Frequency analysis of heavy rainfall related to synoptic weather patterns in Kyushu Island, Japan by using the Self-Organizing Map (SOM)

Norihiro Ohashi

Division of Water Resources Engineering
Department of Building and Environmental Technology
Lund University
Frequency analysis of heavy rainfall related to synoptic weather patterns in Kyushu Island, Japan by using the Self-Organizing Map (SOM)

By:
Norihiro Ohashi

Master Thesis
Division of Water Resources Engineering
Department of Building & Environmental Technology
Lund University
Box 118
221 00 Lund, Sweden
Water Resources Engineering
TVVR-18/5005
ISSN 1101-9824

Lund 2018
www.tvrl.lth.se
English title: Frequency analysis of heavy rainfall related to synoptic weather patterns in Kyushu Island, Japan by using the Self-Organizing Map (SOM)

Author(s): Norihiro Ohashi
Supervisor: Cintia Bertacchi Uvo
Examiner: Ronny Berndtsson
Language: English
Year: 2018
Keywords: Heavy rainfall frequency, Self-Organizing Map, Pattern recognition, Meteorological fields
Acknowledgement

This paper is a product of my academic experiences and lessons throughout my master courses at both of the department of Urban and Environmental System Engineering, Kyushu University and the department of Building and Environmental Technology, Water Resources Engineering, Lund University.

First and foremost, I would like to express my greatest gratitude to my supervisor, Assistant professor Koji Nishiyama. Mr. Nishiyama has given me numerous significant supports and plenty of advices throughout a year since I came back to Japan from Sweden. I learned a lot of knowledge about meteorology from him and his enthusiasm towards research has always inspired me. I couldn’t have done this study without his supervision.

I also would like to show my greatest gratitude to Associate Professor Yoshinari Hiroshiro. After I had presentations, Mr. Hiroshiro gave me accurate advices which surely boosted the quality of my presentation documents.

From my study abroad experience in Lund University, I would like to show my huge appreciation to all the teachers I met. Initially, I totally couldn’t catch up the classes I took and smoothly discuss with my classmates due to the lack of English skills. In that challenging situation, all the teachers didn’t think less of me and abandon me. Always, they gave me right advices. Because of that, I was able to gradually adapt to a new life and concentrate on my study without of too much anxiety. Thank you very much, although I cannot show this appreciation to them in person.

Finally, I would like to appreciate all my colleagues for discussions and helping me in the lab.
Abstract

Recently heavy rainfall frequency in Kyushu Island, the southwestern part of Japan, during the rainy season, has been increasing due to possibly climate change, global warming. Owing to the recent increase of heavy rainfall frequency, natural disasters related to heavy rainfall including flooding and landslides have occurred and caused serious damages to infrastructures and human lives. Therefore, for recognizing the possibility and the risk of heavy rainfall related natural disasters, it is very important to identify and investigate that what kinds of meteorological fields have frequently caused heavy rainfall so far and significantly contributed to the recent increase of heavy rainfall frequency, specifically in Kyushu Island.

Hence, the objective of this study is to reveal the aforementioned meteorological fields by correlating annual variations in heavy rainfall frequency measured for the study period of 39 years with synoptic weather groups classified by the Self-Organizing Map (SOM) developed by Kohonen (1995) and subsequent two clustering techniques (Ward method and K-means method). The SOM algorithm is an unsupervised neural network algorithm that can non-linearly classify complex high-dimensional input data into visually-understanding patterns on a two dimensional map.

In this study, 19032 meteorological fields recorded for 39 years were classified into 45 synoptic weather groups depending on the similarity. Subsequently, annual variations in heavy rainfall frequency corresponding to 45 synoptic weather groups have been investigated. The result shows that stationary front, typhoon, a combination of stationary front and low pressure, and moist tongue have frequently caused heavy rainfall in Kyushu Island so far. Moreover, these meteorological fields have highly contributed to the recent increase of heavy rainfall frequency. Therefore, when those meteorological fields appear in Kyushu Island, it is considerably important to carefully check weather conditions for recognizing the possibility and risk of natural disasters associated to heavy rainfall. Note that whereas heavy rainfall frequency has increased recently in those meteorological fields, the increase trend of appearance frequencies of those fields cannot be recognized because the occurrence of heavy rainfall is often controlled by meteorological factors such as regional topographical features. Therefore, further investigations on smaller-scales will be needed in the future work.
# Table of Contents

1 Introduction ......................................................................................................................... 1  
   1.1 Background ..................................................................................................................... 1  
   1.2 Objective ......................................................................................................................... 3  
   1.3 Application of SOM to synoptic Meteorology ................................................................. 4  
2 Clustering techniques .......................................................................................................... 5  
   2.1 The Self-Organizing Map (SOM) ...................................................................................... 5  
   2.2 QE (Quantization error) and TE (Topographic error) ....................................................... 9  
   2.3 Ward method .................................................................................................................... 10  
   2.4 K-means method .............................................................................................................. 13  
   2.5 DBI (Davies-Bouldin Index) ............................................................................................. 13  
   2.6 Automated Meteorological Data Acquisition System (AMeDAS) ..................................... 14  
3 Methodology ......................................................................................................................... 15  
   3.1 Overview of methodology in this study ............................................................................ 15  
   3.2 Specification of metrological fields ................................................................................... 16  
   3.3 Annual variations in heavy rainfall frequency ................................................................. 17  
   3.4 Heavy rainfall frequency corresponding to meteorological field ...................................... 20  
   3.5 Determination of the SOM structure ................................................................................ 20  
4 Result ................................................................................................................................... 24  
   4.1 SOM structure used for this study ................................................................................... 24  
   4.2 Heavy rainfall frequency in the map ................................................................................. 25  
   4.3 Annual variation in heavy rainfall frequency for each group ............................................ 28  
   4.4 Details of heavy rainfall frequency in G16, G44 and G29 ............................................... 31  
   4.5 Annual variation in appearance frequency of the top 3 synoptic weather groups (G16, G44 and G29) ................................................................................................................. 34  
5 Conclusion ............................................................................................................................. 36  
References ............................................................................................................................... 37  
Appendix ................................................................................................................................. 39
Chapter 1
Introduction

1.1 Background

Heavy rainfall frequency has been increasing recently on a global scale. Many studies have suggested that climate change including global warming can be considered as one of possible reasons of this increase. The third report (IPCC, 2001) of the Intergovernmental Panel on Climate Change suggested that heavy rainfall and its intensity have increased on the time scale of several decades to a century in numerous regions all over the world. Until the present, extreme heavy rainfalls have increased with the increase of surface air temperature in many regions of the world (e.g., Trenberth, 2011).

In Japan, the increase of rainfall frequency has been recognized over the past decade. The investigation of long-term changes in extreme hourly rainfall in Japan conducted by Fujibe (2015) found that the increase in extreme rainfall from 1979 to 2013 was recognized all over Japan, and the increase is related to the increase in air temperature and sea-surface temperature. Moreover, JMA (2016) investigated the annual number of events with hourly rainfall of $\geq 50$ mm/h and $\geq 80$ mm/h, respectively, as shown by Figure 1.1. These data have been observed from 1976 to 2016 at approximately 1000 regional meteorological observation stations located all over Japan, which is known as the Automated Meteorological Data Acquisition System (AMeDAS). The result shows the frequency of heavy rainfall events have gradually increased during past 40 years after the onset of AMeDAS observation.

Figure 1.1: Annual number of events with hourly precipitation of $\geq 50$mm/h and $\geq 80$mm/h per 1000 AMeDAS stations depicted by JMA (2016). The annual numbers of events are converted to those per-1000-stations because the number of AMeDAS stations has increased until the present.
Especially, over Kyushu Island (target area for this study) which is located in the western part of Japan, intense heavy rainfalls have frequently occurred during the mature stage of the rainy season (Ninomiya and Mizuno, 1987). Figure 1.2 describes the annual variation in heavy rainfall frequency (≥ 70 mm/h) in Kyushu Island, based on the data from AMeDAS stations that have observed continuously during for 39 years (1979-2017). The result shows drastic increase trend of heavy rainfall frequency in Kyushu Island.

