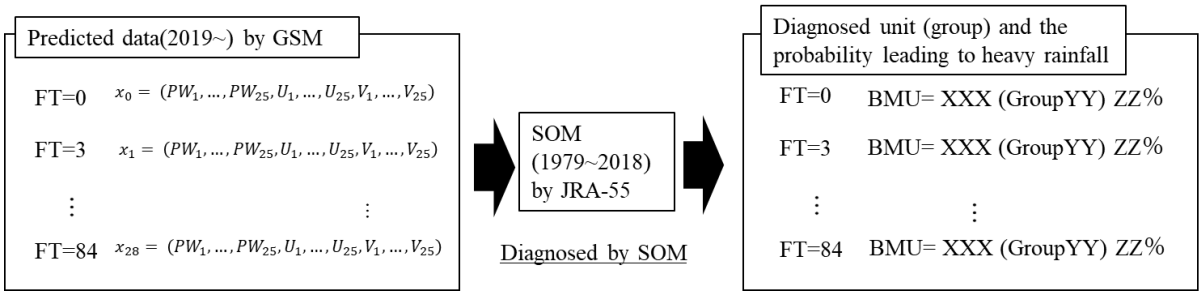


Figure 3.9: The flow of heavy rainfall diagnoses employing predicted data in each heavy rainfall event after 2019

The predicted data has a little variability due to the uncertainty of atmospheric circulation. (i.e.) It can be difficult to estimate the exact value of PW and wind factors. Then, the assigned unit on the SOM would be wrong and it influences the result of probability leading to heavy rainfall. Therefore, for the predicted data, the probability is calculated based on the assigned unit depicted in orange and its surrounding units depicted in yellow, as is shown in Figure 3.10. The formula is expressed by Eq. (3.5.). This aggregated possibility is shown as "p_ag" in Chapter 5 and it can be distinguished from the individual probability given by one unit in Chapter 4.

$$p\_ag = \frac{ag\_frequency(heavy\ rainfall)}{ag\_frequency(heavy\ rainfall+non\ heavy\ rainfall)\}} \tag{3.5}$$

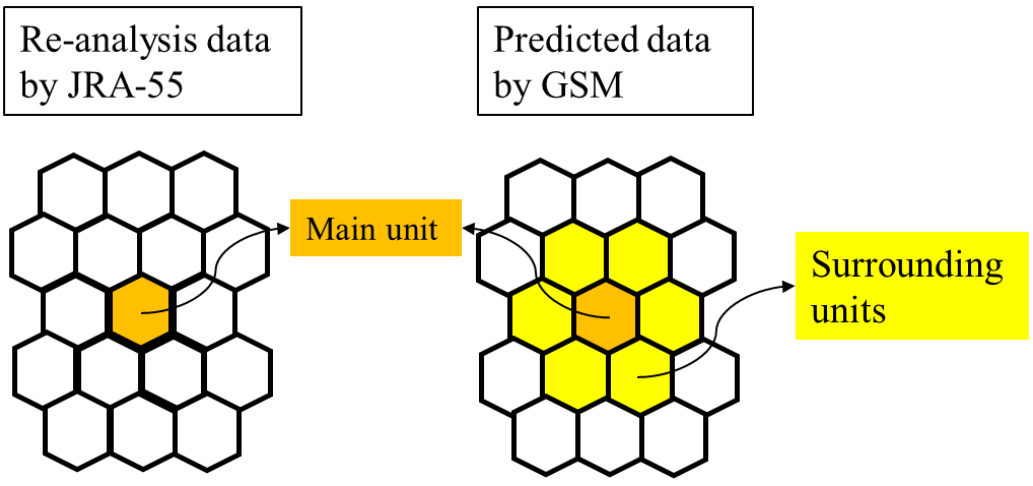


Figure 3.10: The possibility(p) by one unit and the aggregated possibility(p_ag) by seven units

Chapter 4: Result 1 (Classification of heavy rainfall patterns)

4.1 The SOM structure used for this study

According to the previous section, 19520 meteorological fields observed for 40 years were classified into around 700 synoptic weather patterns on two-dimensional map by following the heuristic rule. Then, these patterns were divided into groups by employing two clustering methods of Ward and K-means method. After the classification, they are evaluated by DBI and the SOM structure (X-axis=23, Y-axis=30) is decided in this study. The boundary of each group on the SOM is depicted in black bold line in Figure 4.1. Moreover, the meteorological fields of all synoptic weather patterns are shown in Appendix.

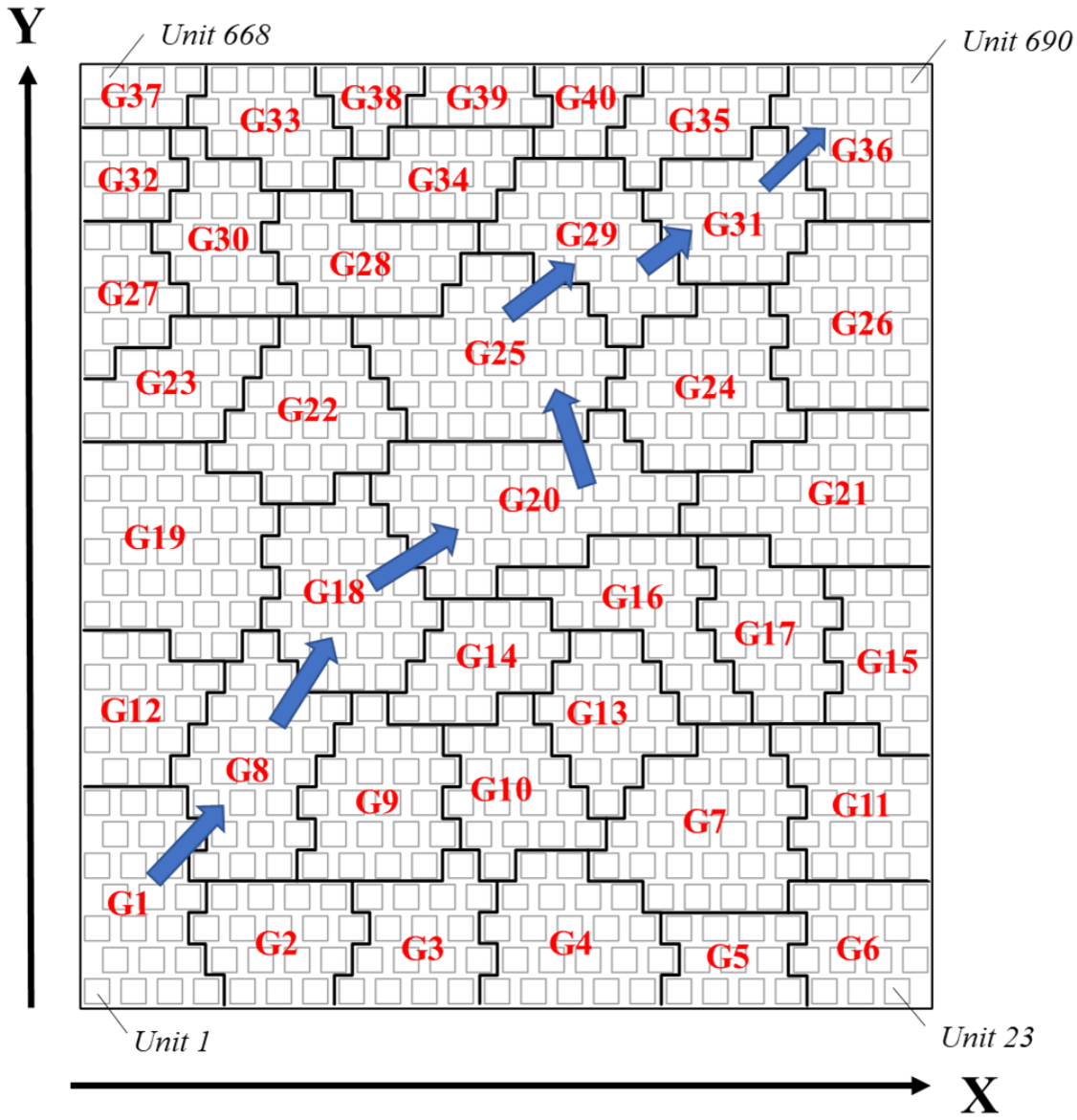


Figure 4.1: The units and groups specified on the SOM

As can be seen in Figure 4.2, the SOM has easily and visually comprehensive unique feature that neighboring groups on the SOM have the similar feature to each other and distant groups are dissimilar to each other. For instance, the neighboring groups such as G31 and G36 have similar features in terms of the distribution of high precipitable water and strong wind from the south-west. On the other hand, comparing distant groups, G01 and G36, there is a clear difference because G01 can be characterized by the intrusion of dry and weak wind from the north mostly in spite of G36 characterized by the humid and strong wind from the south-west. Therefore, both groups are quite opposites in each other. In the following section, the SOM can be used to identify what kind of groups are highly contributed to the frequency of heavy rainfall and natural disaster events related to heavy rainfall.

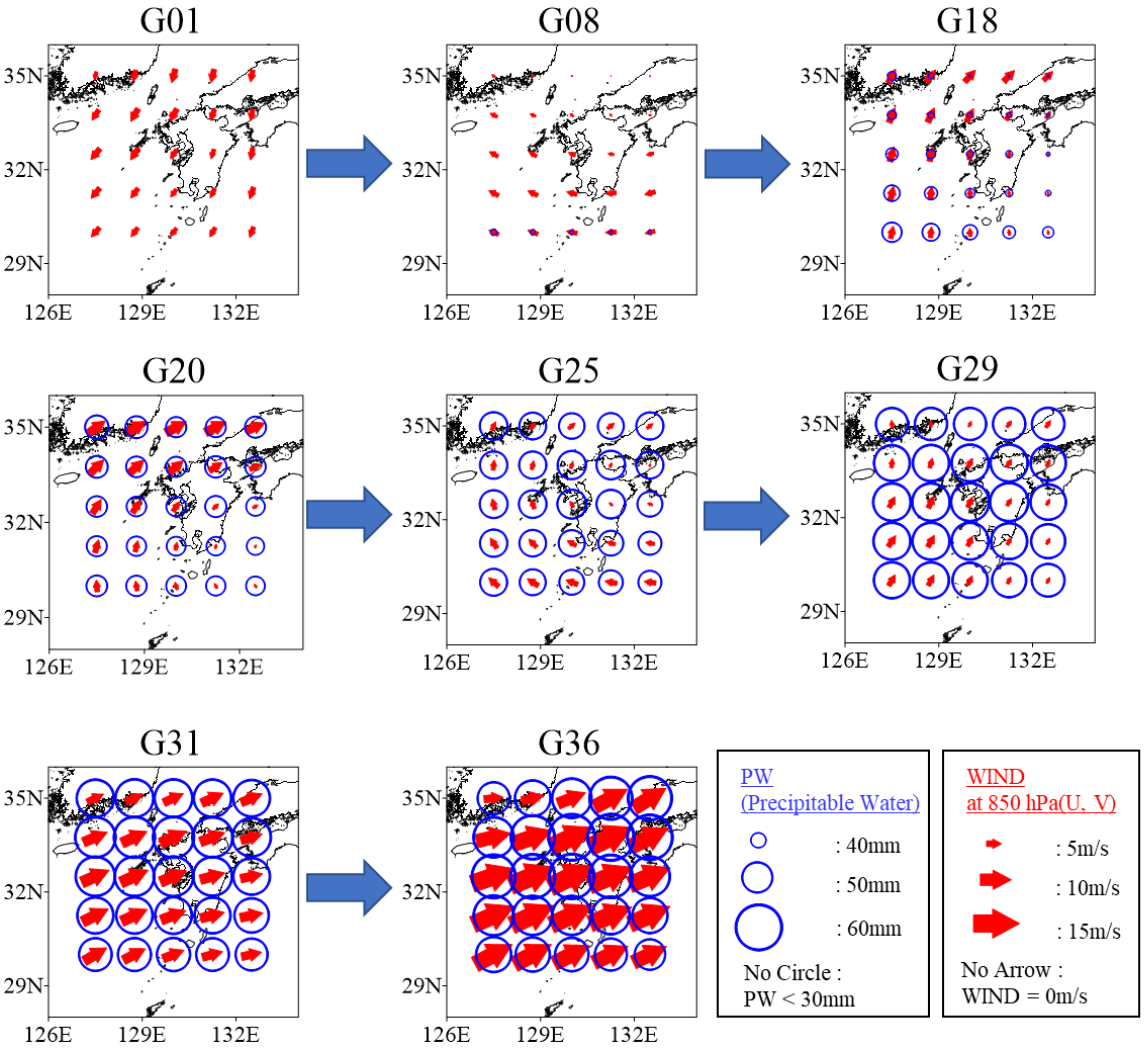


Figure 4.2: Transition of features of synoptic weather groups with distance from G01 to G36 depicted by the arrow in Figure 4.1

4.2 The SOM for heavy rainfall frequency and possibility

The heavy rainfall frequency measured by the R/A from 2006 to 2018 are related to the meteorological fields and the number of that are allocated to each unit on the SOM. The color of each unit on the SOM means the frequency (upper figure) and the possibility (lower figure) of having heavy rainfall events in Figure 4.3. The frequency (upper figure) can be used to evaluate what kind of groups and units are highly contributed to heavy rainfall events quantitatively. Specific number of the total frequency in each group is shown in Table 4.1. On the other hand, the possibility (lower figure) can be used to understand the risk of hazard in each unit on the SOM. Moreover, the target area is divided into the northern and the southern area in Kyushu to understand the local features of heavy rainfall events

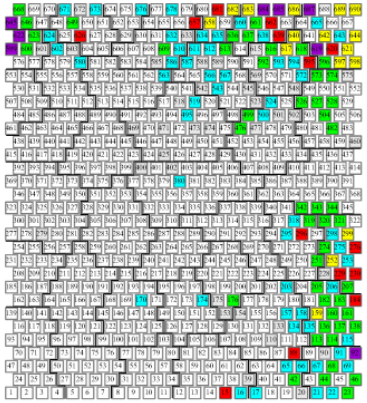
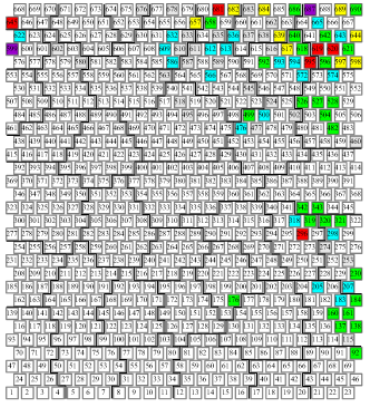
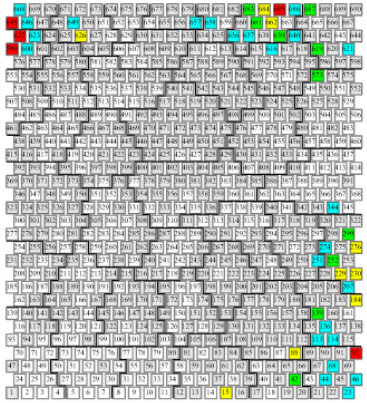
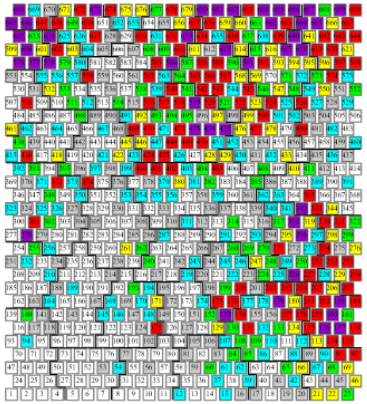
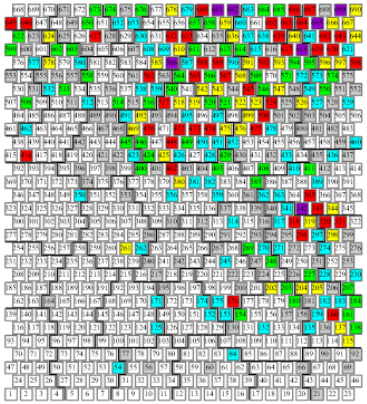
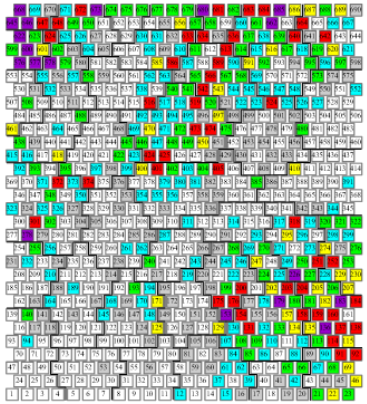
≥50mm/h	All area (N 31~34.2°)	Northern area (N 32.6~34.2°)	Southern area (N 31~32.6°)
Frequency(times) 300 to 500 500 to 1000 1000 to 3000 3000 to 5000 5000 to 10000 10000 to 300000			
Possibility(%) 10 to 20 20 to 30 30 to 40 40 to 50 50 to 70 70 to 101			

Figure 4.3: (Upper figure) The frequency causing heavy rainfall (≥50 mm per hour) for 40 years in each unit on the SOM (Lower figure) The possibility causing heavy rainfall (≥50 mm per hour) for 40 years in each unit on the SOM

4.3 The top 10 synoptic weather patterns which have caused heavy rainfall

The frequency of the precipitation with the intensity over 50 mm per hour in the top 10 synoptic weather groups explains 76.8% in all, 81.8% in the northern and 81.4% in the southern Kyushu of all heavy rainfall frequency among all weather groups, respectively. Moreover, in terms of the intensity over 80 mm per hour, top 10 groups represents 83.9% in all, 88.6% in the northern, 93.1% in the southern Kyushu area, respectively. Therefore, it can be assumed that these top 10 synoptic weather groups have frequently induced heavy rainfall events and it is significant to understand the characteristics of each group. Table 4.1 (right) shows the total frequency in the top 10 groups and it shows G36 explains approximately 20% of all heavy rainfall frequency.

Table 4.1: The top 10 synoptic weather patterns related to heavy rainfall frequency (over 50 mm per hour) in all, the northern and the southern Kyushu area, respectively (left), The total frequency in the top 10 synoptic weather patterns (right)

No	All area	Northern area	Southern area
1	G36	G36	G35
2	G35	G31	G15
3	G32	G32	G32
4	G15	G35	G6
5	G11	G40	G11
6	G6	G11	G37
7	G31	G26	G7
8	G37	G15	G36
9	G40	G17	G33
10	G26	G29	G4

No	Group	ALL area(times) R>=50mm/h
1	G36	66991
2	G35	57146
3	G32	45680
4	G15	38583
5	G11	32310
6	G6	25136
7	G31	24033
8	G37	22416
9	G40	17079
10	G26	12804
ALL		342178

Table 4.1 (left) shows the synoptic weather groups such as G11, G32 and G35 have caused heavy rainfall events in all Kyushu area. Focusing on the local characteristics, G31, G36 and G40 are typically shown in the northern area, whereas, G6, G15 and G37 are oriented in the southern area. Figure 4.4 shows the frequency distribution of heavy rainfall events and indicates these characteristics clearly.