![Figure 1.2: Annual variation in heavy rainfall frequency (≥ 70 mm/h) during a warm season over the last 39 years in Kyushu Island](image)

During the rainy season, heavy rainfall frequently occurs along or in the southern region of the stationary front, which has west-to-east oriented rainfall zone characterized by the supply of warm and humid air into the front with the enhancement of atmospheric instability. The rainy season inherent in Japan usually from June to July is called as the Baiu. The Baiu stationary front is recognized in the middle of May in the southern area of Japan between the wet Pacific high pressure region and dry high pressure region located in the north. The front gradually shifts northward from June to July. In the end of July, it disappears in the northern part of Japan with the expansion of the Pacific high pressure region towards Japan islands (Ninomiya and Murakami, 1987).

Under these circumstances in Japan, natural disasters including floods and landslides have often occurred especially during the rainy season and caused serious damages to infrastructures and human lives during the Baiu season. For example, in July 2012, heavy rainfall occurred with maximum intensity of approximately 800 mm/day and 108 mm/h in the northern Kyushu. This heavy rainfall induced serious damages in Fukuoka, Kumamoto and Oita Prefecture. Moreover, in July 2017, heavy rainfall (referred to as Northern Kyushu Heavy rainfall) continued for 9 hours in the specific area of north Kyushu. The event
brought extremely serious damages in both Fukuoka and Oita prefecture. The recorded maximum daily and hourly rainfall (AMeDAS) were 545 mm/day and 129.5 mm/h, respectively, in Asakura City of Fukuoka prefecture. Moreover, the maximum daily rainfall analyzed by both of ground rainfall and meteorological radar shows the record of approximately 1000 mm. The daily record is comparable to approximately 50% of annual rainfall (2011 mm) observed in 2017 in Asakura city.

Focusing on the issue of future global warming in the world, it has been reported that there are strong relationships relation between global warming and rainfall increase. The fourth assessment report of IPCC (2007) indicated that the estimation of the global average surface air temperature is within approximately 1.8-4.0 degree C during the 21st century depending on the increase of greenhouse gases such as CO2, NH4. Moreover, the fifth assessment report of IPCC (2013) indicated that a large number of extreme weather and climate change have been observed since 1950 and, in the future, the frequency and intensity of heavy rainfall will become higher in many regions all over the world.

For the influence of future global warming on rainfall properties in Japan, the climate projection for the 21st century conducted by Kimoto et al. (2005) shows 10% increase of mean precipitation compared with the present in warm seasons. The climate simulation based on A1B scenario of Kusunoki and Mizuta (2008) shows the increase of precipitation intensity in Western Japan with the tendency to delay the termination of the Baiu season in the future. The current and future climate projection of the most serious scenario RCP8.5 using the regional spectral model (RSM) downscaled over East Asia by Ham et al. (2016) showed the enhancement of the East Asian monsoon and the Pacific subtropical high pressure system leads to rainfall increase in summer. Moreover, some studies reported that the increase of heavy rainfall events is caused by global warming (e.g. Meehl et al., 2005; Yoshizaki et al., 2005).

The results of these climate projections are related to the enhancement of stationary frontal activity in Japan Island during the rainy season in the future. In other words, this implies the increase of water vapor flux, which is further supply of a large amount of water vapor into Japan Island with the enhancement of the East Asian monsoon. Therefore, it may be expected that the increase trend shown in Figure 1.2, and recent serious natural disasters occurring in Kyushu Island are caused by global warming. However, it is premature to clearly recognize these issues as the influence of global warming because recent increase of heavy rainfall frequency may incorporate the influence of climate change within a decadal scale.

1.2 Objective

Based on the above-mentioned background, it is very significant to monitor current patterns of weather situations related to the occurrence of heavy rainfall and associated natural disasters. Therefore, in this study, 1) a large number of meteoro logical fields observed for several decades are classified into synoptic weather patterns by the Self-Organizing Map (SOM) developed by Kohonen (1995). In the next step, 2) the
Synoptic weather patterns are divided into groups by subsequent clustering techniques consisting of Ward method and \( K \)-means method. Finally, 3) a time-series in heavy rainfall frequency obtained from AMeDAS in Kyushu Islands is divided into synoptic weather groups.

Through these procedures, the main objectives of this study are to identify what kinds of meteorological fields have frequently caused heavy rainfall so far, and what kinds of meteorological fields have significantly contributed to the recent increase of heavy rainfall frequency in Kyushu Island by quantitatively investigating significant relationships between the synoptic weather groups and associated annual variations in heavy rainfall frequency, according to Hewitson and Crane (2002), Crane and Hewitson (2003), and Alexandar (2010), which applied the SOM technique for evaluating time-series of synoptic weather patterns.

1.3 Application of SOM to synoptic Meteorology

In this study, Self-Organizing Map (SOM) is applied for the pattern recognition of synoptic weather affecting Kyushu Island in Japan. The SOM was developed by Kohonen (1995) as a kind of unsupervised artificial neural networks (ANNs) technique. The most important feature of the SOM is to extract visually- and easily-detectable information by classifying high-dimensional complicated data into two-dimensional patterns. Therefore, the SOM has been widely applied to many kinds of fields requiring pattern recognition techniques, including meteorological studies. The detail of the application of the SOM to meteorological studies is described by Sheridan and Lee (2011), Liu and Weisberg (2011), and Nishiyama et al. (2013).

The SOM is also available for synoptic meteorology, which investigates significant relationships between meteorological variables (e.g., heavy rainfall frequency, strong wind) in a target area and features of meteorological fields detected in a synoptic scale, as introduced by Hewitson and Crane (2002). For example, Nishiyama et al. (2007) classified the pattern of complicated synoptic weather situations and related them to the heavy rainfall frequency in western Japan. Hewitson and Crane (2002), and Crane and Hewitson (2003) applied the SOM to the analysis of time-series of meteorological local variables relating to synoptic weather.
Chapter 2

Clustering techniques

2.1 The Self-Organizing Map (SOM)

2.1.1 Basic information of the SOM

The Self-Organizing Map (SOM) was developed by Kohonen (1995) as an unsupervised Artificial Neural Network (ANN) to easily and visually understand complicated high dimensional data. Until the present, the SOM has become one of the most popular ANN algorithms in the field of information science. As for unsupervised learning, it aims to find structures and patterns in unlabeled data. No back-propagation is needed because the data is unlabeled without any input-output pairs. This type of learning is commonly used for clustering and data reduction in statistics. The SOM can project high-dimensional input data onto an arbitrary low-dimensional output space (usually it is two-dimensional and referred to as a map). On the map, the topology of high-dimensional input data is preserved. Therefore, by looking at the map, complicated high-dimensional input data can be visually and two-dimensionally interpreted.

2.1.2 The SOM network structure

Figure 2.1 describes the basic structure of the SOM network. The SOM network is composed of an input layer and a competition layer (called as the SOM or just a map). Input vectors classified through the SOM training in the competition layer are allocated to the input layer. On the other hand, the
competition layer consists of two dimensionally-arranged units as shown in Figure 2.1. Similar input vectors are classified into a specific unit. The unit is characterized by the weight vector, which represents average features (pattern) of input vectors classified into the unit, and has the same number of dimensions as the input vector. Therefore, the weight vector is useful and important for understanding features of input vectors in the unit.

Although there are several options for the shape of the units depending on a research topic, regularly-arranged hexagonal shape of units is widely used because the Euclidean distance between the a unit and the nearest six neighbor units is the same on the whole map, as illustrated in Figure 2.1. In this study, regular hexagonal shape of units is applied for the SOM training.

2.1.3 The SOM training

Data preparation before the SOM training

After determining the SOM size, the normalization of the input vectors and the initialization of the weight vectors on the SOM are implemented. The input vector \( \tilde{x}(t) \) and normalized input vector \( x(t) \) are given as Eqs. (2.1) and (2.2), respectively.

\[
\tilde{x}(t) = (\tilde{x}_1(t), \tilde{x}_2(t), \tilde{x}_3(t), \ldots, \tilde{x}_I(t), \ldots \tilde{x}_n(t))
\]

\[
x(t) = (x_1(t), x_2(t), x_3(t), \ldots, x_I(t), \ldots x_n(t))
\]

where subscript \( n \) is the number of elements in each input vector. \( t \) is the input vector’s number from \( t=1 \) to \( t=T \). \( T \) is the total number of input vectors used for the SOM training. \( x(t) \) for the SOM training will be obtained by normalizing \( \tilde{x}(t) \) with Eq. (2.3).

\[
x_I(t) = \frac{\tilde{x}_I(t) - \text{min}\{\tilde{x}_I(t)\}}{\text{max}\{\tilde{x}_I(t)\} - \text{min}\{\tilde{x}_I(t)\}}
\]

Subsequently, the initialization of the initial weight vectors \( m_i(0) \) is executed. For the determination of \( m_i(0) \), random numbers between 0 and 1 are given to all elements of \( m_i(0) \). The weight vectors \( m_i(t) \) while the training and updating are given as Eq. (2.4). The final weight vectors \( M_i \) after the training and updating are given as Eq. (2.5).

\[
m_i(t) = (m^i_1(t), m^i_2(t), m^i_3(t), \ldots, m^i_I(t), \ldots \ldots m^i_n(t))
\]
\[
M_i = (M^i_1, M^i_2, M^i_3, \cdots, M^i_1, \cdots M^i_n)
\] (2.5)

where superscript \(i\) is the unit’s number on the SOM. Here, the dimension of the weight vectors is the same as that of the input vectors.

**Training process**

As can be seen in Figure 2.2, the first step is to find out the best matching unit (so-called BMU). The BMU means a weight vector closest to the input vector. After calculating the Euclidean distance \(\|x(t) - m_i(t)\|\) between the input vector \(x(t)\) and all weight vectors \(m_i(t)\) on the SOM, the BMU closest to the input vector is determined by comparing among all the Euclidean distances, as described by Eq. (2.6).

\[
c = \arg \min_i \{\|x(t) - m_i(t)\| \}
\] (2.6)

---

**Start of SOM training**

- Normalization of the input vectors \((\tilde{x}(t) \rightarrow x(t))\) in the input layer
- Initialization of the weight vectors \((m_i(t))\) in the units on the SOM

**End of SOM training**

---

**Figure 2.2: Flow chart of the SOM training**
where \( c \) is the BMU. The next step is the modification of the weight vectors in the units. The weight vectors are respectively updated by repeating the training process through the presentation of the input vector, according to Eq. (2.7).

\[
\mathbf{m}_i(t + 1) = \mathbf{m}_i(t) + h_{ci}
\left(t_e, \| \mathbf{r}_c - \mathbf{r}_i \| \right) \left[ \mathbf{x}(t) - \mathbf{m}_i(t) \right]
\]  

(2.7)

where \( t \) is defined as the learning step. \( t_e \) is the epoch number and \( T_e \) is the maximum epoch number. 1 epoch is defined by 1 learning of all the input vectors from \( t=1 \) to \( t=T_e \). Until the training during one epoch has been finished, the neighborhood function expressed as \( h_{ci}(t_e) \) defined by Eq. (2.8) is unchanged. This means the degree of the modification of the weight vectors is constant through each epoch. For example, if 100 input vectors are trained by the SOM with the maximum number of \( T_e=100 \), \( t_e=1 \) corresponds to the learning from \( t=1 \) to \( t=100 \), \( t_e=2 \) corresponds to the learning from \( t=101 \) to \( t=200 \). The learning will be continued until \( T_e=100 \) (from \( t=9901 \) to \( t=10000 \)). In other words, even if the number of input samples is 100, the total number of the training step becomes 10000 with the maximum number of \( T_e=100 \).

\[
h_{ci}(t_e, \| \mathbf{r}_c - \mathbf{r}_i \|) = \alpha(t_e) \cdot \exp \left( -\frac{\| \mathbf{r}_c - \mathbf{r}_i \|^2}{2\sigma^2(t_e)} \right) \quad (\| \mathbf{r}_c - \mathbf{r}_i \| \leq \sigma(t_e))
\]

\[
h_{ci}(t_e, \| \mathbf{r}_c - \mathbf{r}_i \|) = 0.0 \quad (\| \mathbf{r}_c - \mathbf{r}_i \| > \sigma(t_e))
\]

(2.8)

where \( \mathbf{r}_c \) and \( \mathbf{r}_i \) are the location vectors of units \( c \) and \( i \), respectively. \( \alpha(t_e) \) is the learning rate and \( \sigma(t_e) \) is the neighborhood radius. Based on the neighborhood function shown by Eq. (2.8), the weight vectors in the neighbor units of the BMU within the radius of \( \sigma(t_e) \) are modified. On the other hand, the weight vectors outside the radius of \( \sigma(t_e) \) are not modified. In this study, \( \alpha(t_e) \) and \( \sigma(t_e) \) are given by Eq. (2.9).

\[
\alpha(t_e) = \max \left\{ \alpha(0) \frac{T_e - t_e + 1}{T_e}, \ 0.005 \right\}
\]

\[
\sigma(t_e) = \max \left\{ \sigma(0) \frac{T_e - t_e + 1}{T_e}, \ 1.1 \right\}
\]

(2.9)

As can be seen in Eq. (2.9), once the learning rate and neighborhood radius fall below 0.005 and 1.1, these values are fixed as 0.005 and 1.1, respectively. This means that the degree of the modification of the weight vectors becomes gradually weaker and then becomes constant with the increase of the epoch number. At the same time, the range of the modification also becomes smaller and then gets fixed as the epoch number increases. In the final phase, only the weight vector in the BMU and the weight vectors in the closest neighbor units of the BMU are modified. In this study, the initial learning rate, the initial
neighborhood radius, and the maximum epoch number $T_e$ are set as 0.2, 5 and 50, respectively.

Figure 2.3 shows the interpretation of the SOM after finishing the training. Each unit can be interpreted as a pattern characterized by a weight vector. Similar input vectors are classified in the unit. Therefore, the weight vector has an average feature of the similar input vectors. Looking at the map, it can be interpreted that neighboring units are similar to each other while distant units are dissimilar. In other words, similarity and dissimilarity among input data or patterns (units) can be easily and visually interpreted on the two-dimensional map.

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

There are some indicators for measuring the performance of the SOM training. In our study, two indicators, Quantization error (QE) and Topographic error (TE), are utilized for determining the optimal map size.

QE is a measurement of the average Euclidean distance between an input vector and the BMU for QE can be calculated, as shown in Eq. (2.10).

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

$TE$ measures the topology preservation of the output layer (map), and shows the proportion of input vectors whose first and second BMUs are not adjacent, as given in Eq. (2.11). $u(x_i)$ is 1 if the first and second BMUs are not adjacent, otherwise 0.

$$TE = \frac{1}{N}\sum_{i=1}^{N} u(x_i) \quad (2.11)$$

In general, the decrease of $QE$ and $TE$ with the iteration step through the SOM training proceeds towards the good performance of the SOM classification. However, larger map size leads to higher $TE$ although $QE$ becomes lower. Therefore, we need to determine an optimal map size by finding suitable combination of $QE$ and $TE$.

### 2.3 Ward method

Hierarchical clustering can be mainly divided into two types, which are called as agglomerative clustering and divisive clustering. In this study, Ward method, one of agglomerative clustering, is used for the clustering of many patterns formed by the SOM training. This method does not need to specify the number of clusters in advance, which is the opposite of $K$-means method mentioned in the next subsection. On the other hand, the method is not suitable for the clustering of a vast number of data.

The agglomerative clustering works in a bottom-up manner, which means that it groups small clusters into larger clusters. This process is executed according to Figure 2.4, which consists of eight steps 1~8. The result of clustering can be finally depicted as a tree structure (dendrogram). In the process, the closest clusters are merged into the larger clusters. As can be confirmed by Figure 2.5, the vertical scale in the dendrogram expresses the distance between merged clusters.