Moreover, some groups show the frequency is relatively high in the western area and others have the opposite features, which can be affected by the topographical condition in Kyushu area, as is shown in Figure 4.5. It shows Kyushu mountain range extending from the northern to the southern area in the middle of this area. Then, if warm and humid air flows from west toward Kyushu area, the frequency of rainfall is oriented in the western Kyushu area, which is clearly seen in G6, G11, G15, G31, G35 and G36. However, if the wind blows from east, the frequency of rainfall is oriented in the eastern Kyushu area such as G32 and G37.

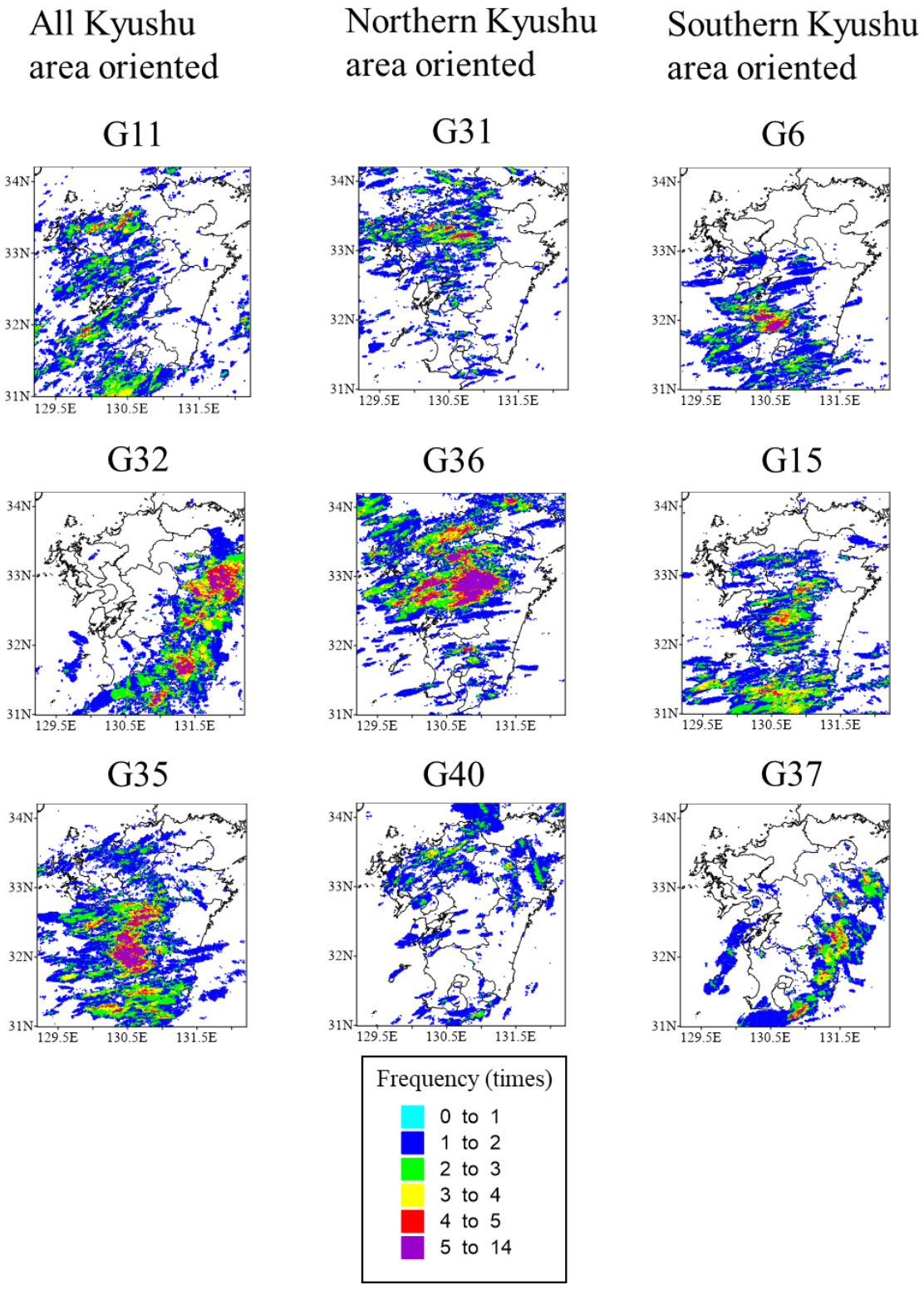


Figure 4.4: The frequency distribution map over 50 mm per hour in Kyushu area

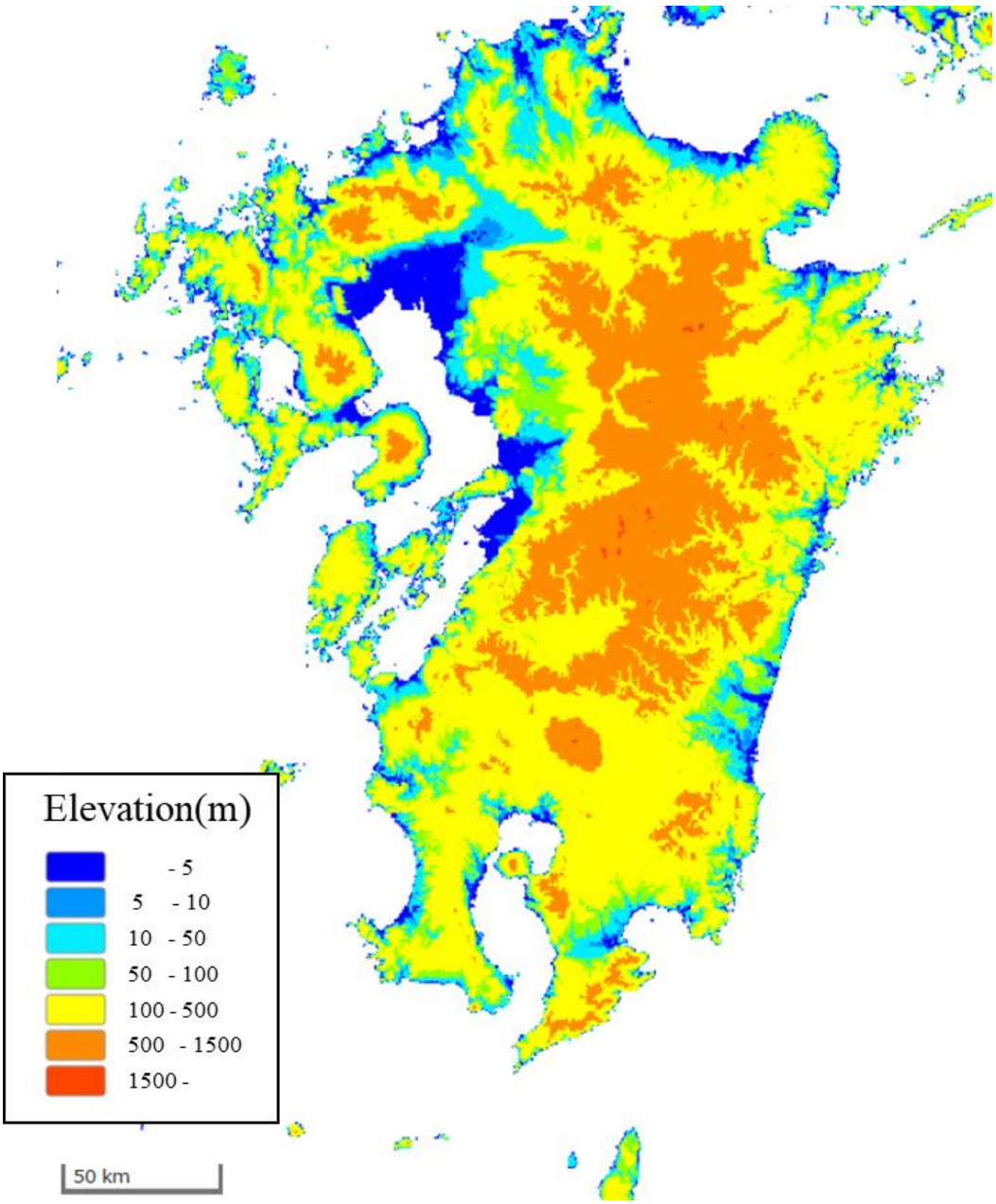


Figure 4.5: The topographical condition of Kyushu area

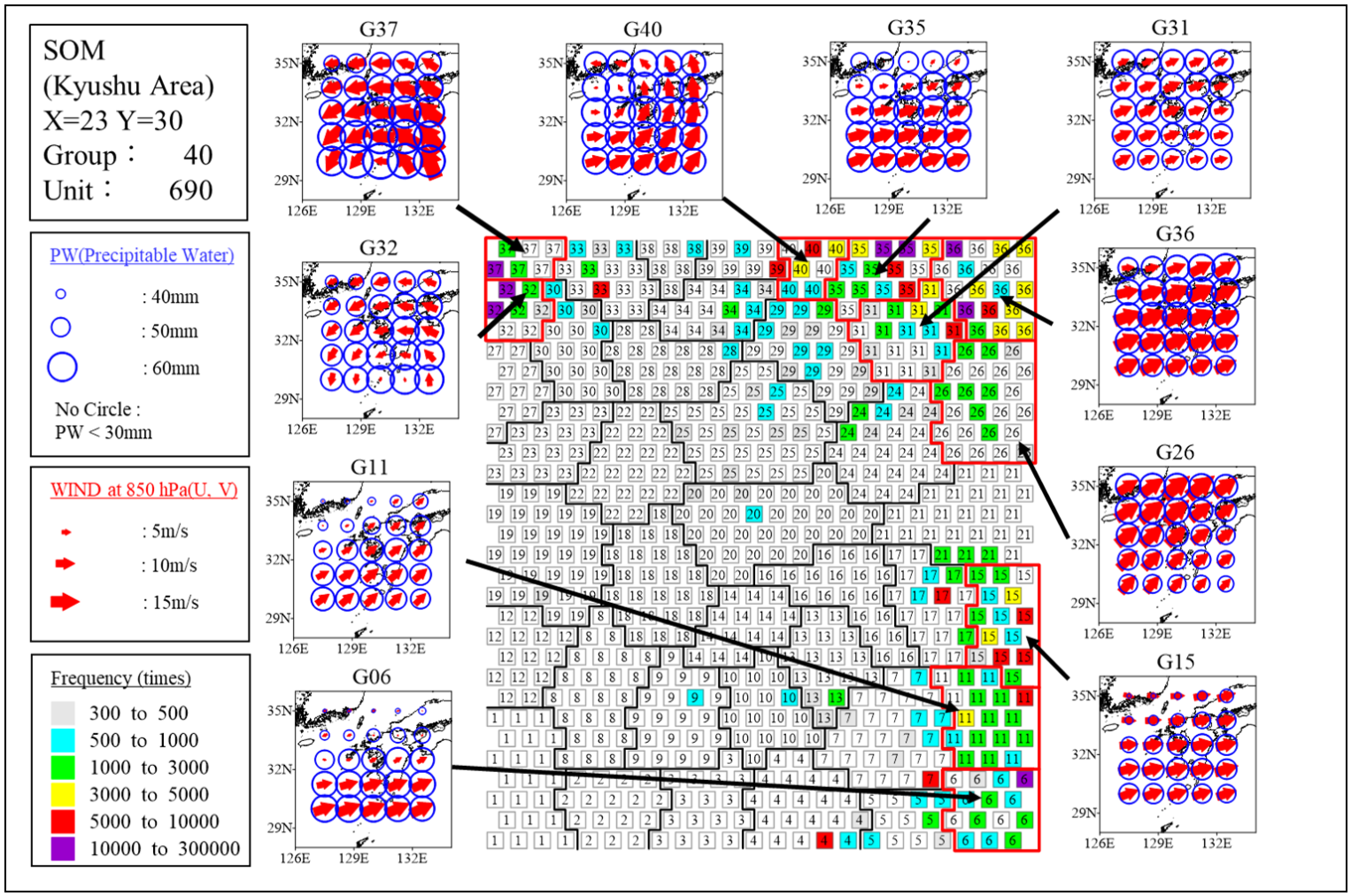


Figure 4.6: The SOM with heavy rainfall frequency over 50 mm/h in each unit (The number in each unit represents the group number), the top 10 synoptic weather groups caused heavy rainfall (depicted in red frame) and the illustration of synoptic weather condition (outside the SOM)

Figure 4.6 shows the top 10 synoptic meteorological groups which are characterized by Precipitable Water (PW) depicted in the blue circle and the 850 hPa wind depicted in the red arrow. These water content and wind condition data can explain the main factors causing heavy rainfall such as the existence of a stationary front, low pressure system, tropical depression, and typhoon, which leads to the comprehensive understanding of synoptic weather groups. Then, the top 10 groups are classified into four patterns discussed below.

Firstly, in the up-right groups (G26, G31, G35 and G36), strong southwest wind and large amounts of PW are observed over the entire Kyushu area, which indicates the existence of the lower-level jet (LLJ). It is recognized as one of typical patterns during the East Asian monsoon period and it has induced heavy rainfall events in Japan. The dominating wind direction is slightly changed from southwest (G26) to west (G31, G35 and G36) as the position of the group on the SOM are moving upward.

Secondly, in the upper-middle group (G40), a counterclockwise circulation over the western coast of Kyushu and an intrusion of moisture tongue in Kyushu area are clearly shown. The wind comes from the south toward the center of the low-pressure system, which transports the large amount of PW through Kyushu area. It is clearly distinguished from the existence of tropical depression or typhoon because the wind speed is much higher as points become distant from the center of low-pressure system. It is completely the opposite feature compared to tropical depression or typhoon.

Thirdly, in the upper-left groups (G32 and G37), the tropical depression or typhoon is assumed to exist because they show cyclonic circulation with a counterclockwise wind direction and largest amount of PW just outside the center. The center is assumed to exist where it is around latitude 30 degrees north and longitude 130 degrees east considering the relatively weak wind around there. Moreover, G37 is assumed to be much more powerful than that of G32 in terms of the stronger wind speed and larger amount of PW. Therefore, G37 indicates the existence of typhoon, whereas, G32 indicate the existence of tropical depression.

Fourthly, in the lower-right groups (G6, G11 and G15), the stationary front which is known as Meiyu-baiu front is assumed to exist over Kyushu area because the gradient of PW is clearly seen between the northern and the southern Kyushu area. In other words, low precipitable water in the northern area and high contained water in the southern area, which is clearly seen in G6. The exact position of a front is located where the gradient of PW reaches its maximum. Among these groups, larger amount of PW is flown into the southern Kyushu area in the group (G6) compared to the group(G15).

Following the classification of the top 10 groups, the meteorological fields with a high frequency of heavy rainfall can be characterized by these four patterns 1) strong southwest wind and large amounts of PW, 2) counterclockwise circulation and intrusion of moisture tongue, 3) tropical depression or typhoon and 4) stationary front known as Meiyu-Baiu front. The features causing heavy rainfall are slightly different between the northern and the southern Kyushu area. The conditions of 1) and 2) are highly contributed to

heavy rainfall events in the northern area, whereas, the conditions of 3) and 4) can be particularly recognized in the southern area.

4.4 The SOM and recent heavy rainfall events

4.4.1 General information

For 6 disaster events caused by heavy rainfall in 20 years, 25 or more people have been killed or missing in each event. Among them, four events happened in the last decade. Table 4.2 shows that Fukuoka prefecture located in the northern Kyushu area have been affected in many events, but the other prefectures are also damaged in these disaster events. Particularly, the event named as Western Japan heavy rainfall happened in 2018 induced devastating damage on Kyushu and neighboring Chugoku region due to flooding and landslide disaster events. Moreover, in July 2020, the Kuma River which flows through Kumamoto Prefecture located in middle part of Kyushu, was flooded and over 80 people has been killed and are still missing. This disaster event is one of recent catastrophic heavy rainfall events in Kyushu area. Dead and missing statistics are slightly different in some reports. This study follows each disaster report published by Fire and disaster Management Agency (FDMA).

Table 4.2: Catastrophic heavy rainfall disasters which are officially named by Japan Meteorological Agency in Kyushu Area after 2000

Disaster Event (Year/month)	Dead and Missing toll in Japan (in Kyushu)	Mainly Damaged Area in Kyushu Island (Prefecture・Area in Kyushu)
2006/7	28 (5)	Miyazaki, Kagoshima (Southern)
2009/7	35 (12)	Fukuoka, Nagasaki (Northern)
2012/7	32 (32)	Fukuoka, Kumamoto (Northern)
2017/7	44 (42)	Fukuoka, Oita (Northern)
2018/7	271 (10)	Fukuoka, Saga, Nagasaki (Northern)
2020/7	86 (79)	Fukuoka, Kumamoto (All)

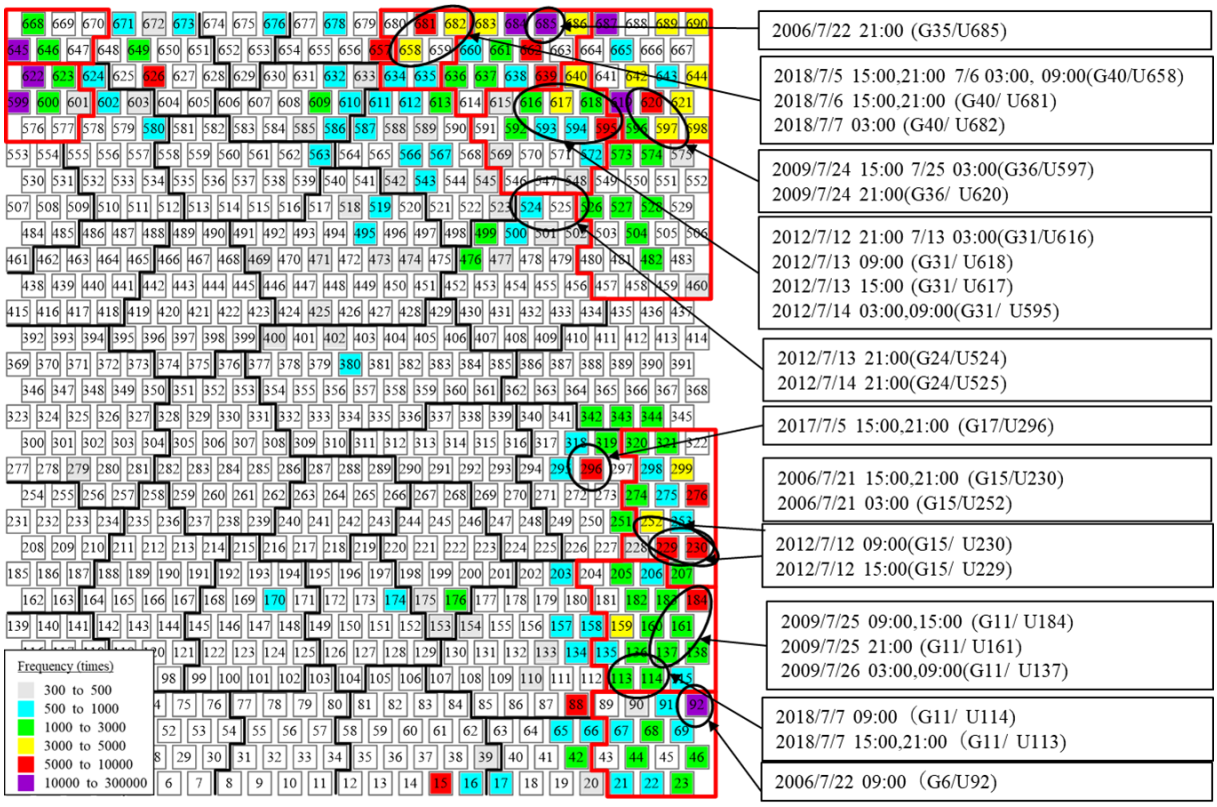


Figure 4.7: Correspondence between the units which represent the frequency of rainfall intensity over 50 mm per hour and when the heavy rainfall disaster happened (groups surrounded by red lines represents the top 10 groups)

Figure 4.7 shows the top 10 groups depicted in the red frame have contributed to catastrophic heavy rainfall events happened in last two decades. However, the characteristic of each event needs to be analyzed in terms of its stability and location on the SOM. For instance, in the event of 2006, the position of corresponding units on the SOM was changed as time passed. This implies that many types of synoptic weather pattern had induced heavy rainfall in Kyushu area. On the other hand, in the event of 2018, synoptic pattern was almost same for 2 days long. This implies that similar weather pattern had continued during this event. Among historical heavy rainfall events, three examples are selected and discussed in the latter part.