Agglomerative clustering steps are as follows:

1. Each element basically begins as one cluster at the initial step. At step 1, element 1 can be considered as a cluster of $C_1$.
2. To find the closest pair of clusters. At step 2, cluster $C_1$ and $C_2$ are closest each other among the other clusters. The way to identify the closest pair is to find the minimum distance by calculating the distance between two clusters.
3. To merge the closest pair of clusters into a bigger cluster. At step 3, cluster $C_1$ and $C_2$ are merged into cluster $C_{12}$.
4. The procedures of (1)-(3) are repeated until there is only one cluster. Consequently, $C_{12345}$ is the final
cluster.

(5) A dendrogram (tree structure) is depicted as the result of clustering. The clustering process can be confirmed in the dendrogram.

Figure 2.4: An example of agglomerative clustering process

In Ward method, the distance between two clusters (assume cluster $C_a$ and $C_b$) is defined as how much the sum of squares will increase when those two clusters are merged. Eq. (2.12) describes the distance between two clusters (cluster $C_a$ and $C_b$). In addition, Figure 2.6 shows a visual illustration of the pair of two
clusters (cluster \( C_a \) and \( C_b \)). If the distance is closest, the two clusters are merged into a bigger cluster.

\[
\Delta(C_a, C_b) = \sum_{i \in C_a \cup C_b} \left\| \mathbf{x}_i - \frac{m_{C_a \cup C_b}}{n_a + n_b} \right\|^2 - \sum_{i \in C_a} \left\| \mathbf{x}_i - \frac{m_{C_a}}{n_a} \right\|^2 - \sum_{i \in C_b} \left\| \mathbf{x}_i - \frac{m_{C_b}}{n_b} \right\|^2
\]

\[= \frac{n_a n_b}{n_a + n_b} \left\| m_{C_a} - m_{C_b} \right\|^2 \quad (2.12)\]

where \( \Delta \) is the merging cost of combining two clusters. \( C_j \) is the name of cluster \( j \). \( m_{C_j} \) is the center of cluster \( j \). \( n_j \) is the number of merged elements (or merged clusters) in cluster \( j \). \( \mathbf{x}_i \) and \( \mathbf{x'}_i \) are the vector and the location vector of an merged element \( i \) (or an merged cluster \( i \)) in cluster \( j \), respectively. \( \| \cdot \|^2 \) is Euclidean distance. When \( \Delta \) is minimum, which means that the distance between two clusters is closest.

**Figure 2.5:** An example of dendrogram corresponding to Figure 2.4

**Figure 2.6:** Distance between two clusters (cluster \( C_a \) and \( C_b \))
2.4 K-means method

K-means algorithm is one of non-hierarchical clustering, which is available for the clustering of a vast number of data unlike the hierarchical clustering. However, the drawback of the K-means is that the number of clusters and the extent of each cluster must be predetermined arbitrarily. The calculation of the K-means is as follows;

1) The number of clusters $K$ and the extent of each cluster are predetermined.
2) According to Eq. (2.13), the mean weight vector $C_j(t')$ in a cluster $(j)$ consisting of the units of $N_j(t')$ is calculated.

$$C_j(t') = \frac{1}{N_j(t')} \sum_{M_i \in Q_j(t')} M_i \quad (j = 1, 2, \ldots, K) \quad (2.13)$$

where $N_j(t')$ is the number of units in a cluster $(j)$, $j$ is the cluster number from $j=1$ to $j=K$. $Q_j(t')$ is the set of units in a cluster $(j)$. $t'$ is the iteration step.

3) According to Eqs. (2.14) and (2.15), all units are relocated into the most similar cluster, which is either in or outside the current cluster. The process is to calculate the Euclidean distance $\|M_i - C_j(t')\|$ between the weight vector in the unit $(i)$ and the mean weight vector in every cluster $(j=1\ldots K)$, and to find cluster $c$ which has the minimum distance closest to unit $(i)$. Then, unit $(i)$ is relocated into the cluster of $c$.

$$d_i(j) = \|M_i - C_j(t')\| \quad (2.14)$$

$$c = \arg \min_j \{d_i(j)\} \quad (2.15)$$

4) The computation procedures of Eqs. (2.13), (2.14) and (2.15) are iterated with the iteration step $t'$ until the relocation of all units to another cluster is finished. In this way, all the units are grouped into several clusters (groups).

2.5 DBI (Davies-Bouldin Index)

DBI index was suggested by Davies and Bouldin (1979) for finding optimum clustering performance. Smaller DBI leads to better clustering. Therefore, a clustering result with the lowest DBI is selected as the optimal number of clusters. The equations for calculating the DBI are as follows;
$$\text{DBI} = \frac{1}{K} \sum_{j=1}^{K} R_j$$  \hspace{1cm} (2.16)

$$R_j = \max_k \left\{ \frac{\left( S_j + S_k \right)}{D_{jk}} \right\} \hspace{0.5cm} (j \neq k)$$  \hspace{1cm} (2.17)

$$S_j = \frac{1}{N_j} \sum_{M_i \in Q_j} \left\| M_i - \tilde{C}_j \right\| \hspace{0.5cm} (j = 1, \ldots, K)$$  \hspace{1cm} (2.18)

$$D_{jk} = \left\| \tilde{C}_j - \tilde{C}_k \right\|$$  \hspace{1cm} (2.19)

where $M_i$ is the weight vector of the unit $(i)$ located in a cluster $(j)$. $\tilde{C}_j$ is the mean weight vector in the units of $N_j$ in a cluster $(j)$. Here, a cluster $(j)$ is the cluster determined after the above mentioned procedures of Ward and K-means are terminated. $\| M_i - \tilde{C}_j \|$ denotes the Euclidean distance between $M_i$ and $\tilde{C}_j$. $N_j$ is the number of units in a cluster $(j)$. $\| \tilde{C}_j - \tilde{C}_k \|$ is the Euclidean distance between $\tilde{C}_j$ and $\tilde{C}_k$.

$S_j$ indicates the degree of the homogeneity of the units in a cluster $(j)$. A value of $S_j$ becomes smaller when the homogeneity of the units in a cluster $(j)$ is high. On the other hand, $D_{jk}$ indicates the degree of the dissimilarity between two clusters $(j$ and $k)$. A value of $D_{jk}$ becomes larger when the similarity between cluster $(j)$ and cluster $(k)$ is small. To put it briefly, a smaller value of DBI means that all units included in a cluster are more similar to each other, and features among clusters on the SOM are more different each other.

### 2.6 Automated Meteorological Data Acquisition System (AMeDAS)

AMeDAS is Automated Meteorological Data Acquisition System, which shows a ground observational meso-network covering the Japan islands, with a mean resolution of about 17 km. AMeDAS data comprise rainfall, wind, temperature, and sunshine duration at 10-min intervals. The AMeDAS observation consisting of 800 stations started in 1976. The number of stations has gradually increased to approximately 1300 in 2016. Therefore, rainfall data during 40 years in Japan are available for constructing an annual variation in heavy rainfall frequency.
Chapter 3
Methodology

3.1 Overview of methodology in this study

Figure 3.1 illustrates an overview of the methodology treated in this study. A large number of metrological fields observed for several decades are classified into synoptic weather groups depending on the similarity of metrological fields by the Self-Organizing Map (SOM) and subsequent clustering techniques (Ward method and K-means method). Finally, an annual variation in heavy rainfall frequency observed from AMeDAS is divided into the synoptic weather groups classified by the SOM and the two clustering methods, as shown by the lower part of Figure 3.1.

Figure 3.1: Overview of methodology in this study
3.2 Specification of metrological fields

Before the SOM training, pre-processing procedures of meteorological fields are required for making an input vector of the SOM. In this study, meteorological fields are extracted on the basis of Nishiyama et al. (2007), which represented a large amount of moisture with low level jet (LLJ), a front formed between dry and wet masses, and rotation of a low pressure system such as typhoon by simply using precipitable water ($PW$) and wind components in lower layers. Therefore, according to Nishiyama et al. (2007), meteorological fields used for input vectors are represented by three components; (1) $U$ m/s (East-West wind speed at the 850hPa level), (2) $V$ m/s (North-South wind speed at the 850hPa level), and (3) Precipitable water (PW mm), which is defined by the water vapor amount contained in a vertical column of the atmosphere. The target area (Figure 3.2) for extracting rainfall data in this study is Kyushu Island located to the west of Japan. The area has been highly affected by stationary fronts and typhoons and associated natural disasters.