4.4.2 2012/7 Northern Kyushu heavy rainfall event

The northern Kyushu area experienced intensive rainfall from 11 to 14 July 2012. This event triggered many shallow landslides, especially on the northeastern rim of Aso caldera. During this event, 30 lives were passed, and another 2 lives were unaccounted, and 13,263 houses were completely or practically damaged. Recorded cumulative rainfall exceeded 500 mm at five sites, with totals ranging from 570.5 to 816.5 mm (Yang, 2015).

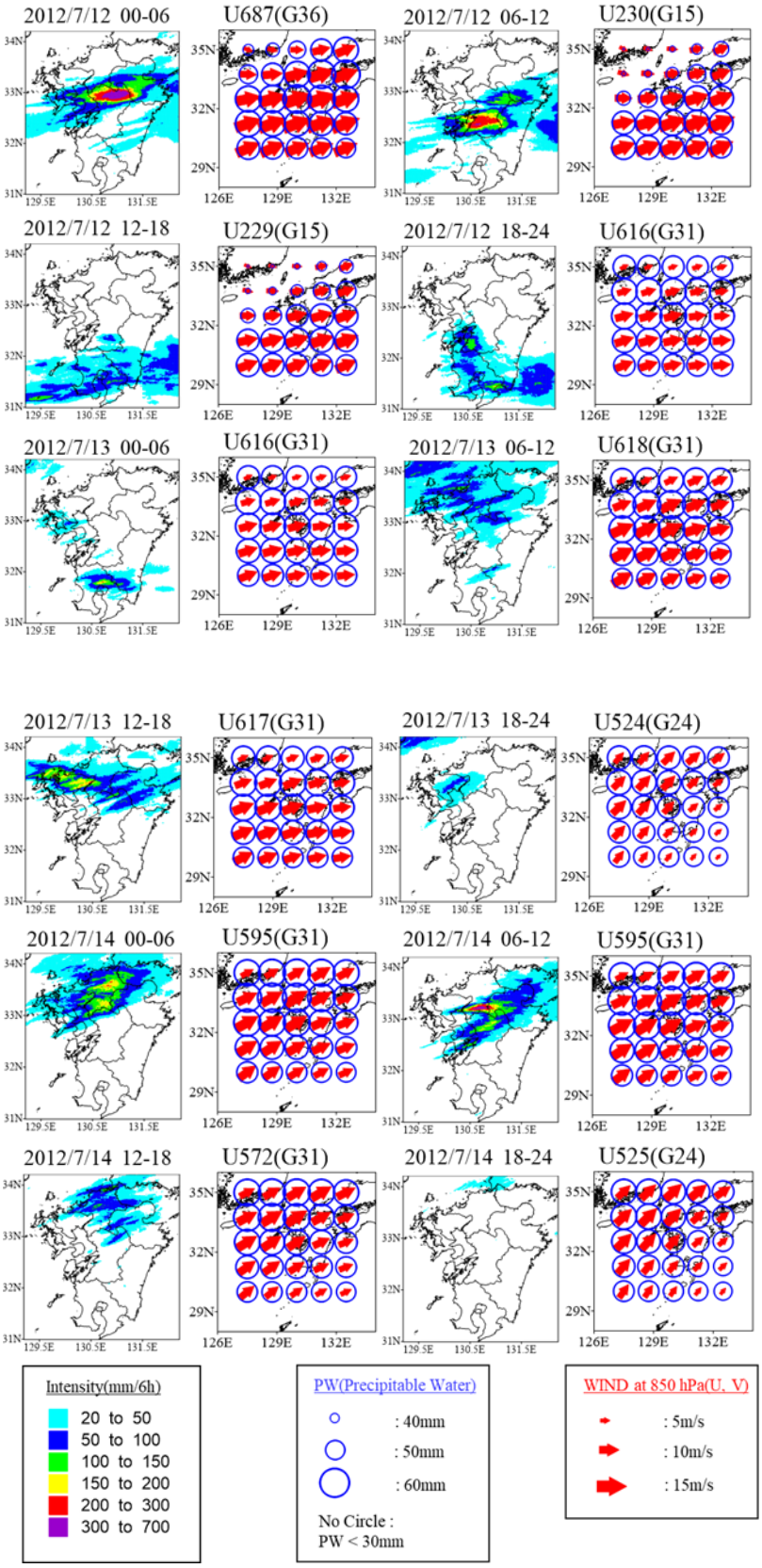


Figure 4.8: The rainfall distribution for each 6 h in Kyushu Area and the representative unit corresponding to the synoptic weather pattern during 2012/7 Northern Kyushu rainfall event

Table 4.3: The unit and the possibility of causing rainfall over 50 mm per hour in All or Northern Kyushu area in the time series during 2012/7 Northern Kyushu rainfall event

No	Date	Unit(Group)	p(≥50mm/h) All area	p(≥50mm/h) Northern Area
1	7/12 03:00	687 (G36)	75	50
2	7/12 09:00	230 (G15)	55	20
3	7/12 15:00	229 (G15)	40	0
4	7/12 21:00	616 (G31)	46.2	15.4
5	7/13 03:00	616 (G31)	46.2	15.4
6	7/13 09:00	618 (G31)	100	100
7	7/13 15:00	617 (G31)	70	50
8	7/13 21:00	524 (G24)	66.7	50
9	7/14 03:00	595 (G31)	46.2	38.5
10	7/14 09:00	595 (G31)	46.2	38.5
11	7/14 15:00	572 (G31)	25	25
12	7/14 21:00	525 (G24)	28.6	14.3

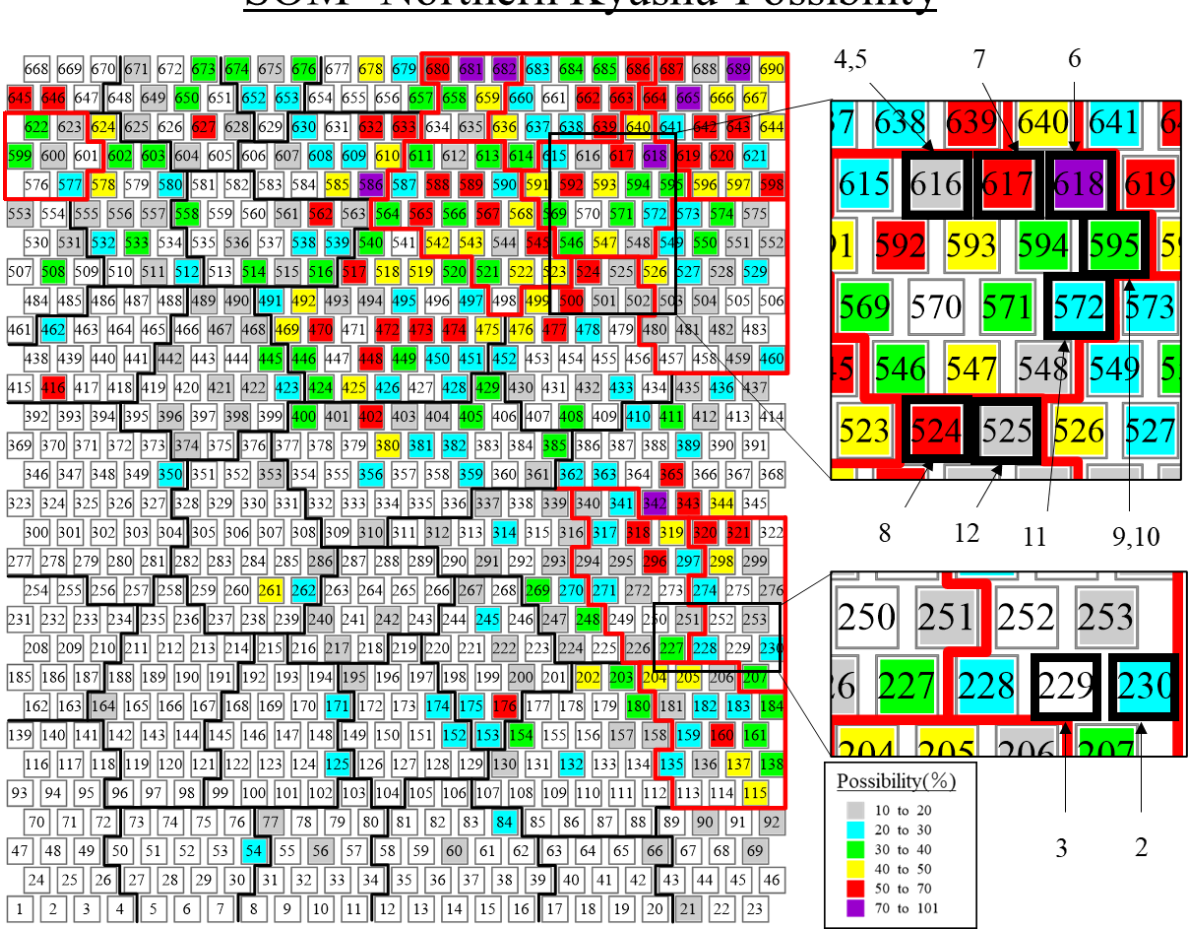


Figure 4.9: The units which represent the possibility of causing heavy rainfall over 50 mm per hour in the northern Kyushu during 2012/7 Northern Kyushu heavy rainfall event (the groups surrounded by red line represent the top 10 groups in the northern Kyushu area and the unit surrounded by black bold line in the enlarged figures represents the classified pattern during this event)

Table 4.3 shows the time series of the unit and possibility of having heavy rainfall over 50mm per hour. The synoptic weather patterns (G15, G31 and G36) which have contributed to heavy rainfall events so far appeared in this event. However, the unit and group are slightly changed in the time series, which should be discussed in the following section.

In Figure 4.8, a stationary front and the intrusion of strong lower-level jet are clearly seen in the morning on 12th July, and they induced torrential rainfall in the northern Kumamoto prefecture. In the afternoon, the position of unit on the SOM was considerably changed from G36 to G15, as is shown in Table 4.3. It means the southward movement of a stationary front and intrusion of dry air into the northern Kyushu area.

On July 13th, the synoptic pattern was almost same (G31) and the stationary front induced heavy rainfall in the northern Kyushu area. The exact position of the unit on the SOM is shown in Figure 4.9. On July 14th, the southern Fukuoka hit the torrential rainfall. The damaged area was very close to the area where it was damaged on July 12th. However, the synoptic pattern was slightly different between two events. In terms of synoptic weather condition, there was a stationary front over the northern area in the first event on 12th July, whereas, there was a moisture tongue over all Kyushu in the second one event on 14th July.

Table 4.3 also shows the possibility of having rainfall over 50mm per hour was changed in the time series. Unit 229 in G15 has the 40% chance of having intensive rainfall in all Kyushu area, however, it is 0% in the northern Kyushu area, which shows the G15 is practically seen in the rainfall event happened in the southern Kyushu. It also matches what is discussed in the chapter 4.3. Moreover, it should be noted that the higher probability cases do not lead to the catastrophic rainfall event directly. For instance, the possibility was 100% in 9am on 13th, July, however, it was relatively weak intensity level and rainfall area was limited compared to 3am on 12th, July when the torrential rainfall had fallen in Aso region.

4.4.3 2017/7 Northern Kyushu heavy rainfall event

Northern Kyushu during July 5-6, 2017 was damaged by a torrential rainfall events known as the July 2017 Northern Kyushu Heavy Rainfall. This event caused multiple disasters and river flooding. In Fukuoka and Oita Prefecture, 41 people were killed or went missing, more than 1400 residents were damaged. A small area, measuring 20km east-west and 5km north-south, received 400mm or more of rainfall for 6 h (Kato, 2018).

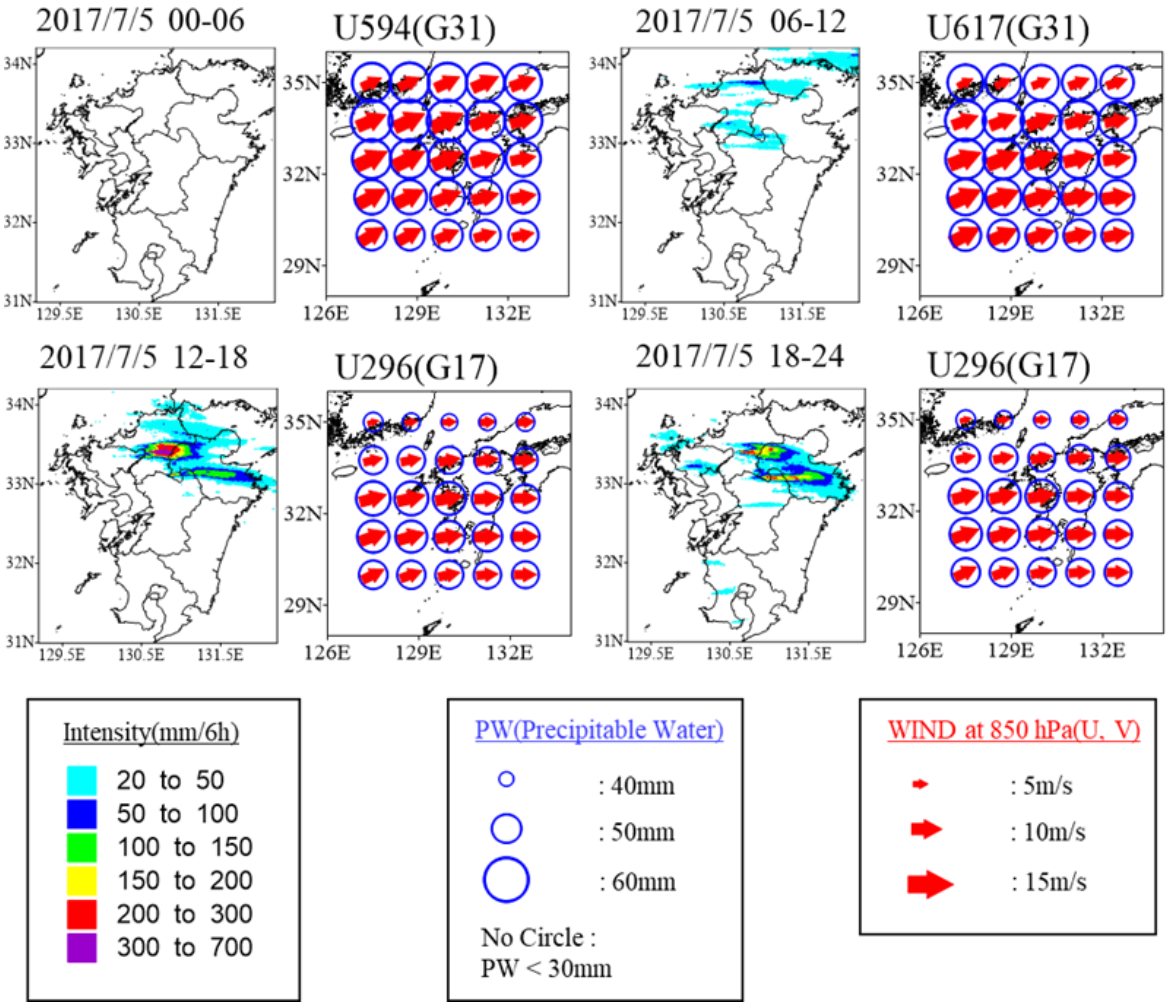


Figure 4.10: The rainfall distribution for each 6 h in Kyushu Area and the representative unit corresponding to the synoptic weather pattern during 2017/7 Northern Kyushu rainfall event

Table 4.4: The unit and the possibility of causing rainfall over 50 mm per hour in All or Northern Kyushu area in the time series during 2017/7 Northern Kyushu rainfall event

No	Date	Unit(Group)	p(≥50mm/h) All area	p(≥50mm/h) Northern Area
1	7/5 03:00	594 (G31)	40.9	31.8
2	7/5 09:00	617 (G31)	70	50
3	7/5 15:00	296 (G17)	71.4	57.1
4	7/5 21:00	296 (G17)	71.4	57.1

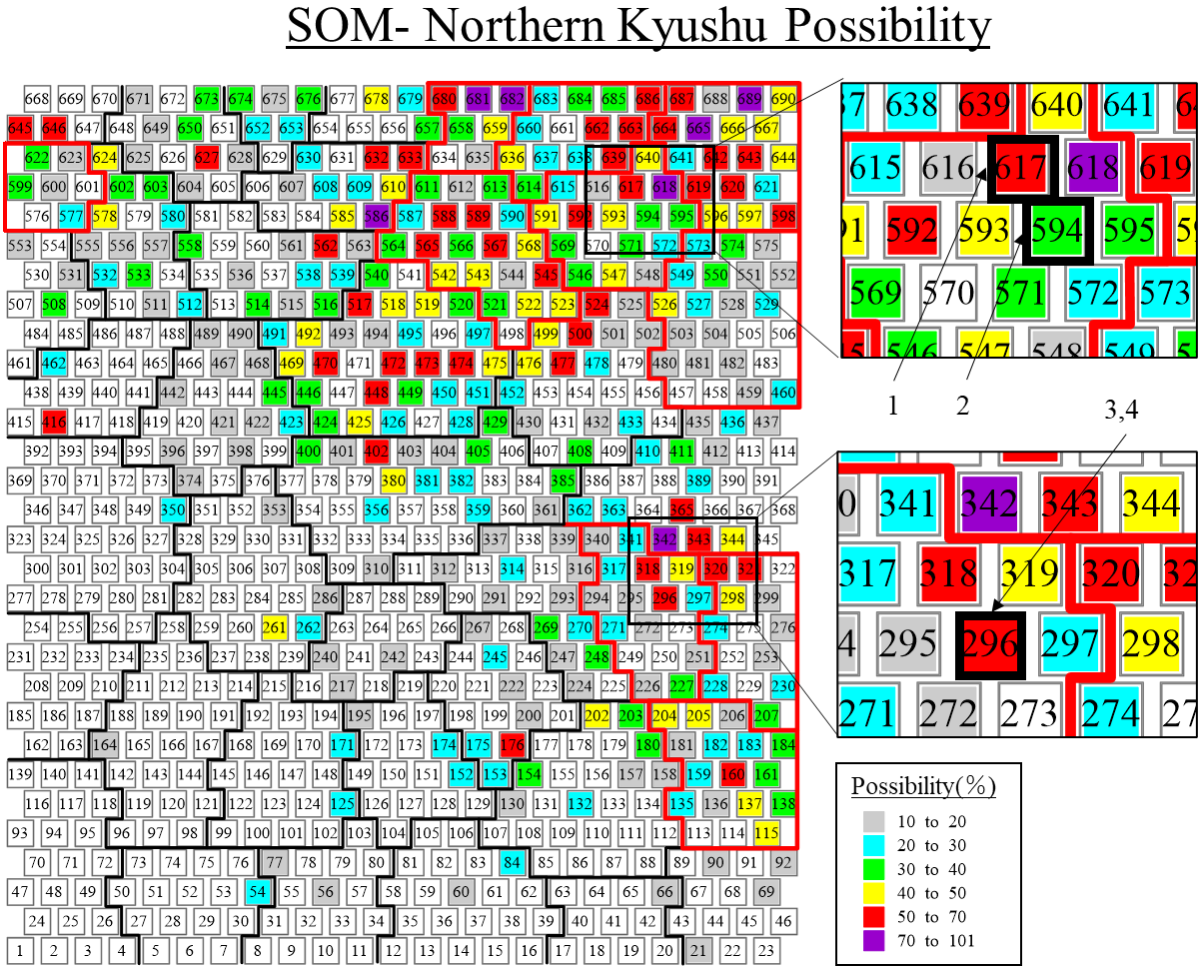


Figure 4.11: The units which represent the possibility of causing heavy rainfall over 50 mm per hour in the northern Kyushu during 2017/7 Northern Kyushu heavy rainfall event (the groups surrounded by red line represent the top10 groups in the northern Kyushu area and the unit surrounded by black bold line in the enlarged figures represents the classified pattern during this event)

Figure 4.10 shows the strong southwest wind and the intrusion of large amounts of PW are clearly seen from the morning on 5th July, which has induced torrential rainfall in the limited area of Fukuoka Prefecture. On this day, the meteorological condition was changed from G31 to G17, which has contributed to heavy rainfall events in the northern Kyushu area so far, as seen in Figure 4.11.

Table 4.4 shows U296 in G17 has 71.4 % chance of intensive rainfall in all Kyushu area and 57.1 % in the northern Kyushu area, respectively. Despite observing intensive rainfall only in the afternoon, the existence of low-level jet was assumed for all day long on 7th July, as is shown in Figure 4.10. It implies there was a potential to have intensive rainfall for all day long. Moreover, the pattern of U617 in G31 observed at 09:00 was also seen in the 2012/7 Northern Kyushu rainfall event.

4.4.4 2018/7 West Japan heavy rainfall event

Southern western Japan was damaged by the worst torrential rainfall event in 2018 during more than three decades in Japan. This event resulted in killing 225 people and forcing millions to evacuate across 15 prefectures. Hiroshima prefecture located in east of Kyushu area had 1243 mudslides, which is much more than the entire nation's total in an average year (Bandaru, 2020).

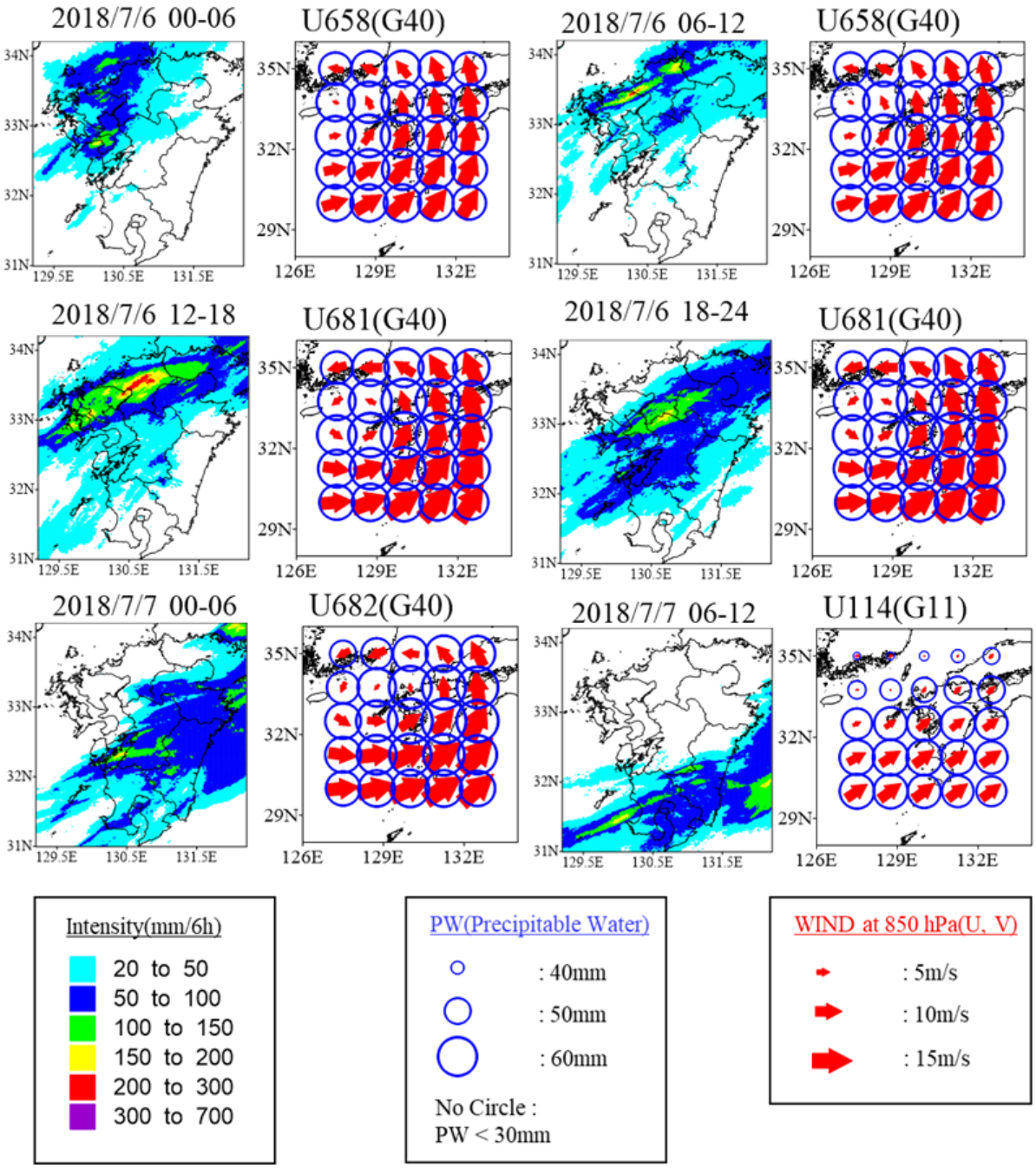


Figure 4.12: The rainfall distribution for each 6 h in Kyushu Area and the representative unit corresponding to the synoptic weather pattern during 2018/7 West Japan rainfall event

Table 4.5: The unit and the possibility of causing rainfall over 50 mm per hour in all or the northern Kyushu area in the time series during 2018/7 West Japan heavy rainfall event

No	Date	Unit(Group)	p(≥50mm/h) All area	p(≥50mm/h) Northern Area
1	7/6 03:00	658 (G40)	57.1	35.7
2	7/6 09:00	658 (G40)	57.1	35.7
3	7/6 15:00	681 (G40)	80	80
4	7/6 21:00	681 (G40)	80	80
5	7/7 03:00	682 (G40)	70	70
6	7/7 09:00	114 (G11)	54.5	9.1

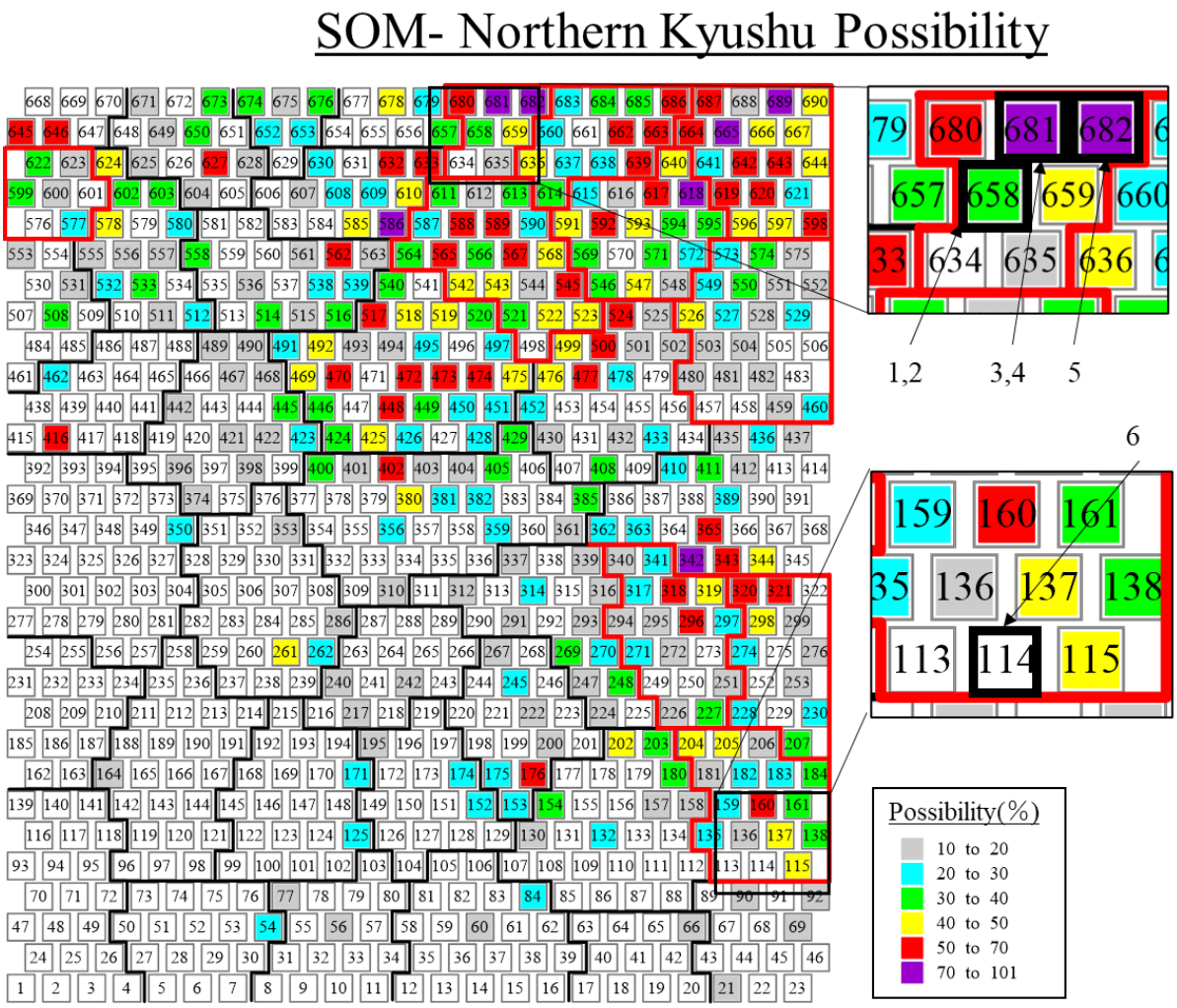


Figure 4.13: The units which represents the possibility of causing heavy rainfall over 50mm per hour in the northern Kyushu during 2018/7 West Japan heavy rainfall event (the groups surrounded by red line represent the top10 groups in the northern Kyushu area and the unit surrounded by black bold line in the enlarged figures represents the classified pattern during this event)

The group (G40) is characterized as the pattern of the intrusion of large amounts of precipitable water due to the existence of lower pressure system over the western coast of Kyushu area and it has contributed to heavy rainfall events in the northern Kyushu area, as is discussed before. Table 4.5 shows U681 of G40 has the 80 % chance of intensive rainfall in all Kyushu area. On 6th July, the stationary front was stalled in the northern Kyushu area and the synoptic pattern was almost same for all day long as shown in Figure 4.12, which means the duration of heavy rainfall was quite long and resulted in severe flooding and land slide disasters. This feature can be clearly and visually recognized by the position of the unit on SOM in Figure 4.13. On 7th July, meteorological pattern was completely changed by the intrusion of the dry air into the northern Kyushu area. The possibility of heavy rainfall in the northern Kyushu area was dropped as the front moved southward.

Chapter 5: Result 2 (Diagnosis of heavy rainfall pattern)

5.1 General information

The frequency and possibility of having intensive rainfall over 50 mm/h in each unit on the SOM, and what kinds of meteorological conditions contributed to heavy rainfall events are revealed in the previous chapter. Then, in this chapter, the structured SOM is combined with the predicted meteorological data given by GSM and it is utilized for the early prediction of heavy rainfall events.

It is examined how the heavy rainfall risk increased before the occurrence of heavy rainfall in terms of the time series. This study employs three heavy rainfall disaster cases happened in 2019 and 2020, as is shown in Table 5.1. In these events, JMA announced the highest-level alert for the heavy rainfall which is operated from 2013. This warning means residents must take action to save lives. Case 2 and 3 can be said as the consecutive case in some research, however the damaged area is different in these cases. Therefore, they are treated as the different cases in this research.

Table 5.1: Three case studies for the prediction of heavy rainfall

Case	DATE	Mainly Damaged Area in Kyushu Island (Prefecture・Area in Kyushu)
1	2019/8/28	Saga, Nagasaki(Northern)
2	2020/7/4	Kumamoto,Kagoshima (Southern)
3	2020/7/6	Fukuoka(Northern)

It is difficult to estimate when the natural disasters related to heavy rainfall happened in the particular area exactly, so the target forecasting time should be considered by following the time when the highest-level warning announced in the area. Then, in this research, the target forecasting time is decided just before the announcement of the warning for heavy rainfall because the human lives were already in danger when the highest-level alert was called. For instance, the river was already submerged, and the landslide disaster was already occurred at that time. Therefore, the target forecasting time means *heavy rainfall time* in this study and the time lags between the initial calculation time in GSM and the target forecasting time (*heavy rainfall time*) are expressed as Lead Time (LT). 0 (LT) shows the *heavy rainfall time*. In this study, the maximum LT is set to 72 h (3days) before the occurrence of heavy rainfall.

Table 5.2: The target forecasting time (heavy rainfall time) and the time when the warning announced

Case	The target forecasting time (*heavy rainfall time*)	The highest-warning time
1	2019/8/28 3:00	2019/8/28 5:50
2	2020/7/4 3:00	2020/7/4 4:50
3	2020/7/6 15:00	2020/7/6 16:30

As is mentioned in the methodology, the probability is calculated based on the assigned unit and its surrounding units to include the uncertainty of predicted data. This aggregated unit is expressed as "p_ag" in this chapter. Moreover, the probability leading to heavy rainfall in other intensity levels such as 30 mm, 70 mm, 80 mm/h are shown in the Appendix.

5.2 Case 1: 2019/8/28 03:00

Northern Kyushu during August 26 to 29, 2019 was hit a torrential rainfall. The total precipitation during this event is 626.5 mm in Hirado (Nagasaki Prefecture) and 533 mm in Karatsu (Saga Prefecture), which is more than double the normal monthly precipitation in August. JMA announced the highest-level alert for the heavy rainfall and flooding in Northern Kyushu area at 5:50 on August 28th. Therefore, this case is examined, and the target forecasting time (*heavy rainfall time*) for early prediction is decided at 3:00 on August 28th.

Table 5.3: The predicted unit on the SOM and the probability leading to heavy rainfall over 50 mm/h in the northern and All Kyushu area, respectively at heavy rainfall time

Initial Time	LT (hour)	Predicted Unit (Group)	p_ag(≥50mm/h)(%) Northern Area	p_ag(≥50mm/h)(%) All Area
8/25 03:00	72	529(G26)	13.04	18.84
8/25 09:00	66	529(G26)	13.04	18.84
8/25 15:00	60	506(G26)	9.09	9.09
8/25 21:00	54	506(G26)	9.09	9.09
8/26 03:00	48	506(G26)	9.09	9.09
8/26 09:00	42	506(G26)	9.09	9.09
8/26 15:00	36	573(G26)	32.5	40
8/26 21:00	30	573(G26)	32.5	40
8/27 03:00	24	644(G36)	39.47	50
8/27 09:00	18	620(G36)	45.21	57.53
8/27 15:00	12	619(G36)	48.53	57.35
8/27 21:00	6	595(G31)	37.65	43.53
8/28 3:00	0	596(G36)	41.98	54.32

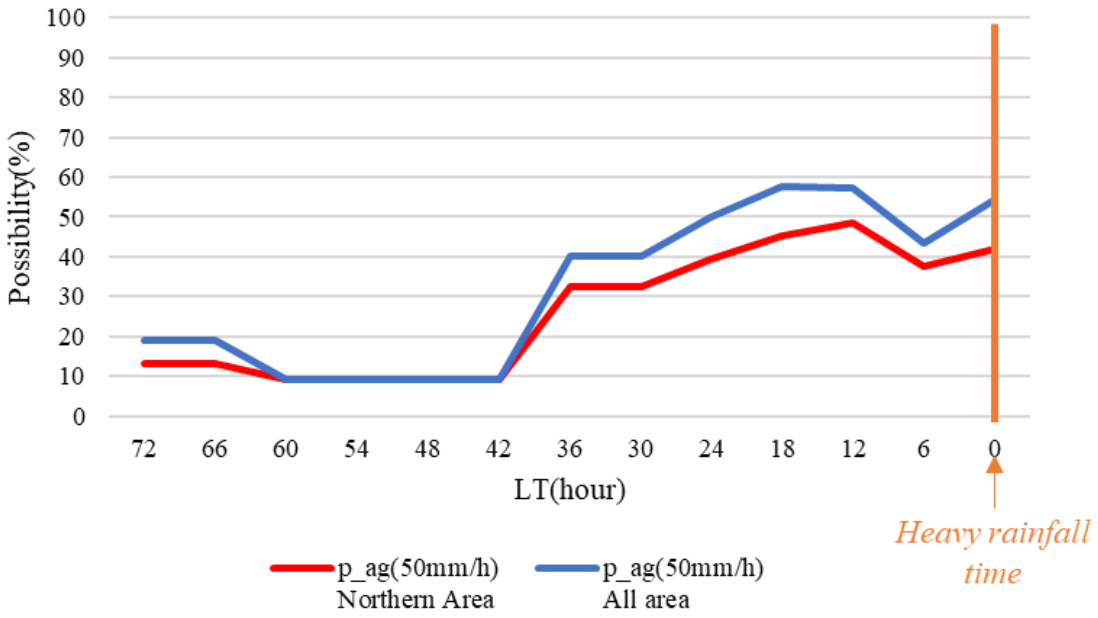


Figure 5.1: The probability leading to heavy rainfall exceeding 50 mm/h at heavy rainfall time in the Northern and all Kyushu area, respectively.