Meteorological data used for this study are obtained from the NCEP/NCAR reanalysis dataset (Kalnay et al., 1996). The dataset are recorded 4 times (3, 9, 15 and 21 JST: JST is UTC+9 hours) per day, every 6 hours. This shows a multi-dimensional grid dataset across the globe describing the condition of the earth’s atmosphere. This study extracts meteorological fields ($PW$, $U$ and $V$) consisting of 16 grid points.
(2.5 grid interval) with the extent of longitudes (125.0° to 132.5°) and latitudes (27.5° to 35.0°) during a warm seasons (June–September) from 1979 until 2017, as shown in Figure 3.2. Therefore, an input vector \( x \) for representing the meteorological field is composed of 48 dimensions (16 grid points, 3 dimensions), as expressed by Eq. 3.1. Thus, 19032 meteorological fields during 39 years are used for the SOM training.

\[
x = (PW_1, ..., PW_{16}, U_1, ..., U_{16}, V_1, ..., V_{16})
\]  

(3.1)

3.3 Annual variations in heavy rainfall frequency

In this study, 126 AMeDAS stations (Figure 3.3) located all over Kyushu Island are used for obtaining heavy rainfall frequency. Hourly rainfall amount for the study period of 39 years are obtained from the AMeDAS stations. Here, to guarantee the accuracy of data, only the stations which have been constantly operating and been located in the same place for 39 years are selected. A few of them were actually relocated just a little distance. In addition, stations installed or removed during the 39 years were excluded.

![Figure 3.3: Locations of AMeDAS stations used for this study in Kyushu Island](image-url)
Figure 3.4: Annual variation in heavy rainfall frequency
(a) $R \geq 30 \text{ mm/h}$, (b) $R \geq 50 \text{ mm/h}$, (c) $R \geq 70 \text{ mm/h}$
In this study, heavy rainfall amounts more than 30, 50 and 70 mm/h during the 39 years (1979-2017) in Kyushu Island are extracted from AMeDAS, respectively. Annual trends of the heavy rainfall frequency are shown by Figure 3.4, respectively. According to these figures, each total number of heavy rainfall frequency (≥ 30, 50 and 70 mm/h) observed all over Kyushu Island during 39 years were calculated as 15338, 2322 and 385 times, respectively.

As discussed in next section, each original time-series of heavy rainfall frequency is divided into many time-series of heavy rainfall frequency characterized by the synoptic weather patterns. For example, if a large number of meteorological fields are classified and grouped into 30 synoptic weather patterns, original time-series of heavy rainfall frequency is divided into 30 time series of heavy rainfall frequency. Consequently, increase or decrease trend of annual variation in heavy rainfall frequency depending on each synoptic weather pattern can be easily discussed.

In order to grasp the increase trend of heavy rainfall frequency (≥ 30, 50 and 70 mm/h) shown in Figure 3.4, the study period of 39 years was divided into period A (1979-1998) and period B (1999-2017). In each case, annual average heavy rainfall frequency of each period and the increasing rate of annual average heavy rainfall frequency between both periods were calculated, as given in Table 3.1. In the all cases, annual average heavy rainfall frequency in period B is higher than that in period A. Moreover, as rainfall becomes more intense, the increasing rate becomes higher. Especially, the increasing rate of heavy rainfall (≥ 70 mm/h) is much higher than the others. This means that the increase trend of extremely heavy rainfall frequency is remarkable for recent period. These annual variations in heavy rainfall frequency are divided into the synoptic weather groups classified by the SOM.

Table 3.1: Annual average heavy rainfall frequency during each period and the increasing rate (%) of annual average heavy rainfall frequency between past and recent periods in the cases of ≥ 30, 50 and 70mm/h

<table>
<thead>
<tr>
<th>Heavy rainfall category</th>
<th>Annual average frequency in period A</th>
<th>Annual average frequency in period B</th>
<th>Increase rate of heavy rainfall between period A and B</th>
</tr>
</thead>
<tbody>
<tr>
<td>R ≥ 30 mm/h</td>
<td>358.5</td>
<td>430</td>
<td>20</td>
</tr>
<tr>
<td>R ≥ 50 mm/h</td>
<td>51</td>
<td>68.6</td>
<td>34.6</td>
</tr>
<tr>
<td>R ≥ 70 mm/h</td>
<td>7.1</td>
<td>12.8</td>
<td>80.1</td>
</tr>
</tbody>
</table>
3.4 Heavy rainfall frequency corresponding to meteorological field

As mentioned before, a meteorological field (NCAR/NCEP reanalysis) used as an input vector of the SOM training are provided 4 times per day (3, 9, 15 and 21 JST, every 6 hours). On the other hand, hourly rainfall data (AMeDAS) are obtained every 1 hour. In order to relate meteorological fields to heavy rainfall frequency (≥ 30, 50 and 70 mm/h, respectively), 24 hours (1 day) are divided into 4 periods, 0-6 JST, 6-12 JST, 12-18 JST and 18-24 JST. Subsequently, the number of observed heavy rainfalls within each time period in the target area (Kyushu Island) corresponds to a meteorological field at 3, 9, 15 and 21 JST, respectively, as shown in Figure 3.5.

![Diagram](image)

Figure 3.5: Calculation of the number of heavy rainfalls corresponding to a meteorological field

3.5 Determination of the SOM structure

It should be noted that suitable way to determine the optimal SOM structure has not been completely established. This study determines the optimal SOM structure by using some conventional methods. The procedure is to determine the optimal number of units, optimal side lengths, and the optimal number of groups.

The first step is to determine the optimal number of units on the SOM using heuristic formula suggested by Vesanto et al. (2000), as presented in Eq. 3.2.
$M \approx 5\sqrt{n}$  \hspace{1cm} (3.2)

where $M$ is the heuristic optimal number of map units and $n$ is the number of input vectors for the SOM training. In this study, by using 19032 input vectors, the optimum number of map units is given as 689.

The second step is the determination of the width ($X$) and length ($Y$) based on the total unit number of 689. In this study, five cases of the width ($X$) and length ($Y$) are selected as the candidates for the optimal SOM structure, as listed in Table 3.2. The total number of units for these cases shows a value near the optimal number of 689. For instance, the map size of $X_{21}Y_{32}$ (672 units) means that the side lengths of width ($X$) and length ($Y$) on the map is 21 and 32, respectively. The result of SOM performance evaluated by Quantization error (QE) and Topographical error (TE) obtained after the SOM training shows that QE and TE differences among the five cases are negligible small. This implies that any case is suitable for a candidate for determining the SOM structure.

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

<table>
<thead>
<tr>
<th>Map size</th>
<th>QE</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{21}Y_{32}(672)$</td>
<td>0.0598</td>
<td>0.0492</td>
</tr>
<tr>
<td>$X_{21}Y_{33}(693)$</td>
<td>0.0595</td>
<td>0.0474</td>
</tr>
<tr>
<td>$X_{24}Y_{28}(672)$</td>
<td>0.0598</td>
<td>0.0499</td>
</tr>
<tr>
<td>$X_{26}Y_{26}(676)$</td>
<td>0.0598</td>
<td>0.0500</td>
</tr>
<tr>
<td>$X_{26}Y_{27}(702)$</td>
<td>0.0595</td>
<td>0.0530</td>
</tr>
</tbody>
</table>

Table 3.3: Specification of smallest groups given in each case of Table 3.2

<table>
<thead>
<tr>
<th>Map size</th>
<th>Initial smallest group</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{21}Y_{32}(672)$</td>
<td>x1y1</td>
</tr>
<tr>
<td></td>
<td>x3y2</td>
</tr>
<tr>
<td>$X_{21}Y_{33}(693)$</td>
<td>x1y1</td>
</tr>
<tr>
<td></td>
<td>x3y3</td>
</tr>
<tr>
<td>$X_{24}Y_{28}(672)$</td>
<td>x1y1</td>
</tr>
<tr>
<td></td>
<td>x2y2</td>
</tr>
<tr>
<td>$X_{26}Y_{26}(676)$</td>
<td>x1y1</td>
</tr>
<tr>
<td></td>
<td>x2y2</td>
</tr>
<tr>
<td>$X_{26}Y_{27}(702)$</td>
<td>x1y1</td>
</tr>
<tr>
<td></td>
<td>x2y3</td>
</tr>
</tbody>
</table>
Therefore, for the third step, clustering procedure consisting of Ward method and subsequent K-means method is applied for all the SOM cases shown in Table 3.2. The evaluation is based on the DBI. The minimum value of DBI gives the optimal clustering. The procedure is conducted in the next manner according to Figure 3.6.