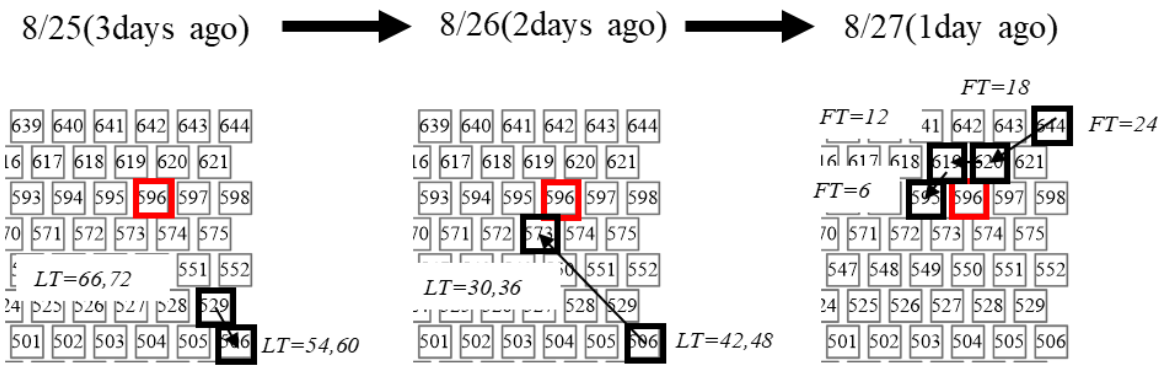


Figure 5.2: The trajectory of the predicted unit on the SOM at each initial calculation time. The red colored frame shows the actual unit at heavy rainfall time.

Table 5.3 shows the predicted unit and the probability leading to heavy rainfall over 50mm/h at 3:00 on August 28th. The prediction started from August 25th (i.e.) 3 days before the *heavy rainfall time* in this area. On August 25th, the predicted probability is relatively low at that time and the predicted unit on the SOM is far away from the actual unit which is depicted in red color, as is shown in Figure 5.2.

On August 26th, the probability has gradually increased as time advances. For instance, the initial value at 03:00 and 09:00 (LT=48,42) shows 9.09% of leading to heavy rainfall in Northern area, which can be regarded as a low risk. However, the probability increases to 32.5% at 15:00 and 21:00 (LT= 36, 30). Then, the risk of heavy rainfall can be announced at that time. Moreover, the predicted unit on the SOM is located adjacent to the actual unit spatially.

On August 27th, the position on the SOM is relatively stable and the probability has gradually increased as time approaches *heavy rainfall time* (LT=0). These probabilities indicate a high risk of heavy rainfall. The probability is 48.53% in the Northern and 57.35% in all Kyushu area (LT=12), which can be used for the early prediction for heavy rainfall events.

Moreover, the predicted unit on the SOM (U596) is so close to the unit which was seen in 2012/7 Northern Kyushu rainfall event (U595 and U618), which resulted in the flood disasters along Yabe river located in Fukuoka prefecture in both cases.

5.3 Case 2: 2020/7/4 03:00

In early July 2020, an extreme rainfall event generated precipitation over 1000 mm in Kyushu district. Especially, in central Kyushu district, there is an event with precipitation over 600 mm in 13 h (Hirockawa, 2020). JMA announced the highest-level alert for the heavy rainfall in Southern Kyushu area at 4:50 on July 4th. Therefore, this case is examined, and the target forecasting time (*heavy rainfall time*) for early prediction is decided at 3:00 on July 4th.

Table 5.4: The predicted unit on the SOM and the probability leading to heavy rainfall over 50 mm/h in the southern and All Kyushu area, respectively at heavy rainfall time

Initial Time	LT (hour)	Predicted Unit (Group)	p_ag(≥50mm/h)(%) Southern Area	p_ag(≥50mm/h)(%) All Area
7/1 03:00	72	15 (G04)	17.91	17.91
7/1 09:00	66	207 (G15)	50.7	57.75
7/1 15:00	60	44 (G06)	28.17	29.58
7/1 21:00	54	207 (G15)	50.7	57.75
7/2 03:00	48	207 (G15)	50.7	57.75
7/2 09:00	42	685 (G35)	62.5	78.12
7/2 15:00	36	685 (G35)	62.5	78.12
7/2 21:00	30	686 (G35)	52.63	77.19
7/3 03:00	24	686 (G35)	52.63	77.19
7/3 09:00	18	686 (G35)	52.63	77.19
7/3 15:00	12	686 (G35)	52.63	77.19
7/3 21:00	6	685 (G35)	62.5	78.12
7/4 3:00	0	685 (G35)	62.5	78.12

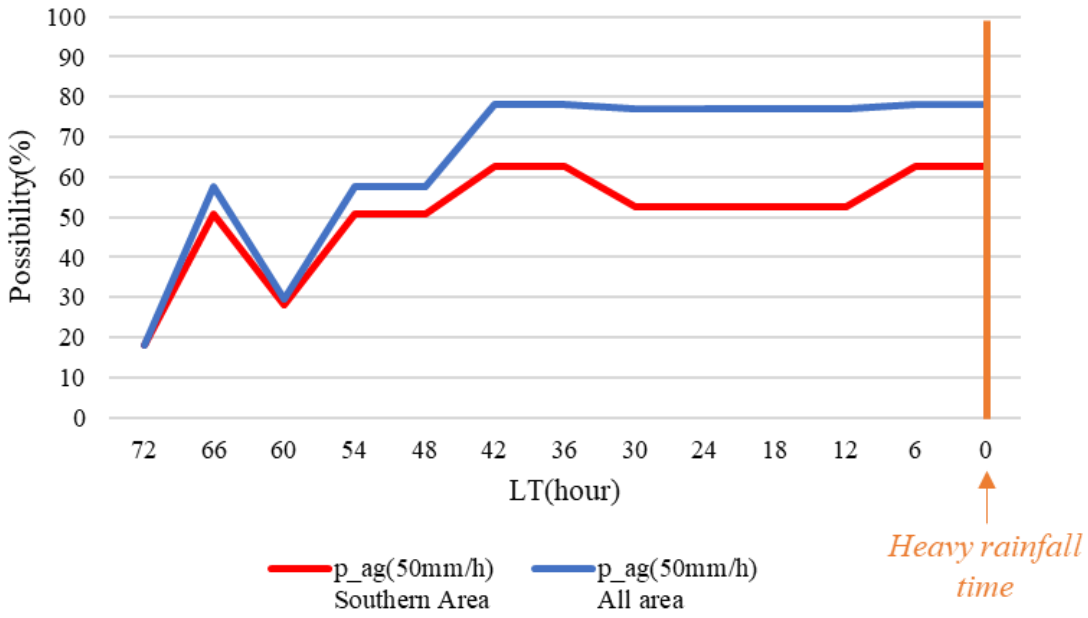


Figure 5.3: The probability leading to heavy rainfall exceeding 50 mm/h at heavy rainfall time in the southern and all Kyushu area, respectively.

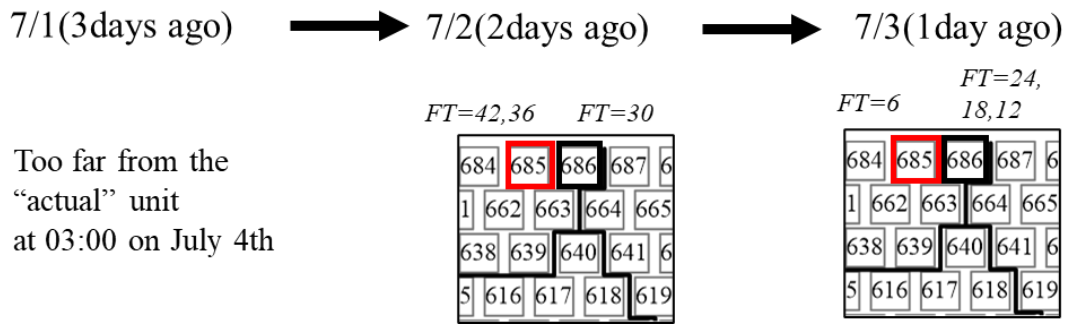


Figure 5.4: The trajectory of the predicted unit on the SOM at each initial calculation time. The red colored frame shows the actual unit at heavy rainfall time.

As is same in Case1, the initial calculation started from 3 days (LT=72) before the *heavy rainfall time* (LT=0) in the target area. On July 1st, the probability is unstable and the position on the SOM is far away from the actual unit at 3:00 on July 4th.

On July 2nd, the unit is located far away at 3:00 (LT=48), however the position considerably changed from U207 to U685 at 9:00 (LT=42). At that time, the probability leading to heavy rainfall is 62.5% in the northern and 78.12% in all Kyushu area. As is shown in Figure 5.4, the position on the SOM is stable after that time, which indicates the prediction result is relatively stable compared to what is observed on July 1st. Therefore, it is possible to announce the heavy rainfall risk at that time.

On July 3rd, the predicted meteorological condition is almost stable, and it is same as after 9:00 on July 2nd. For one and a half day, the predicted unit is U685 or U686. The probability is slightly changed in

the time series, however, it is over 50% during this period (0≤LT≤42), which indicates the high risk of upcoming rainfall.

The unit (U685) was also observed in 2006/7 heavy rainfall event. The total precipitation during 2006/7 heavy rainfall event is over 1000 mm and it can be said as severe heavy rainfall event. In case 2, torrential rainfall was observed from middle to southern Kyushu area where heavy rainfall was seen in this historical event.

5.4 Case 3: 2020/7/6 15:00

After the heavy rainfall events on July 4th, the Baiu-front moved northward and caused torrential rainfall in Northern Kyushu area such as precipitation around 250 mm in 3 h. JMA announced the highest-level alert for the heavy rainfall in Northern Kyushu area at 16:30 on July 6th. Therefore, this case is examined, and the target forecasting time (*heavy rainfall time*) for early prediction is decided at 15:00 on July 6th.

Table 5.5: The predicted unit on the SOM and the probability leading to heavy rainfall over 50 mm/h in the northern and All Kyushu area, respectively at heavy rainfall time

Initial Time	LT (hour)	Predicted Unit (Group)	p_ag(≥50mm/h)(%) Northern Area	p_ag(≥50mm/h)(%)All Area
7/3 15:00	72	598 (G36)	34.09	47.73
7/3 21:00	66	598 (G36)	34.09	47.73
7/4 03:00	60	598 (G36)	34.09	47.73
7/4 09:00	54	620 (G36)	45.21	57.53
7/4 15:00	48	686 (G35)	49.12	77.19
7/4 21:00	42	621 (G36)	44.44	55.56
7/5 03:00	36	621 (G36)	44.44	55.56
7/5 09:00	30	598 (G36)	34.09	47.73
7/5 15:00	24	621 (G36)	44.44	55.56
7/5 21:00	18	621 (G36)	44.44	55.56
7/6 03:00	12	621 (G36)	44.44	55.56
7/6 09:00	6	621 (G36)	44.44	55.56
7/6 15:00	0	621 (G36)	44.44	55.56

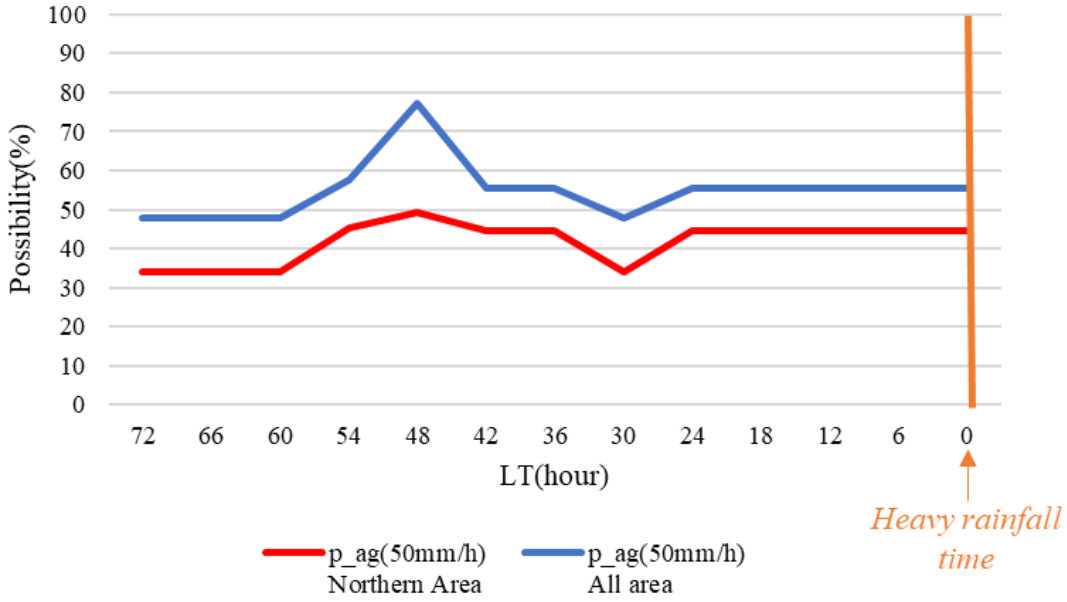


Figure 5.5: The probability leading to heavy rainfall exceeding 50 mm/h at heavy rainfall time in the Northern and all Kyushu area, respectively.

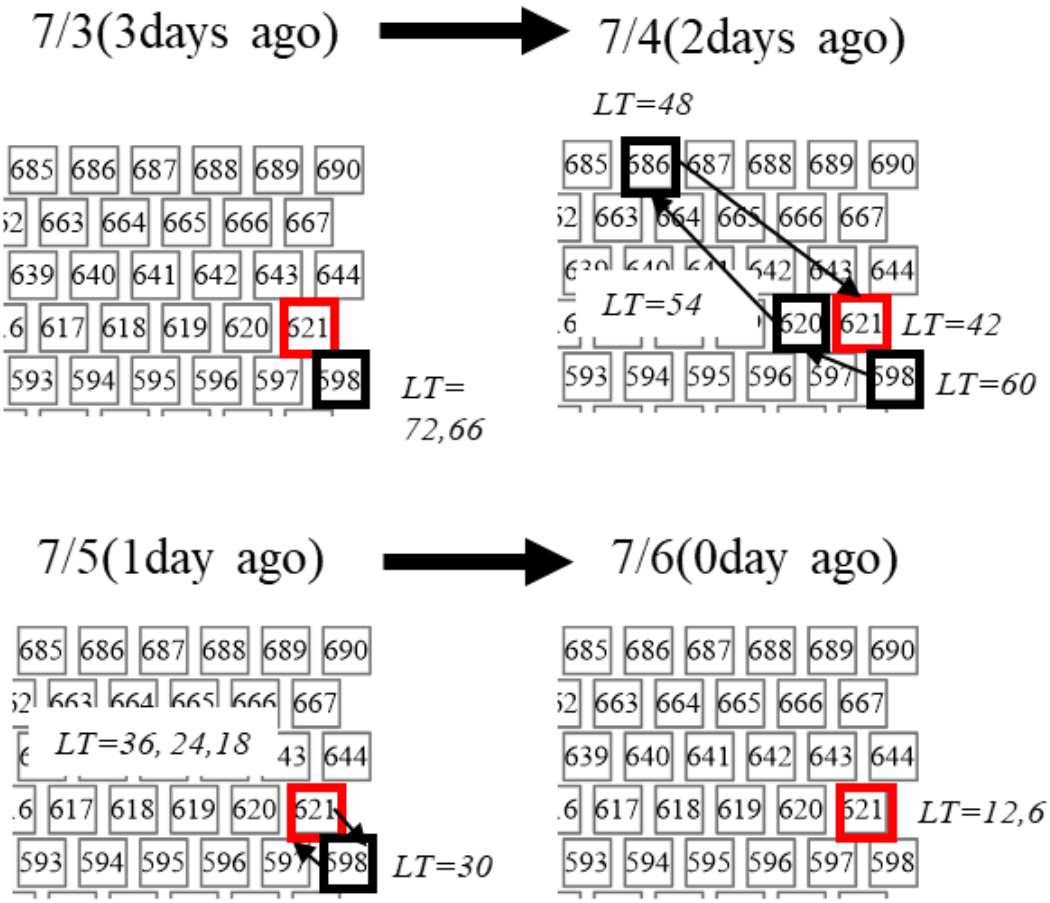


Figure 5.6: The trajectory of the predicted unit on the SOM at each initial calculation time. The red colored frame shows the actual unit at heavy rainfall time.

Apart from two study cases, on July 3rd (i.e.) 3 days before the *heavy rainfall time*, the probability is already high and the unit on the SOM is adjacent to the actual unit depicted in red color in Figure 5.6. It is because the stationary front is expected to stall around Kyushu area during 3 days before the *heavy rainfall time.*

On July 4th, there is a high probability, however the position of the unit is changed as time advances, which indicates each predicted meteorological condition is likely to have heavy rainfall. These results would be a challenging situation for decision makers who announces the warning of heavy rainfall in advance because there are some uncertainties in the predicted value. However, in any predicted units, they show the high-risk of heavy rainfall and the warning should be announced at that time.

On July 5th, the probability and the position are stable compared to the previous day. The probability is 44.44% in the northern and 55.56% in all Kyushu area, respectively at 03:00, 15:00, 21:00 (LT=36, 24, 18). The predicted unit corresponds to the actual unit at *heavy rainfall time* after 15:00 (LT=24).

The similar units (U597, U620) which are adjacent to the actual unit (U621) in Case 3, are observed in 2009/7 Chugoku・Kyushu rainfall event. In 2009/7 event, it caused extreme rainfall in Fukuoka, and Nagasaki as is same in Case 3. Therefore, this SOM database can relate diagnosed results to historical events happened before 2018, which are summarized in the conclusion part.

Chapter 6: Discussion and Conclusion

As a preliminary step, 19520 meteorological fields observed from 1979 to 2018 are classified into around 700 synoptic patterns by the heuristic rule proposed by Vesanto. Then, these patterns are divided into 40 groups by employing Ward and K-means method. The classification performance is evaluated by the Davies-Bouldin Index (DBI) and the SOM (X=23, Y=30, Groups=40, Units=690) is selected and employed in the following results. Then, the structured SOM relates meteorological fields to the frequency of heavy rainfall measured by R/A from 2006 to 2018. Therefore, the possibility leading to heavy rainfall in each unit on the SOM is calculated.

In this research, 1) what kinds of meteorological fields have highly contributed to extreme precipitation events, and natural disaster events such as flooding and landslide in Kyushu area so far are revealed in Chapter 4. Then, by combining the SOM structured in the previous chapter and predicted meteorological field by GSM, it is discussed 2) whether it is possible to diagnose heavy rainfall risk in advance and relate the predicted meteorological data to historical heavy rainfall events in Chapter 5.

6.1 Result 1

The SOM consists of 40 groups and 690 units. The top 10 groups of heavy rainfall frequency are selected to analyze synoptic weather condition because the top 10 groups explain 76.8% in entire Kyushu, 81.8% in northern and 81.4% in southern area, respectively of all heavy rainfall frequency among 40 groups. Moreover, it is revealed that these groups have contributed to recent heavy rainfall events named by JMA after 2000. Considering synoptic weather patterns characterized by precipitable water and wind factors, the top 10 groups can be interpreted by four reasons such as 1) strong southwest wind and large amounts of PW, 2) counterclockwise circulation with moisture tongue, 3) tropical cyclone and 4) stationary front. The features causing heavy rainfall are slightly different between the northern and the southern Kyushu area. The conditions of 1) and 2) are highly contributed to heavy rainfall events the in the northern area , whereas, 3) and 4) can be particularly recognized in the southern area, which can be interpreted as local characteristics in Kyushu area.