Figure 3.6: Clustering procedure based on Ward method and subsequent K-means method

1) The smallest group given as an initial group is predetermined in each cases of Table 3.3. Here, for example, x3y2 means the smallest group consists of 6 units with 3 units (x-axis) and 2 units (y-axis) on the map.

2) First, Ward clustering is applied to 11 cases shown in Table 3.3. The clustering starts from the smallest group, and continues until all the units on the map are classified into predetermined final groups (10–70 groups).

3) Next, K-means clustering is applied for the rearrangement of units locating near boundaries among groups because the ‘boundary’ units may be more similar to features of another group.

4) Finally, the value of DBI is calculated per a predetermined final group (10–70 groups) after the application of the K-means.

Figure 3.7 shows the dependency of DBI on the number of groups on the map in all the cases of Table 3.3. According to Figure 3.7, the DBI decreases with an increase in the number of groups. The map of X21Y33 (x1y1) with 70 groups takes the lowest DBI. Following the rule that the number of groups with the lowest DBI is optimal, the map of X21Y33 with 70 groups would be selected as the optimal SOM structure. However, the application of many groups to the map interrupts the extraction of remarkable features included in an annual variation of heavy rainfall frequency depending on each synoptic weather group because the total number of heavy rainfalls ≥ 70 mm/h observed in Kyushu Island is only 385 times during 39 years. Therefore, for obtaining reliable annual variation in heavy rainfall frequency, a group number less than 70 should be selected.
Again, looking at Figure 3.7, the DBI decreases drastically from 1.3 to 1.05 with an increase in the number of groups between 10 and 40. On the other hand, the DBI after 45 groups shows a moderate decrease, and the variability of the DBI after 45 groups is small. In other words, the tendency of the DBI after 45 groups is stable within the extent of 0.95 and 1.05. Focusing on the feature, the number of groups and the map structure used for this study is selected by the comparison of the DBI among all cases with 45 or 50 groups. The result shows that the DBI with 45 groups for X24Y28 (x3y2) is the lowest values, as compared with the other cases less than 50 groups. Thus, the map of X24Y28 with 45 groups has been selected for this study.

![Figure 3.7](chart.png)

Figure 3.7: The dependency of DBI on the number of groups in all the cases of Table 3.3
Chapter 4

Result

4.1 SOM structure used for this study

According to the previous section, using the SOM algorithm, 19032 meteorological fields observed for 39 years were classified into 672 synoptic weather patterns on two-dimensional map (X=24, Y=28), which were divided into 45 groups by applying the two clustering techniques of Ward method and K-means method. The fields of the 45 synoptic weather groups are depicted in Appendix.

As can be seen in Figure 4.2, the SOM has visually-and easily comprehensible unique feature that adjacent groups on the SOM are similar to each other and distant groups are dissimilar. For example, the synoptic weather groups of adjacent G37 and G44 are similar. On the other hand, comparing G44 and G07, which are distant between both groups, G44 is characterized by the intrusion of humid air with strong wind from the southwest. G07 is characterized by the intrusion of dry air with weak wind from the north. Actually, both groups are quite opposites of each other.
As shown in Figure 3.4 and Table 3.1, for the all case of annual heavy rainfall frequency ≥ 30, 50, and 70 mm/h in Kyushu Island, the increase trend of heavy rainfall frequency during the recent period (1999-2017) is remarkable as compared with the past period (1979-1998). Especially, the increasing rate of heavy rainfall frequency ≥ 70 mm/h is much higher than the others. Therefore, this study focuses on features of significant relationships annual heavy rainfall frequency ≥ 70 mm/h and synoptic weather patterns or groups.

Figure 4.2: Transition of features of synoptic weather groups with distance from G44 to G07 depicted by the arrow in Figure 4.1

### 4.2 Heavy rainfall frequency in the map

As shown in Figure 3.4 and Table 3.1, for the all case of annual heavy rainfall frequency ≥ 30, 50, and 70 mm/h in Kyushu Island, the increase trend of heavy rainfall frequency during the recent period (1999-2017) is remarkable as compared with the past period (1979-1998). Especially, the increasing rate of heavy rainfall frequency ≥ 70 mm/h is much higher than the others. Therefore, this study focuses on features of significant relationships annual heavy rainfall frequency ≥ 70 mm/h and synoptic weather patterns or groups.
Heavy rainfalls measured during the study period of 39 years are classified into 672 units and plotted on each unit on the SOM, as can be seen in Figure 4.3. The number in each unit in the map means heavy rainfall frequency caused by a synoptic weather pattern characterizing each unit. The number of heavy rainfalls for each group is shown in Table 4.1. By looking at the map, it can be visually and easily identified which units and groups have frequently caused heavy rainfall so far. As a feature of the map, the groups located on the left side, especially the upper left side, of the map have caused more heavy rainfall. Therefore, the map is very useful as a first step for finding out synoptic weather groups relating the occurrence of heavy rainfall frequency.

![Figure 4.3: Heavy rainfall frequency (≥ 70mm/h) for 39 years in each unit on the SOM](image)

**Table 4.1: Heavy rainfall frequency ≥ 70mm/h over 39 years for each group**

<table>
<thead>
<tr>
<th>Group</th>
<th>Frequency</th>
<th>Group</th>
<th>Frequency</th>
<th>Group</th>
<th>Frequency</th>
<th>Group</th>
<th>Frequency</th>
<th>Group</th>
<th>Frequency</th>
<th>Group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>G16</td>
<td>45</td>
<td>G23</td>
<td>14</td>
<td>G21</td>
<td>3</td>
<td>G28</td>
<td>1</td>
<td>G15</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G44</td>
<td>42</td>
<td>G14</td>
<td>12</td>
<td>G24</td>
<td>3</td>
<td>G02</td>
<td>0</td>
<td>G22</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G29</td>
<td>34</td>
<td>G19</td>
<td>10</td>
<td>G33</td>
<td>3</td>
<td>G03</td>
<td>0</td>
<td>G35</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G18</td>
<td>31</td>
<td>G17</td>
<td>9</td>
<td>G34</td>
<td>3</td>
<td>G04</td>
<td>0</td>
<td>G39</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G32</td>
<td>26</td>
<td>G38</td>
<td>9</td>
<td>G45</td>
<td>3</td>
<td>G05</td>
<td>0</td>
<td>G42</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G37</td>
<td>23</td>
<td>G10</td>
<td>8</td>
<td>G08</td>
<td>2</td>
<td>G06</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G41</td>
<td>22</td>
<td>G26</td>
<td>8</td>
<td>G20</td>
<td>2</td>
<td>G07</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G30</td>
<td>18</td>
<td>G43</td>
<td>8</td>
<td>G27</td>
<td>2</td>
<td>G09</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G36</td>
<td>15</td>
<td>G31</td>
<td>7</td>
<td>G12</td>
<td>1</td>
<td>G11</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G40</td>
<td>15</td>
<td>G01</td>
<td>5</td>
<td>G25</td>
<td>1</td>
<td>G13</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In this study, one of the main aims is to identify what kinds of weather patterns have frequently caused heavy rainfall so far. Therefore, this study focuses on the top 10 groups which have caused heavy rainfalls during 39 years. From Table 4.1, the synoptic weather group of G16 has caused most heavy rainfall during 39 years in Kyushu Island. The top 10 groups can be recognized as G16, G44, G29, G18, G32, G37, G41, G30, G36, and G40. According to Table 4.1, the total heavy rainfall frequency of all the 45 groups during 39 years is 385 times. On the other hand, the frequency of the top 10 groups is 271 times. Namely, the frequency of the top 10 groups explains 70.4% of the total frequency of all the 45 groups. This implies that the top 10 groups have frequently induced heavy rainfalls in Kyushu Island so far.