As is mentioned-above, the target area (Kyushu) is divided into a northern and southern area, and the SOM provides the characteristics of meteorological field inducing heavy rainfall in each area, however, it should be downscaled to detect the local features such as in the prefecture-level. More detailed dataset such as ERA-5 published by ECMWF should be utilized in further research. Moreover, there are other options for input variables such as equivalent potential temperature at specific altitude (e.g.) 500 hPa and 850 hPa, and geopotential height, instead of precipitable water. Especially, equivalent potential temperature is commonly used for analyzing the potential of heavy rainfall and it should be discussed in further research.

6.2 Result 2

It is examined how the heavy rainfall risk increased before the occurrence of heavy rainfall in terms of the time series. This study employs three heavy rainfall cases that happened in 2019 and 2020. Figure 6.1 shows the probability leading to heavy rainfall exceeding 50 mm/h at *heavy rainfall time* in the target area and it indicates the probability stabilizes 36 h before the *heavy rainfall time* at a high probability. In other words, the trend of relatively high probability leading heavy rainfall is seen between 36(LT) and 0(LT). The probability is around 40 to 80 % in three case studies. Therefore, if the probability of heavy rainfall exceeds 40% up to 36 h, the early warning should be announced for residents to secure evacuation time.

The SOM structured in this study can relate these diagnosed results to historical rainfall events which are already classified in the units of the SOM, as is shown in Table 6.1. In other words, the SOM can provide another effective information to judge whether the diagnosed unit shows a high-risk pattern. For instance, the predicted unit (U596) on the SOM in Case 1 is so close to the unit seen in 2012/7 Northern Kyushu rainfall event (U595), which resulted in the flood disasters in Fukuoka prefecture. Therefore, decision makers can learn where was mostly affected in the predicted unit on the SOM and estimate where will be affected in the upcoming rainfall events.

For the practical use, it should be noted that a short delay occurs in the calculation process of predicting meteorological data. Then, the predicted data can be obtained around 3.5 h after the initial time. Moreover, after obtaining the result, it takes much more time for decision makers to consider whether they should announce the warning of heavy rainfall. Therefore, it is needed to publish a set of guidelines for the interpretation of the result and further cases should be analyzed to increase the reliability of combined prediction method (i.e.) GSM and SOM based method. In the further research, some index tools such as false-alarm ratio should be employed to evaluate the performance of this prediction method. After completing the evaluation, this approach will be used effectively for decision makers in near future.

Table 6.1: The correspondence between three study cases (2019,2020) and historical heavy rainfall cases (~2018)

Case	Danger time and corresponding unit	Similar unit in the historical heavy rainfall cases happened before 2018
1	2019/8/28 3:00(U596)	2012/7/14 3:00,9:00(U595)
2	2020/7/4 3:00(U685)	2006/7/22 21:00(U685)
3	2020/7/6 15:00(U621)	2009/7/24 15:00,21:00(U597,U620) 2009/7/25 3:00(U597)

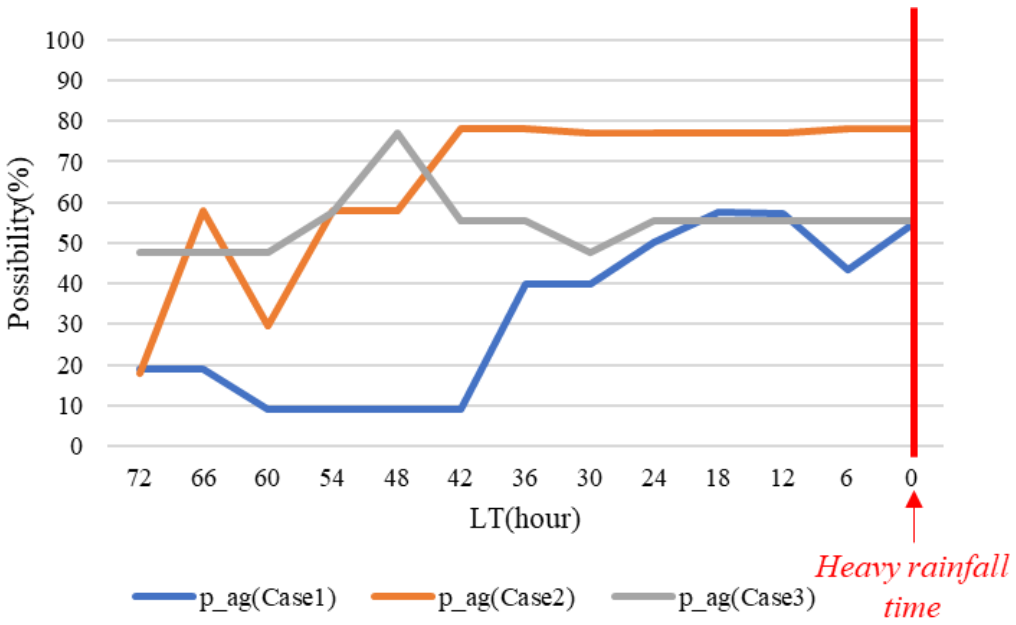


Figure 6.1: The probability leading to heavy rainfall exceeding 50 mm/h at heavy rainfall time in the Kyushu area in three study cases

References

Bandaru, S., Sano, S., Shimizu, Y., Seki, Y., Okano, Y., Sasaki, T., Wada, H., Otsuki, T. and Ito, T. (2020) 'Impact of heavy rains of 2018 in western Japan: disaster-induces health outcomes among the population among the population of Innoshima Island', *Heliyon*, **6**, e03924, DOI:10.1016/j.heliyon.2020.e03942

Cassano, E. N., Lynch, A.H., Cassano, J.J and Koslow, M.R. (2006) 'Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska, USA', *Climate Research*, **30**, 83-97, 2006

Cavazos, T. (1999) 'Large-scale circulation anomalies conductive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas', *Journal of Climate*, **12**, 1506-1523

Hewitson, B. C. and Crane, R. G. (2002) 'Self-organizing maps: applications to synoptic climatology', *Climate Research,* **22,** 13-26

Hilario, L. G. and Ivan, M.G. (2004) 'Self-organizing map and clustering for wastewater treatment monitoring', *Engineering Applications of Artificial Intelligence,* **17**, 215-225, DOI: 10.1016/j.engappai.2004.03.004

Hirockawa, Y., Kato, T., Araki K. and Mashiko, W. (2020) 'Characterisitics of an Extreme Rainfall Event in Kyushu District, Southwestern Japan in Early July 2020', *SOLA*, **16**, 265-270 DOI:10.2151/sola.2020-044

IPCC (2014): Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151pp

JMA (2020): Climate Change Monitoring Report 2019, 100pp,

Kato, R., Shimose, K. and Shimizu, S. (2019) 'Predictability of Precipitation caused by linear precipitation systems during the July 2017 Northern Kyushu heavy rainfall event using a cloud-resolving numerical weather prediction model', *Journal of Disaster Research*, **13(5)**, 846-859

Kiviluoto, K. (1996) 'Topology preservation in self-organizing maps', *Proc. International Conf. Neural Networks*, Washington, DC, 294-299

Kobayashi, S., Ota Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kanamori, H., Kobayashi, C., Endo, H., Miyaoka, K. and Takahashi, K. (2015) 'The JRA-55 Reanalysis: General Specifications and Basic Characteristics', *Journal of Meteorological Society of Japan*, **93(1)**, 5-48, DOI:10.2151/jmsj.2015-001

López, M., Valero, S., Senabre, C. and Aparicio, J. (2011) 'A SOM neural network approach to load forecasting. Meteorological and time frame influence', 2011 International Conference on Power Engineering, Energy and Electrical Drives, Malaga, pp. 1-5, DOI:10.1109/PowerEng.2011.6036553.

Matsumoto, S., Ninomiya, K. and Yoshizumi, S. (1971) 'Characteristic Features OF "Baiu" front Associated with Heavy Rainfall', *Journal of the Meteorological Society of Japan*, **49**, 267-281

Ministry of the Environment, Ministry of Education, Culture, Sports, Science and Technology, Ministry of Agriculture, Forestry and Fisheries, Ministry of Land, Infrastructure, Transport and Tourism and Japan Meteorological Agency (2018) 'Climate Change in Japan and Its Impacts': Synthesis Report on Observations, Projections and Impact Assessments of Climate Change

Mizuta, R., Oouchi, K., Yoshimura, H., Noda, A., Katayama, K., Yukimoto, S., Hosaka, M., Kusunoki, S., Kawai, H. and Nakagawa, M. (2006) '20-km Mesh Global Climate Simulations Using JMA-GSM Model - Mean Climate States-', *Journal of the Meteorological Society of Japan*, **84(1)**, 165-185

Nguyen-Le, D., Yamada, T. J. (2019) 'Using weather pattern recognition to classify and predict summertime heavy rainfall occurrence over the upper Na river basin northwestern Thailand', *Weather and forecasting*, **34**, 346-360, DOI:10.1175/WAF-D-18-0122.1

Nguyen-Le, D., Yamada. T. J. and Tran-Anh, D (2017) 'Classification and forecast of heavy rainfall in northern Kyushu during Baiu season using weather pattern recognition', *Atmospheric science letters*, **18**, 324-329, DOI:10.1002/asl.759

Ninomiya, K. and Shibasaki, Y. (2007)'Multi-Scale of the Meiyu-Baiu Front and Associated Precipitation Systems', *Journal of the Meteorological Society of Japan*, **85B**, 103-122

Nishiyama, K., Endo, S., Jinno, K., Uvo, C.B., Olsson, J. and Berndtsson, R. (2007) 'Identification of typical synoptic patterns causing heavy rainfall in the rainy season in Japan by a self-organizing map', *Atmospheric Research*, **83**, 185-200, DOI:10.1016/j.atmosres.2005.10.015

Ohashi, N. (2018) 'Frequency analysis of heavy rainfall related to synoptic weather patterns in Kyushu Island, Japan by using the Self-Organizing Map (SOM)', Master's thesis, Lund University, Sweden

Ohba, M., Kadokura, S., Nohara, D. and Toyoda, Y. (2016) 'Rainfall downscaling of weekly ensemble forecasts using self-organizing maps', *Tellus A: Dynamic Meteorology and Oceanography*, **68:1**, 20293, DOI: 10.3402/tellusa.v68.29293

Ohba, M., Kadokura, S., Yohsida, Y., Nohara, D. and Toyoda, Y. (2015) 'Anomalous weather patterns in relation to heavy precipitation events in Japan during the baiu season', *Journal of hydrometeorology*, **16**, 688-701, DOI:10.1175/JHM-D-14-0124.1

Oki, R., Sumi A. and Short, D.A. (1997) 'TRMM Sampling of Radar-AMeDAS Rainfall Using the Threshold Method', *Journal of applied meteorology*, **36**, 1480-1492

Paul, C.L., Lintner, B.R. and Sweeney, A. (2017) 'Characterizing Large-Scale Meteorological Patterns and Associated Temperature and Precipitation Extremes over the Northwestern United States Using Self-Organizing Maps', *Journal of Climate*, **30**, 2829-2847, DOI:10.1175/JCLI-D-16-0670.1

Reusch, D. B., Alley, R. B. and Hewitson, B. C. (2005) 'Relative performance of self-organizing maps and principle component analysis in pattern extraction from synthetic climatological data', *Polar Geography*, **29**, No.3, 188-212

Schwalm, C.R., Glendon, S. and Duffy, P.B. (2020) 'RCP8.5 tracks cumulative CO2 emissions', *PNAS*, **117**, 19656-19657, DOI:10.1073/pnas.2007117117

Sheridan, S.C. and Lee, C.C. (2011) 'The self-organizing map in synoptic climatological research', *Progress in Physical Geography*, **35(1)**, 109-119, DOI:10.1177/0309133310397582

Tamaki, Y., Inatsu, M., Nguyen-Le, D. and Yamada, T. J. (2018) 'Heavy rainfall duration in dynamical downscaling and its related synoptic patterns in summertime asian monsoon', *Journal of applied meteorology and climatology*, **57**, 1477-1496, DOI:10.1175/JAMC-D-17-0116.1

Wang, S., Xu, T., Nie, W., Jiang, C., Yang, Y., Fang, Z., Ki, E. and Zhang,Z (2020) 'Evaluation of precipitable water vapor from five reanalysis products with ground based GNSS observations', *Remote Sensing*, **12**, 1817 , DOI:10.3390/rs12111817

Yang, H., Wang, F., Vilmek, V., Araiba, K. and Asano, S. (2015) 'Investigation of rainfall-induced shallow disasters on the northeastern rim of Aso caldera, Japan, in July 2012', *Geoenvironmental Disasters,* **2:20**, DOI:10.1186/s40677-015-0028-3

Appendix

A: All synoptic weather groups

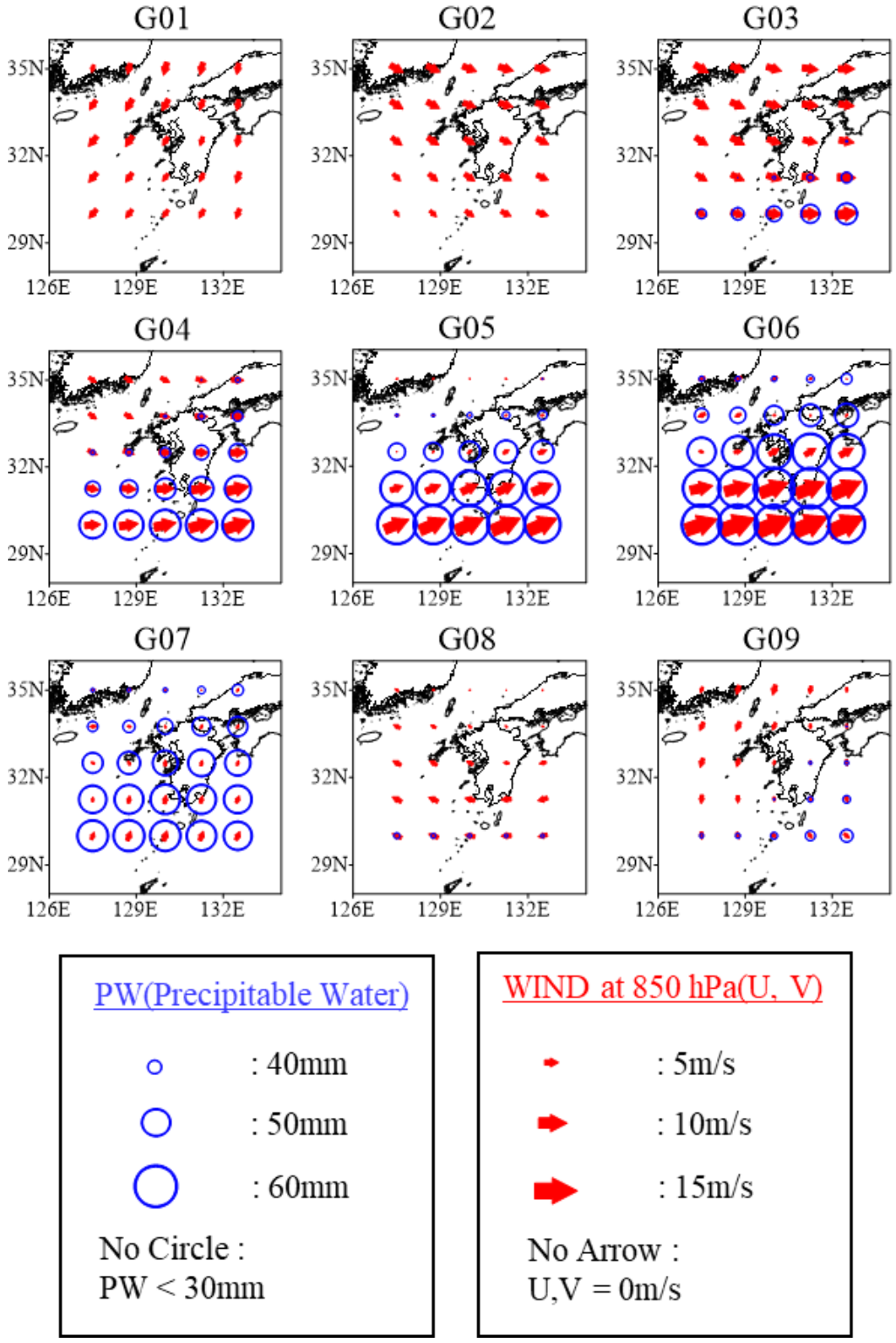


Figure A.1: The synoptic weather groups from G01 to G09

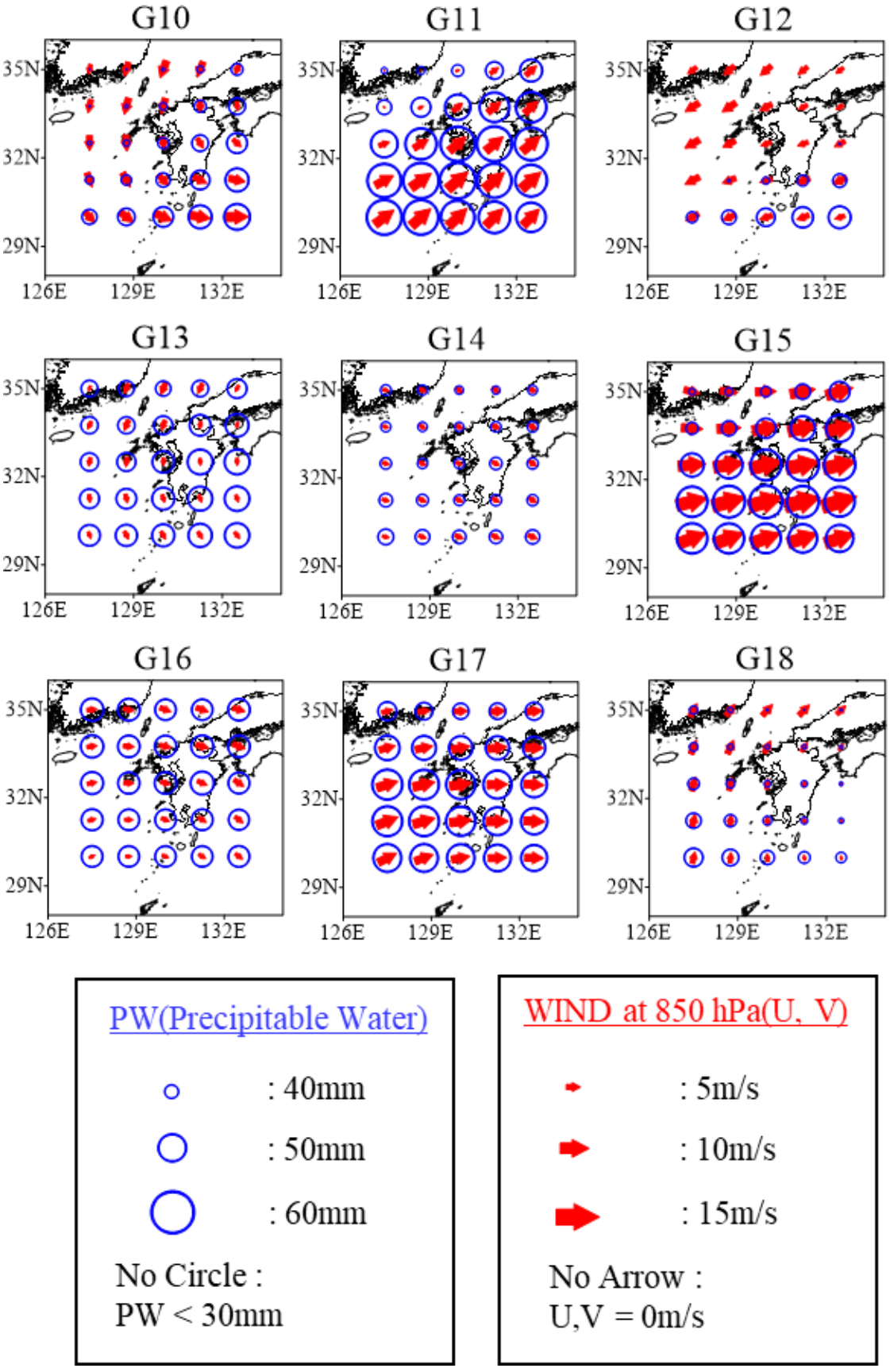


Figure A.2: The synoptic weather groups from G10 to G18

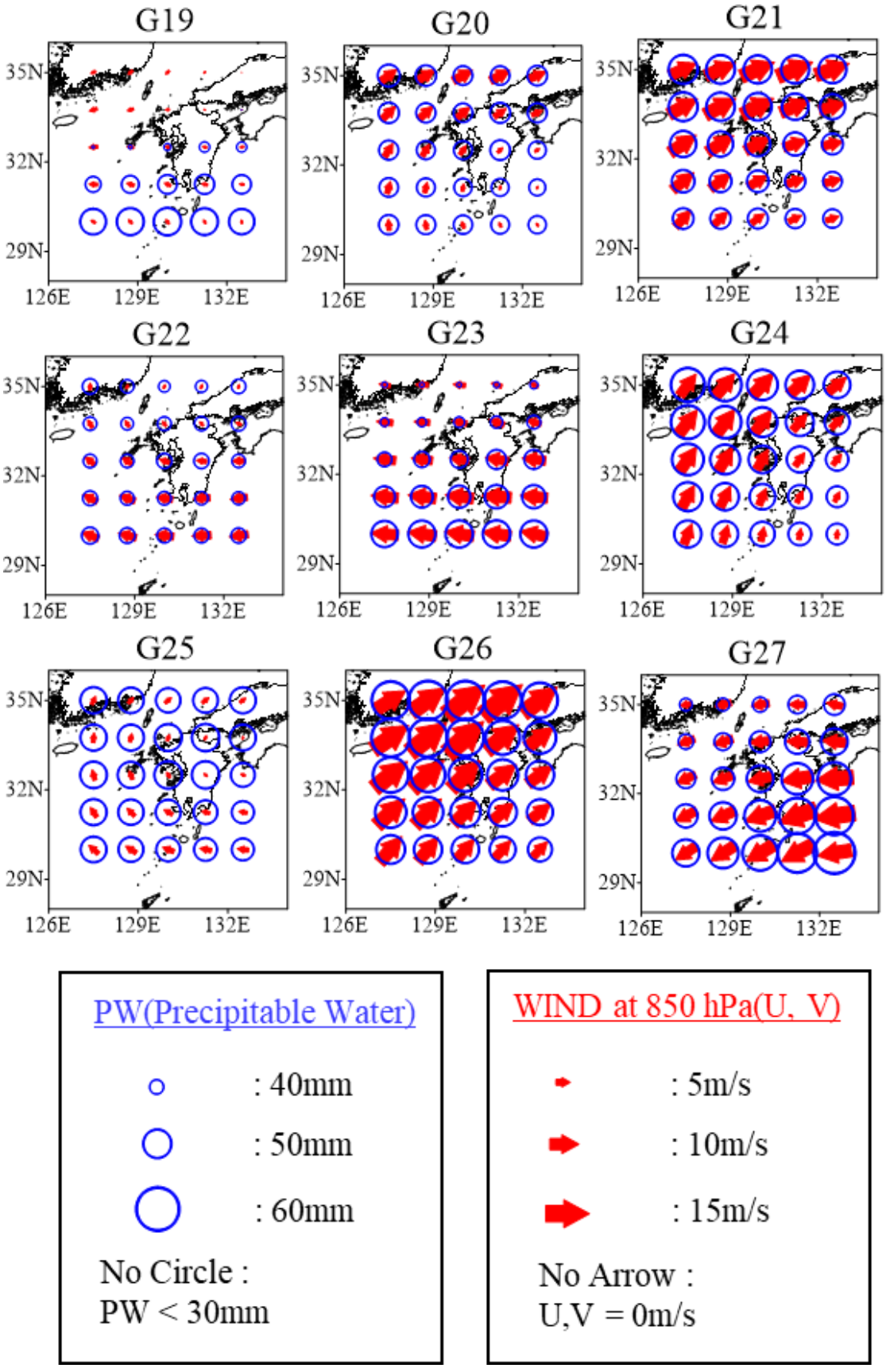


Figure A.3: The synoptic weather groups from G19 to G27

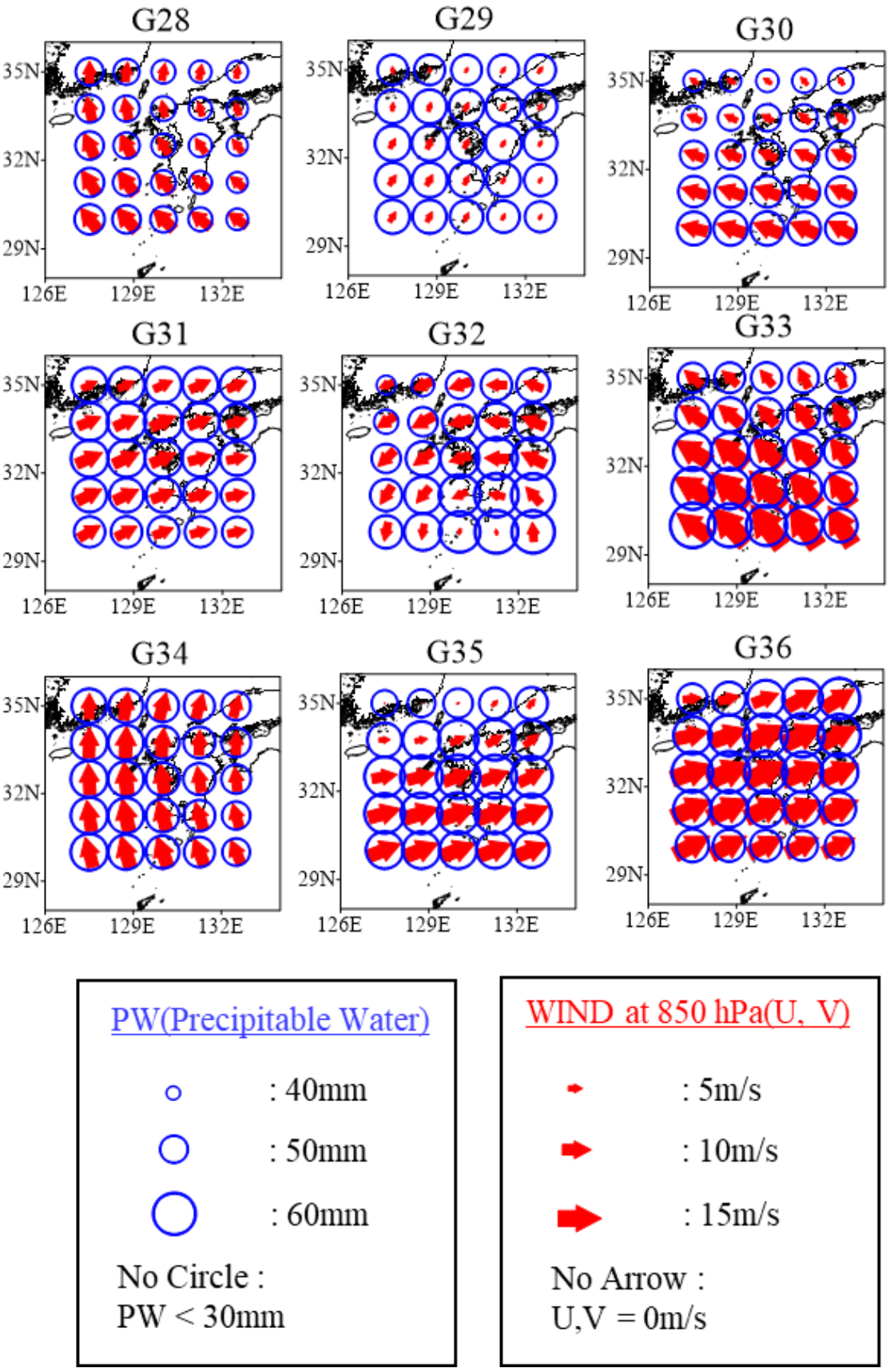


Figure A.4: The synoptic weather groups from G28 to G36

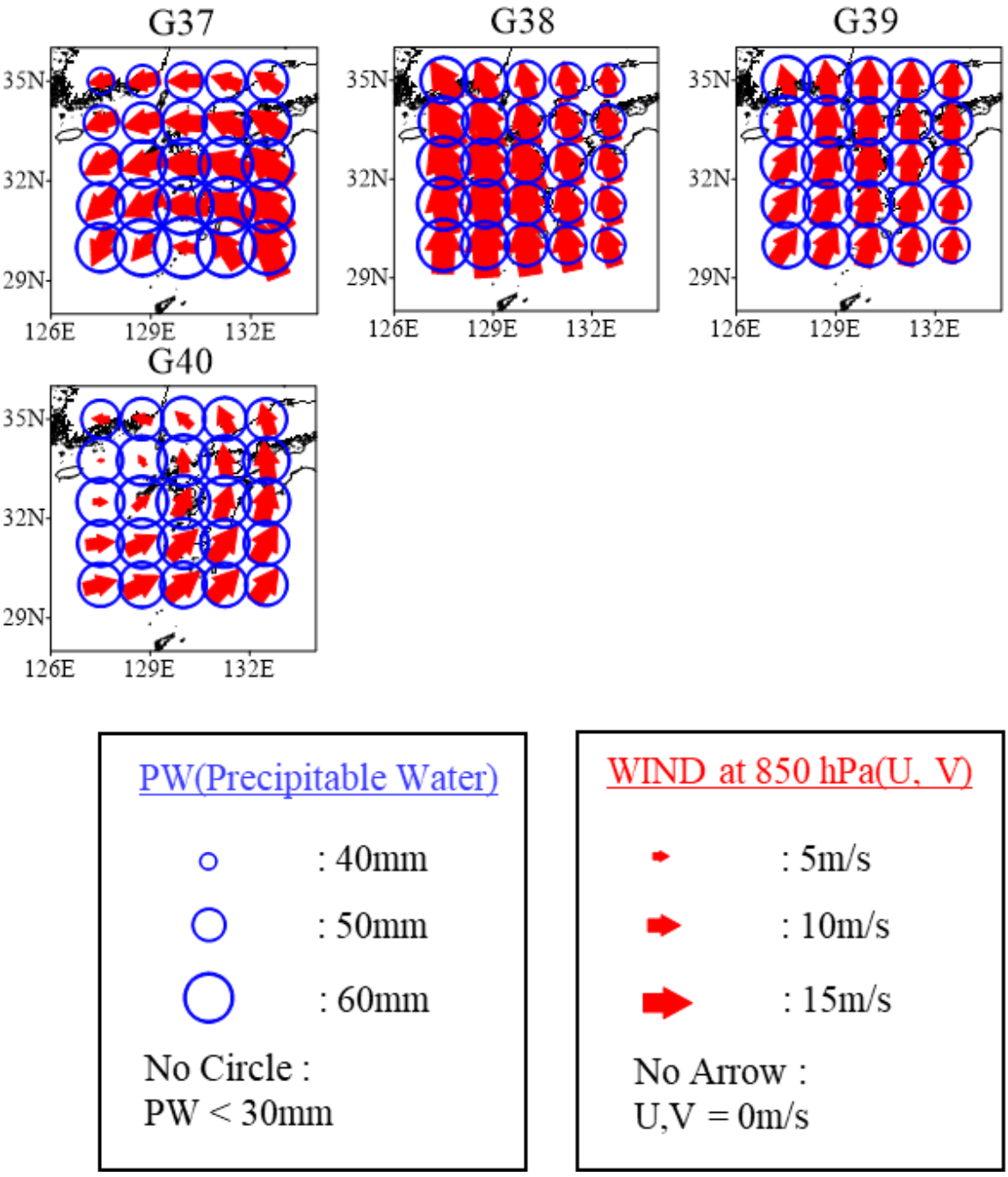


Figure A.5: The synoptic weather groups from G37 to G40

B: The SOM in each rainfall intensity (30, 70, 80 mm/h)

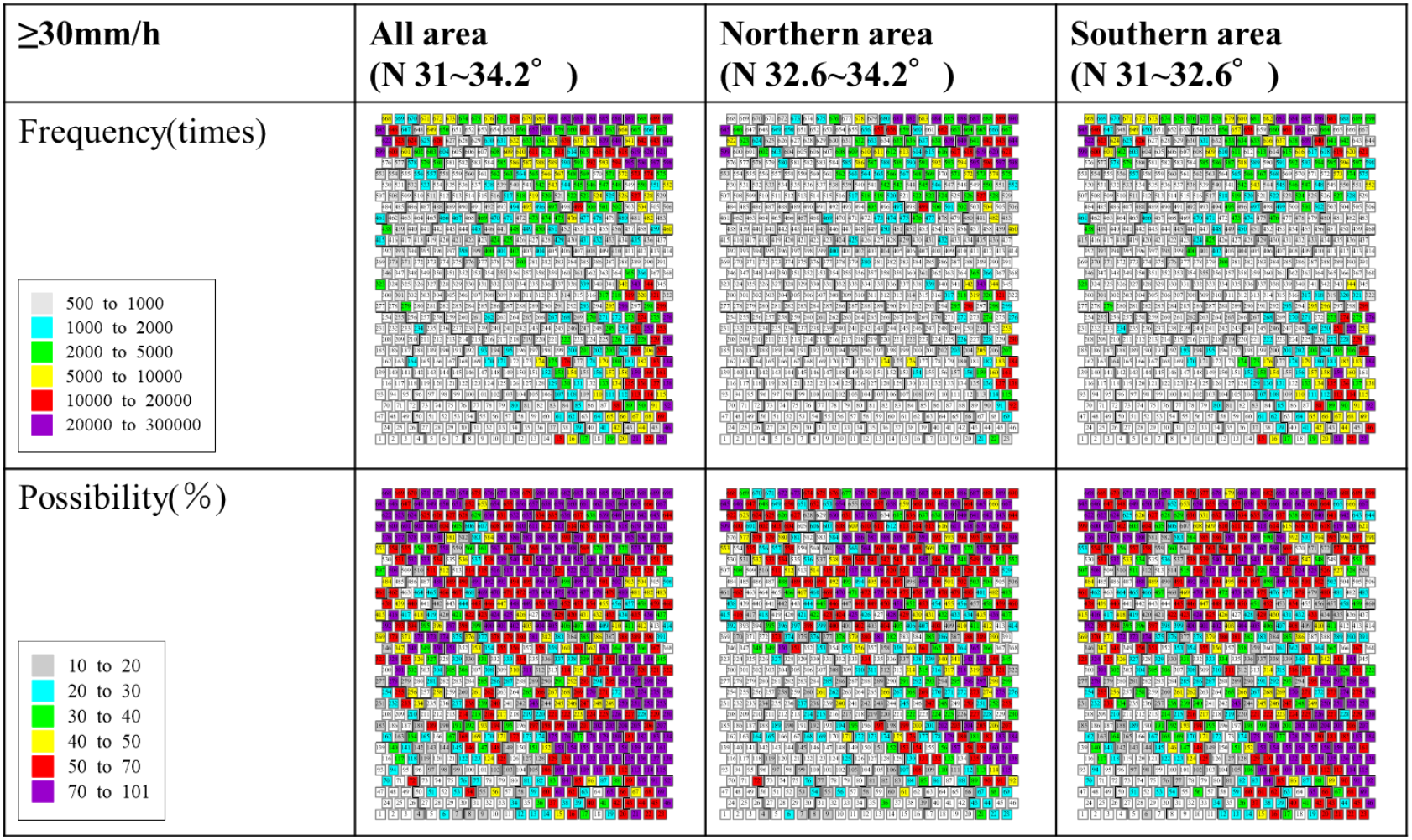


Figure B.1: Heavy rainfall frequency(upper) and possibility(lower) (≥30 mm per hour) for 40 years in each unit on the SOM

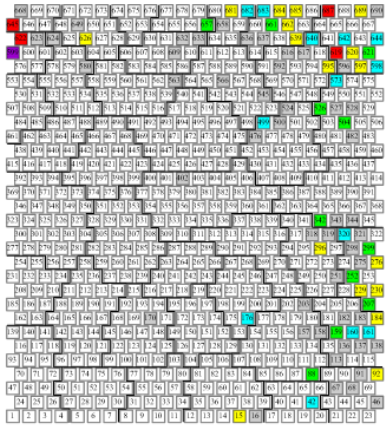
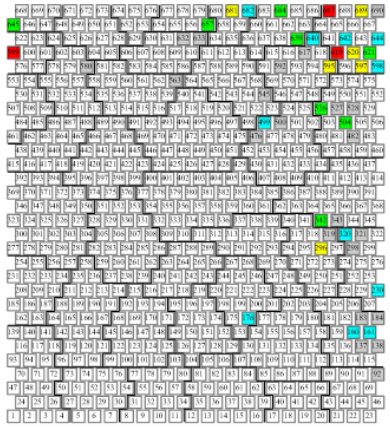
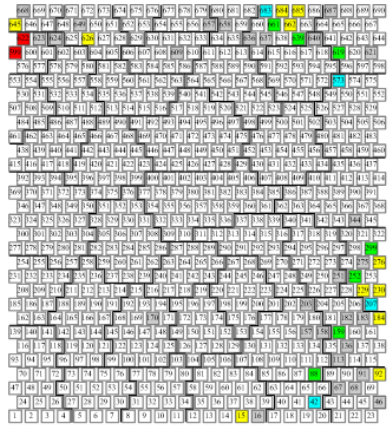
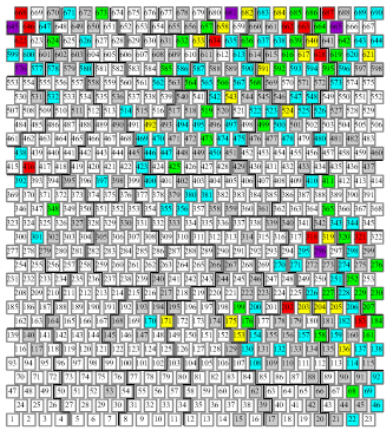
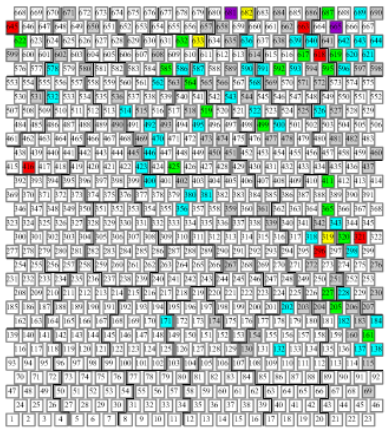
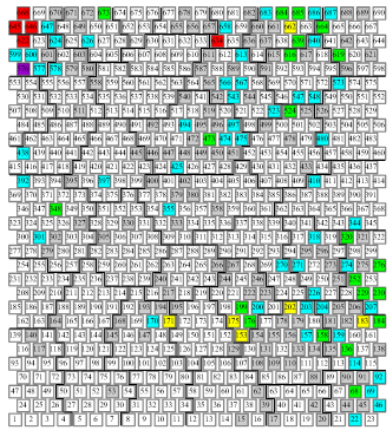
≥70mm/h	All area (N 31~34.2°)	Northern area (N 32.6~34.2°)	Southern area (N 31~32.6°)
Frequency(times) 100 to 300 300 to 500 500 to 1000 1000 to 3000 3000 to 5000 5000 to 30000			
Possibility(%) 10 to 20 20 to 30 30 to 40 40 to 50 50 to 70 70 to 101			