Figure 4.4 illustrates the top 10 synoptic weather groups relating to heavy rainfall occurrence. G16 and G30 are characterized by the existence of stationary front formed along the boundary between humid air mass in the southern area and dry air mass in the northern area. G18 represents the approach of tropical cyclone (Typhoon) towards Kyushu Island. The approach of the typhoon transports a large amount of water vapor into mountainous areas ranging from the north to south around the middle of Kyushu Island. Therefore, strong tendency to cause heavy rainfall along the east side of Kyushu Island can be clearly
recognized, as also shown by Nishiyama (2007). G29 illustrates the existence of small-scale low-pressure system formed above a stationary front. G36, G37, G41 and G44 are known as typical synoptic weather fields causing heavy rainfall, and called as moist tongue, which means the inflow of a large amount of water vapor with strong wind (LLJ: low level jet) in lower layers. G40 shows that Kyushu Island is located in the edge of the Pacific high-pressure system. In this case, heavy rainfall events tend to occur under the influence of potentially-unstable atmospheric situation and the inflow of moist air.

On the other hand, there were no heavy rainfalls ≥ 70mm/h observed by G02-G07, G09, G11, G13, G15, G22, G35, G39, and G42 during 39 years. The synoptic groups excepting G35, G39, and G42 are surrounded by dry air, which suppresses heavy rainfall due to the enhancement of atmospheric stability. On the other hand, G39 and G42 are characterized the existence of moist air under the influence of the Pacific high-pressure system. In this case, relatively strong rainfall in the afternoon is expected.

4.3 Annual variation in heavy rainfall frequency for each group

Figure 4.5 shows the contribution of all the synoptic weather groups to annual variation in heavy rainfall frequency ≥ 70mm/h. Moreover, the figure is divided into the contribution of the top 10 group (Figure 4.6) and the other 35 groups combined (Figure 4.7). As can be seen in Figure 4.5 and 4.6, the total frequency of the top 10 groups has dramatically increased recently. On the other hand, recent increase in the frequency of the other 35 groups is not so remarkable, as shown by Figure 4.7. Thus, the features of the top 10 groups can be recognized as major factors of the recent increase of heavy rainfall frequency.

![Figure 4.5: Contribution of all the synoptic weather groups to annual variation in heavy rainfall frequency ≥ 70mm/h](image_url)
Figure 4.6: Contribution of the top 10 synoptic weather groups to annual variation in heavy rainfall frequency ≥ 70mm/h

Figure 4.7: Contribution of the other 35 synoptic weather groups to annual variation in heavy rainfall frequency ≥ 70mm/h

Figure 4.8 describes annual variation in heavy rainfall frequency ≥ 70mm/h of the top 10 groups. The study period of 39 years was divided into period A (1979-1998) and period B (1999-2017), as is the case with Figure 3.4. Then, by comparing these annual average frequencies of both periods for each group, this study recognizes what kinds of weather groups have highly contributed to the recent increase of heavy rainfall frequency. In Figure 4.8, the red numbers mean that annual average heavy rainfall frequency during
period B is higher than during period A. On the other hand, the blue numbers mean that annual average frequency during period B is lower than during period A. The result shows that annual average frequency of period B is higher than that of period A in the 8 groups out of the top 10 groups. Therefore, the result implies that 8 synoptic weather groups have significantly contributed to the recent increase of heavy rainfall frequency. Especially, the top 3 groups (G16, G44, and G29) show remarkable recent increase trend.

![Figure 4.8: Annual variations in heavy rainfall frequency ≥ 70mm/h of the top 10 groups. Target period in this study is divided into period A (1979-1998) and B (1999-2017). The two numbers in each graph show annual average heavy rainfall frequency during each period. The red and blue numbers show increase and decrease trend of the recent period B compared with the past period A, respectively.](image-url)
4.4 Details of heavy rainfall frequency in G16, G44 and G29

This subsection investigates the reasons of drastic increase in heavy rainfall frequency of the groups of G16, G44 and G29 during the recent period. Tables 4.2, 4.3, and 4.4 show heavy rainfall frequency per one heavy rainfall event ($\geq 70$mm/h) of G16, G44 and G29, respectively. As already described in chapter 3, heavy rainfall frequency in one event is defined by the number of observed heavy rainfalls within each time period (6 hours interval) corresponding to a meteorological field at 3, 9, 15 and 21 JST in the target area (Kyushu Island), respectively, as shown in Figure 3.5.

As can be seen in these tables, heavy rainfall frequency per one heavy rainfall event during the recent period is sometimes higher than during the past period. In other words, for any group, it can be recognized that the number of events recording heavy rainfall frequency $\geq 2$ times has increased during the recent period, as compared with in the past period. For example, G44 includes the event (18-24JST, June 20, 2006) with the frequency of 11 times in the recent period, and, on the other hand, there are no events recording heavy rainfall frequency $\geq 2$ times in the past period. Moreover, in any group, the increase in the number of events causing heavy rainfall ($\geq 70$mm/h) during the recent period can be clearly recognized. Therefore, these two reasons contribute to the recent increase in heavy rainfall frequency.
Table 4.2: Heavy rainfall frequency per one event (≥ 70mm/h) of G16

<table>
<thead>
<tr>
<th>Period</th>
<th>Date of heavy rainfall event (year/month/day/hour)</th>
<th>Unit</th>
<th>Group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1987/08/08/03</td>
<td>217</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1989/07/10/09</td>
<td>267</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1991/06/13/09</td>
<td>241</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1993/06/18/09</td>
<td>218</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1995/07/03/21</td>
<td>265</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1996/06/18/03</td>
<td>193</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1996/07/03/15</td>
<td>267</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>B</td>
<td>2003/07/19/09</td>
<td>194</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2003/08/26/03</td>
<td>243</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2005/07/10/03</td>
<td>241</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>2005/07/30/09</td>
<td>243</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2005/07/30/15</td>
<td>267</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2006/07/02/03</td>
<td>241</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>2006/07/05/21</td>
<td>265</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>B</td>
<td>2006/07/21/21</td>
<td>266</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2006/07/22/09</td>
<td>266</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2007/07/11/03</td>
<td>217</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>2007/07/11/09</td>
<td>194</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2008/06/19/09</td>
<td>218</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2010/06/20/15</td>
<td>241</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2010/07/14/03</td>
<td>217</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2010/07/14/09</td>
<td>218</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2012/07/12/03</td>
<td>241</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>B</td>
<td>2012/07/12/09</td>
<td>242</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>2016/06/21/03</td>
<td>241</td>
<td>16</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4.3: Heavy rainfall frequency per one event (≥ 70mm/h) of G44

<table>
<thead>
<tr>
<th>Period</th>
<th>Date of heavy rainfall event (year/month/day/hour)</th>
<th>Unit</th>
<th>Group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1983/07/05/03</td>
<td>650</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1985/06/28/03</td>
<td>602</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1985/07/09/09</td>
<td>604</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1990/06/15/09</td>
<td>626</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1991/06/30/15</td>
<td>651</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1991/06/30/21</td>
<td>651</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1993/08/01/09</td>
<td>628</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1993/08/01/15</td>
<td>628</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1995/07/02/21</td>
<td>601</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2001/06/29/03</td>
<td>652</td>
<td>44</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2006/07/05/03</td>
<td>649</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2007/07/06/09</td>
<td>651</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2008/06/21/15</td>
<td>627</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2008/06/21/21</td>
<td>627</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2009/07/24/15</td>
<td>601</td>
<td>44</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2010/07/22/21</td>
<td>602</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2010/07/03/03</td>
<td>649</td>
<td>44</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2010/07/03/09</td>
<td>649</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2011/07/06/09</td>
<td>651</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2012/07/11/09</td>
<td>649</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2012/07/11/21</td>
<td>649</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2012/07/13/03</td>
<td>604</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2014/07/06/15</td>
<td>603</td>
<td>44</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2016/06/20/21</td>
<td>652</td>
<td>44</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2016/06/22/21</td>
<td>651</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2017/08/15/03</td>
<td>629</td>
<td>44</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.4: Heavy rainfall frequency per one event (≥ 70mm/h) of G29

<table>
<thead>
<tr>
<th>Period</th>
<th>Date of heavy rainfall event (year/month/day/hour)</th>
<th>Unit</th>
<th>Group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1979/07/17/09</td>
<td>409</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1979/07/17/15</td>
<td>385</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1991/09/14/09</td>
<td>385</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1993/06/13/21</td>
<td>435</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1998/07/19/03</td>
<td>385</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1998/07/26/03</td>
<td>433</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1999/06/29/03</td>
<td>458</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1999/06/29/09</td>
<td>434</td>
<td>29</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>2004/09/07/09</td>
<td>433</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2004/09/29/09</td>
<td>385</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2006/09/17/15</td>
<td>433</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2009/07/24/21</td>
<td>458</td>
<td>29</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2009/07/25/03</td>
<td>458</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2014/07/03/09</td>
<td>410</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2014/08/15/09</td>
<td>435</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2015/08/12/09</td>
<td>458</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2015/08/25/03</td>
<td>433</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2015/08/25/09</td>
<td>433</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2016/09/07/15</td>
<td>386</td>
<td>29</td>
<td>1</td>
</tr>
</tbody>
</table>
4.5 Annual variation in appearance frequency of the top 3 synoptic weather groups (G16, G44 and G29)

This subsection investigates whether the increase trend of heavy rainfall frequency is related to that of the appearance frequencies of the synoptic weather groups causing heavy rainfall. Figure 4.9 shows annual variations in appearance frequencies of top 10 groups depicted as color columns, and the other 35 groups combined. The result shows that there is no increase trend of appearance frequency for the top 10 synoptic weather groups during 39 years regardless of the increase trend of heavy rainfall frequency shown by Figure 4.6.

Therefore, by focusing on the top 3 synoptic weather groups (G16, G44, and G29) causing most heavy rainfall, features of appearance frequency are confirmed. Figure 4.10 shows annual variation in heavy rainfall frequency and appearance frequency of (a) G16+G44+G29, (b) G16, (c) G44, and (d) G29. For G16+G44+G29, drastic increase (1.4/year to 4.89/year) in heavy rainfall frequency between two periods can be clearly recognized. For these three groups, heavy rainfall frequencies in the recent period are more than 3 times higher than in the past period. However, the increase trend of appearance frequencies of the groups cannot be recognized. In other words, there is no relationship that the increase trend of heavy rainfall frequency corresponds to that of appearance frequencies of these groups. The reason cannot be interpreted by just using NCAR/NCEP reanalysis data with 2.5 degree grid (longitude and latitude) because the occurrence of heavy rainfall is often controlled by meteorological factors such as regional topographical features. Although the investigation of these features is beyond the scope of this study, further analysis on smaller-scales (referred to as meso-scale) will be required.
Figure 4.10: The left figures show annual variation in heavy rainfall frequency of (a) G16+G44+G29, (b) G16, (c) G44, and (d) G29. The right figures show annual variations in appearance frequencies of these four cases. The red and blue numbers show increase and decrease trend of the recent period B compared with the past period A, respectively.
Chapter 5

Conclusion

In this study, 1) a large number of metrological fields observed for several decades were classified into synoptic weather patterns by the Self-Organizing Map (SOM) developed by Kohonen (1995). In the next step, 2) the synoptic weather patterns were divided into groups by subsequent clustering techniques consisting of Ward method and K-means method. Finally, 3) a time-series in heavy rainfall frequency obtained from AMeDAS in Kyushu Islands are divided into synoptic weather groups.

Consequently, this study identified what kinds of synoptic weather groups have frequently caused heavy rainfall so far and significantly contributed to the recent increase of heavy rainfall frequency. Our frequency analysis of heavy rainfall results in as follows;

1) Approximately 70% of the total heavy rainfall frequency \( \geq 70 \text{mm/h} \) observed in Kyushu Island for the last 39 years has been caused by the top 10 groups, which have also highly contributed to the recent increase of heavy rainfall frequency. Among them, especially, the synoptic weather groups of stationary front, a combination of stationary front and low pressure, and moist tongue, have frequently caused heavy rainfalls during the recent period. Therefore, when those weather groups causing heavy rainfall appear in Kyushu Island, it is considerably important to carefully check weather conditions for recognizing the possibility and risk of natural disasters associated to heavy rainfall.

2) The reasons of drastic increase in heavy rainfall frequency were investigated. Heavy rainfall frequency per one heavy rainfall event during the recent period is sometimes higher than during the past period. In other words, for any group, it can be recognized that the number of events recording heavy rainfall frequency \( \geq 2 \) times has increased in the recent period, as compared with in the past period. Moreover, in any group, the increase in the number of events causing heavy rainfall \( (\geq 70 \text{mm/h}) \) during the recent period can be clearly recognized. Therefore, it was found out that these two reasons contribute to the recent increase in heavy rainfall frequency.

3) Focusing the top 3 groups causing most heavy rainfall, heavy rainfall frequencies in the recent period are more than 3 times higher than that in the past period. However, the increase trend of appearance frequencies of the groups cannot be recognized. The reason cannot be interpreted by just using NCAR/NCEP reanalysis data with 2.5 degree grid (longitude and latitude) because the occurrence of heavy rainfall is often controlled by meteorological factors such as regional topographical features. Therefore, further investigations on smaller-scales (referred to as meso-scale) will be needed in the future work.
References


K. Nishiyama: Diagnosis of climate and weather, Climate Change Modeling, Mitigation, and Adaptation, ASCE books, American Society of Civil Engineers, chapter 17, 2013.


Appendix

Synoptic weather groups

The size of arrow shows the wind speed and direction. Moreover, the size of circle describes the amount of precipitable water (PW). The scale values are as follows;

20m/s → 10m/s

60mm ○ 40mm

Figure A.1: Synoptic weather group for G01 / Figure A.2: Synoptic weather group for G02

Figure A.3: Synoptic weather group for G03 / Figure A.4: Synoptic weather group for G04
Figure A.5: Synoptic weather group for G05
Figure A.6: Synoptic weather group for G06

Figure A.7: Synoptic weather group for G07
Figure A.8: Synoptic weather group for G08

Figure A.9: Synoptic weather group for G09
Figure A.10: Synoptic weather group for G10
Figure A.11: Synoptic weather group for G11 /Figure A.12: Synoptic weather group for G12

Figure A.13: Synoptic weather group for G13 /Figure A.14: Synoptic weather group for G14

Figure A.15: Synoptic weather group for G15 /Figure A.16: Synoptic weather group for G16
Figure A.17: Synoptic weather group for G17

Figure A.18: Synoptic weather group for G18

Figure A.19: Synoptic weather group for G19

Figure A.20: Synoptic weather group for G20

Figure A.21: Synoptic weather group for G21

Figure A.22: Synoptic weather group for G22
Figure A.23: Synoptic weather group for G23

Figure A.24: Synoptic weather group for G24

Figure A.25: Synoptic weather group for G25

Figure A.26: Synoptic weather group for G26

Figure A.27: Synoptic weather group for G27

Figure A.28: Synoptic weather group for G28
Figure A.29: Synoptic weather group for G29

Figure A.30: Synoptic weather group for G30

Figure A.31: Synoptic weather group for G31

Figure A.32: Synoptic weather group for G32

Figure A.33: Synoptic weather group for G33

Figure A.34: Synoptic weather group for G34
Figure A.35: Synoptic weather group for G35 /Figure A.36: Synoptic weather group for G36

Figure A.37: Synoptic weather group for G37 /Figure A.38: Synoptic weather group for G38

Figure A.39: Synoptic weather group for G39 /Figure A.40: Synoptic weather group for G40
Figure A.41: Synoptic weather group for G41
Figure A.42: Synoptic weather group for G42
Figure A.43: Synoptic weather group for G43
Figure A.44: Synoptic weather group for G44
Figure A.45: Synoptic weather group for G45