Figure B.2: Heavy rainfall frequency(upper) and possibility(lower) (≥70 mm per hour) for 40 years in each unit on the SOM

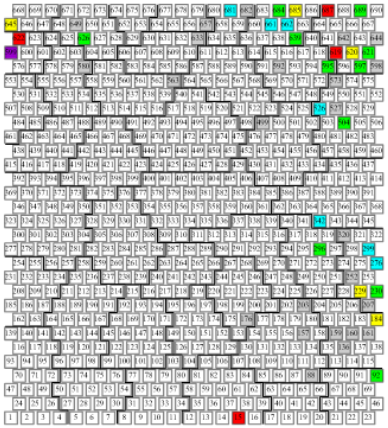
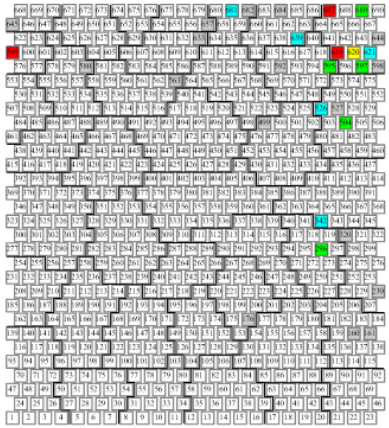
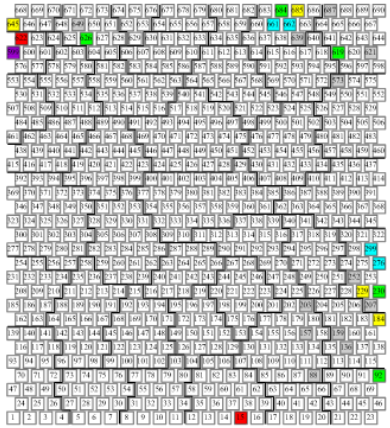
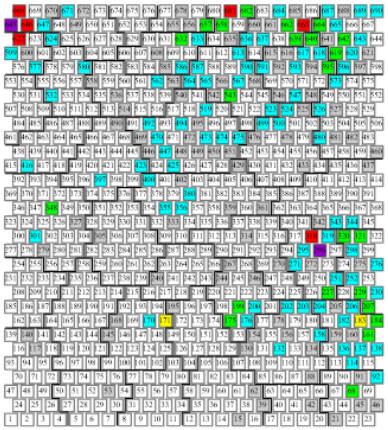
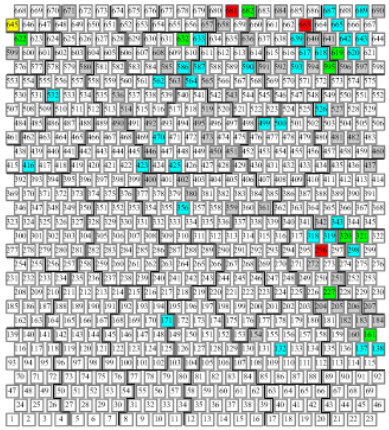
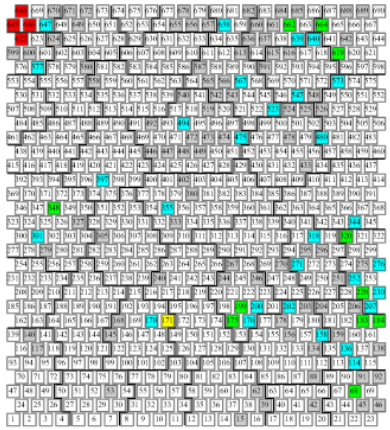
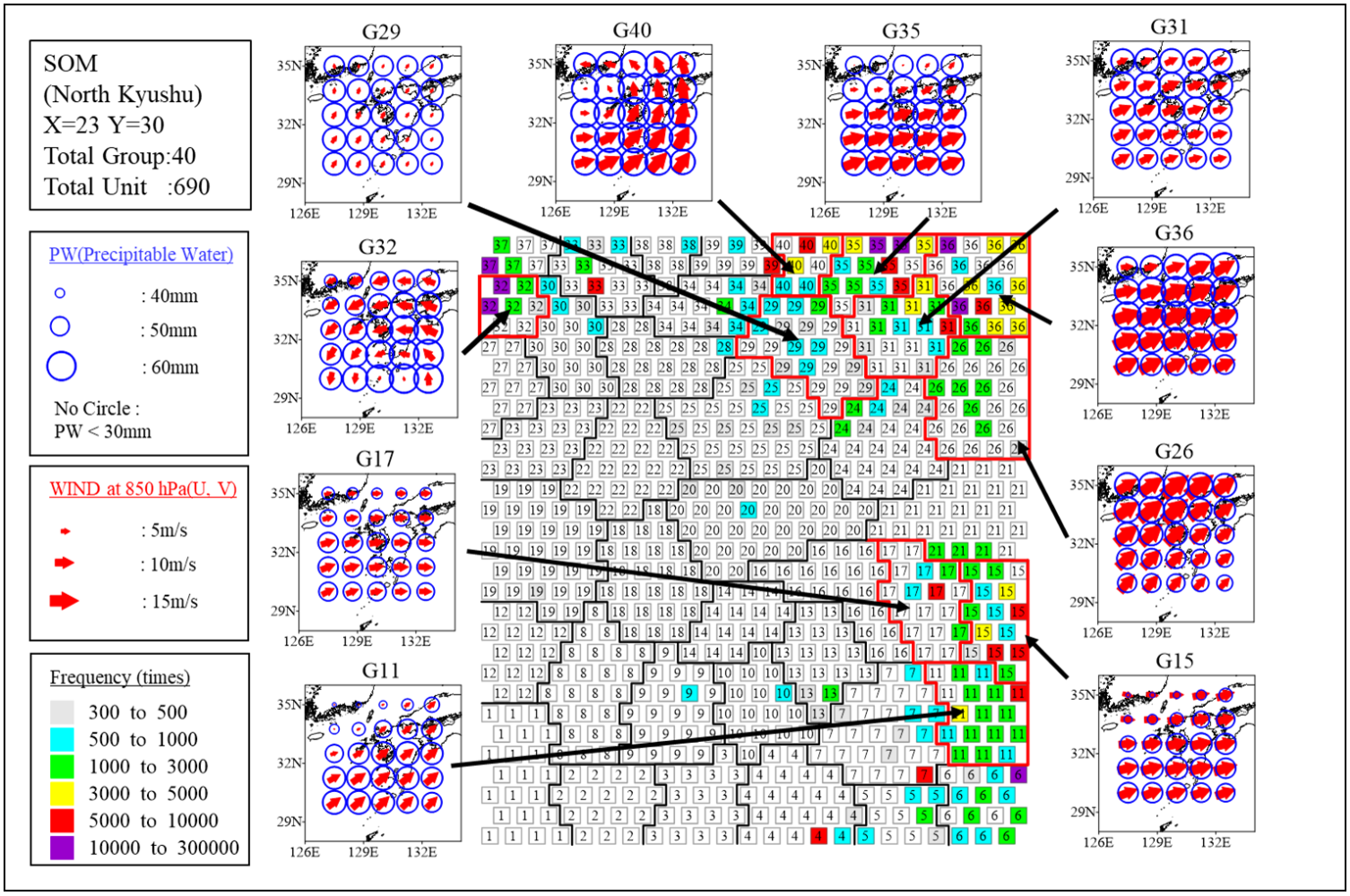
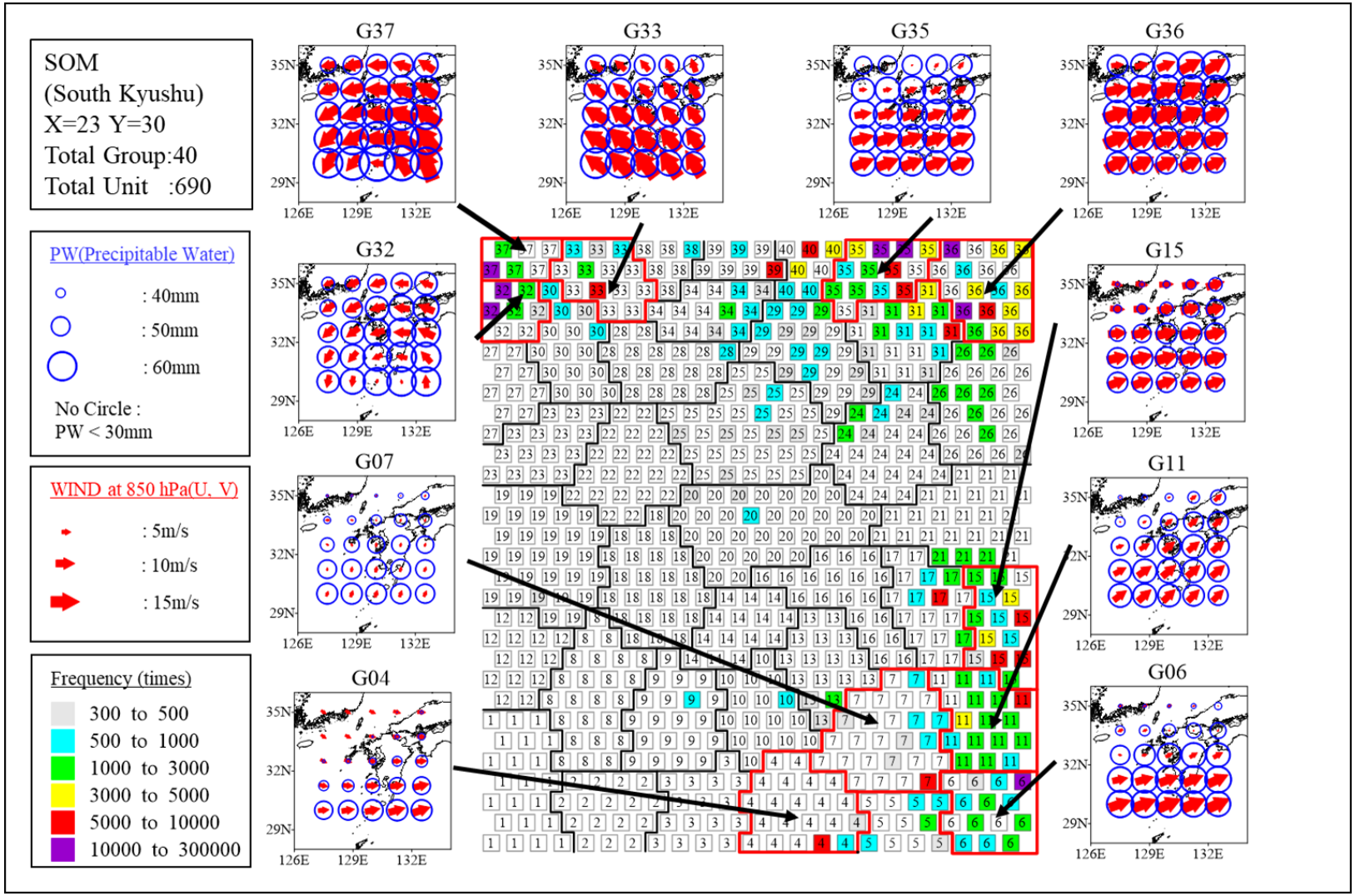
≥80mm/h	All area (N 31~34.2°)	Northern area (N 32.6~34.2°)	Southern area (N 31~32.6°)
Frequency(times) 100 to 300 300 to 500 500 to 1000 1000 to 2000 2000 to 3000 3000 to 30000			
Possibility(%) 10 to 20 20 to 30 30 to 40 40 to 50 50 to 70 70 to 101			

Figure B.3: Heavy rainfall frequency(upper) and possibility(lower) (≥80 mm per hour) for 40 years in each unit on the SOM

C: The SOM and the top 10 synoptic weather patterns in each area (50 mm/h)



FigureC.1: The SOM and the top 10 synoptic weather patterns in the northern Kyushu area



FigureC.2: The SOM and the top 10 synoptic weather patterns in the southern Kyushu area

D: The probability leading to heavy rainfall events exceeding each rainfall intensity (30, 70, 80 mm/h)

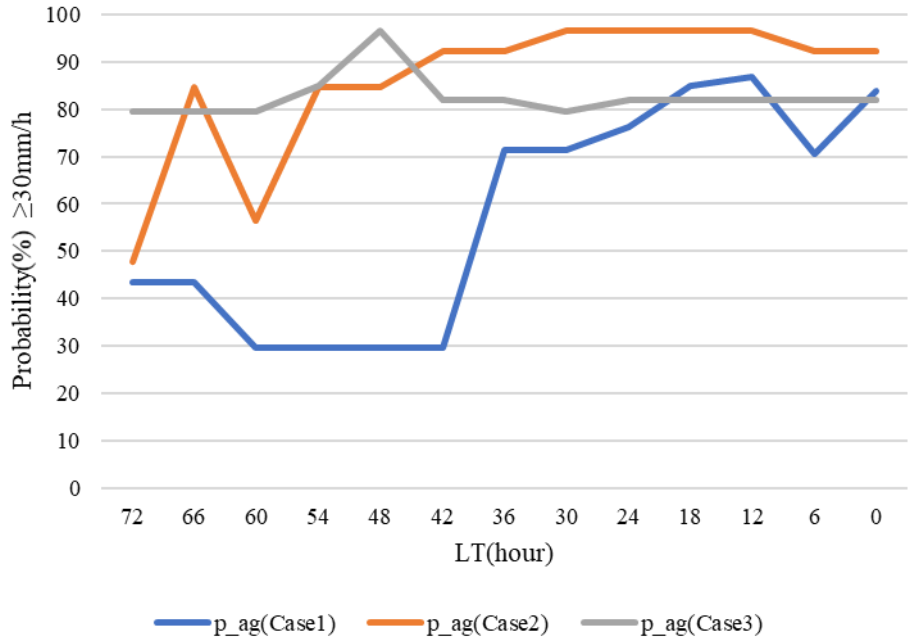


Figure D.1: The probability leading to heavy rainfall events exceeding 30 mm/h at heavy rainfall time

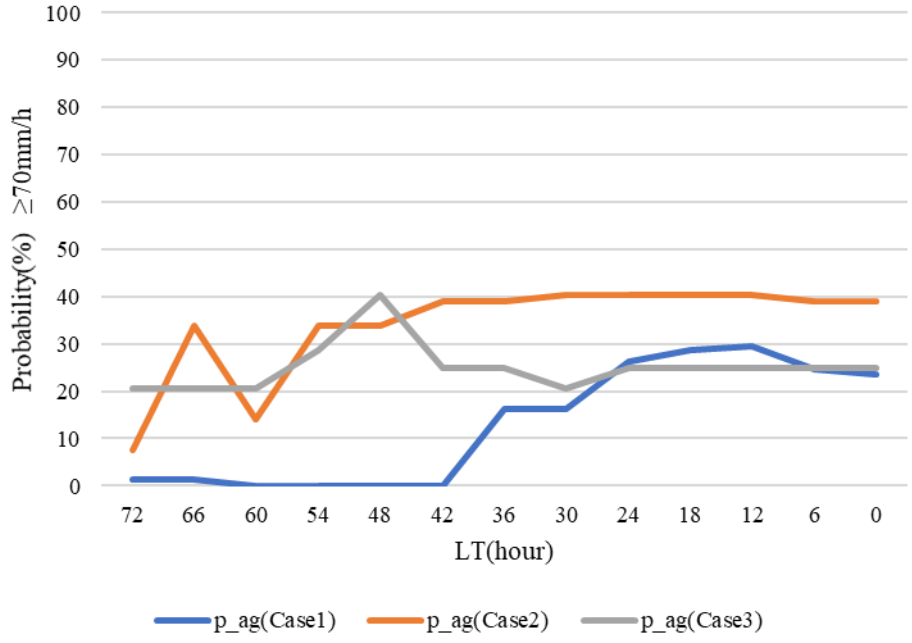


Figure D.2: The probability leading to heavy rainfall events exceeding 70 mm/h at heavy rainfall time

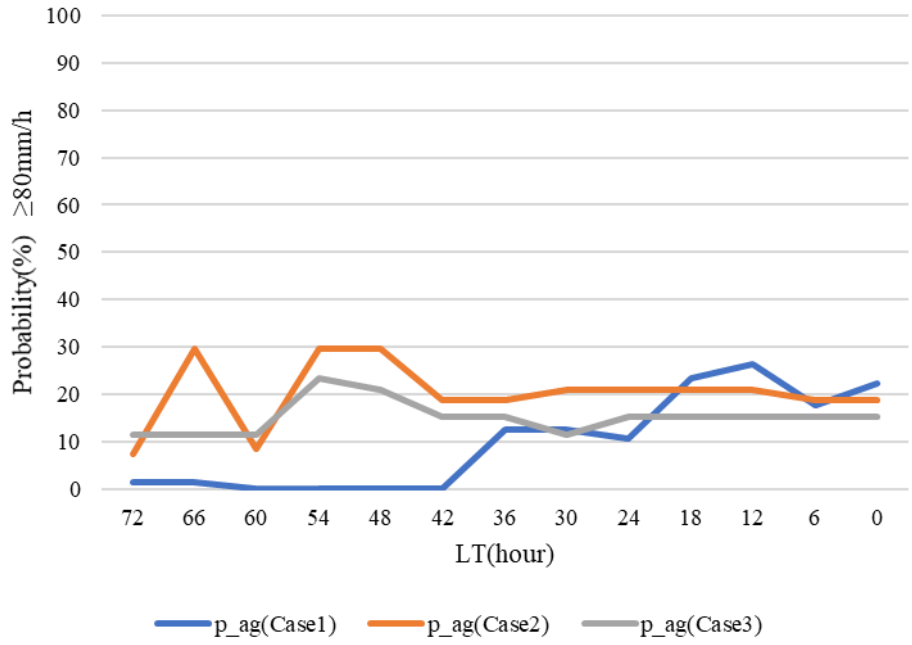


Figure D.3: The probability leading to heavy rainfall events exceeding 80 mm/h at heavy rainfall time