Dissertation

Development of Combined Downscaling Method for High-resolution Rainfall Estimation and River-Runoff and Inundation Simulation under Global Warming Condition

Graduate School of

Natural Science and Technology

Kanazawa University

Division of
Environmental Design

Student ID No: 1524052013

Name: Tran Anh Quan

Chief Advisor: Associate Professor. Kenji TANIGUCHI

Date of submission: 2018/06

Development of Combined Downscaling Method for High-resolution Rainfall Estimation and River-Runoff and Inundation Simulation under Global Warming Condition

By

Tran Anh Quan

A dissertation submitted in partial fulfillment of the requirements for the Degree of Doctor of Engineering

Examination Committee	Assoc. Prof. Kenji Taniguchi (Chairman) Prof. Masatoshi Yuhi Prof. Masakatsu Miyajima Prof. Yasuo Chikata Assoc. Prof. Shinya Umeda
Nationality	Vietnamese
Previous Degree	Bachelor of science Vietnam National University Master of Engineering Kanazawa University
Scholarship Donor	Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT)

Kanazawa University
Graduate School of Natural Science and Technology
Kanazawa, Japan
June 2018

Abstract

In this study, the present and future rain-fall-runoff and inundation conditions of Cau-Thuong-Luc Nam river basin during rainy season (Jun-July-August) were also examined using the combination of the Weather Research and Forecasting (WRF) and River Run-off and Inundation model (RRI). We investigated the rainfall-runoff and inundation characteristics of the watershed in connection with the correspondence climate condition of the present (2000-2009) and future (2060-2069). The RRI model was used for the simulation of watershed hydrological characteristics. The essential future precipitation inputs for RRI were achieved by using the WRF model nested inside GFDL-CM3, and MIROC-5 models. Both WRF and RRI models are capable of deploying further assessment on the future river basin condition. The future downscaling results by GFDM-CM3 and MIROC-5, indicated heavier rainfall conditions in the mid-21st century and consequently cause severe inundation with higher depth, wider radius and longer period. Heavy rainfall and inundation were expected to increase in the second half of rainy season. In both GFDL-CM3 and MIROC-5, the impacted areas due to flood will increase in both spatial and temporal extent, intensity, and density. Future inundation condition will affect mostly the agricultural and residential areas in the lower Cau-Thuong-Luc Nam river basin.

The hybrid dynamical-statistical downscaling approach is an effort to combine the ability of dynamical downscaling to resolve fine-scale climate changes with the low computational cost of statistical downscaling. In this study, we propose a dynamical-statistical downscaling technique by incorporating WRF with artificial neural networks (ANN) to downscale rainfall information over the Red River Delta in Vietnam. First, WRF downscaling was performed to produce nested 30-km and 6-km resolution simulations. Then, in the statistical downscaling step, the ANN was trained to predict rainfall in the 6-km domain based on weather predictors in the 30-km simulation. The evaluation shows that the WRF modeling system can reproduce temporal variation in the JJA daily rainfall reasonably well, but underestimates the total precipitation. Owing to the higher precision of WRF, the region appears to have more drizzle, resulting in significantly fewer dry days than were observed. The best performing ANN model produced high-resolution rainfall patterns that are highly correlated with WRF ($r = 0.91$) and have low RMSE (12 mm/day). High-resolution rainfall in each grid cell was downscaled by taking the climatological variables from the four grid cells in the coarse domain. ANN was configured as an MLP-BG network with three hidden layers using the hyperbolic tangent sigmoid activation function. Running 30-km WRF and using ANN to downscale to 6 km is 89% less expensive than running nested 30-km and 6-km WRF simulations.

Acknowledgment

Firstly, I would like to express my sincere gratitude to my supervisor Assoc. Prof. Kenji TANIGUCHI for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my Ph.D. study.

Besides my supervisor, I would like to thank the other member of my thesis examination committee: Prof. Masatoshi YUHI, Prof. Masakatsu MIYAJIMA, Prof. Yasuo CHIKATA, and Assoc. Prof. Shinya UMEDA for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

Special thanks to all the fantastic people that I have met especially in the Lab, for all their support and friendship, especially my sincere thanks also goes to Dr. Tran Quoc Lap, Mr. Koki Nakaya, and Dr. Nguyen Trinh Chung for their great support in my research and living in Japan.

I am indebted to the Japan Ministry of Education, Culture, Sports, Science and Technology, who funded the study with half a year of research student and 3-years Ph.D. scholarship in Kanazawa University, and I want to thank Hanoi University of Mining and Geology - Department of Geo-Ecology and Environmental Technology, who allowed me to persuit my study.

Last but not the least, I must express my very profound gratitude to my parents and to my wife and son for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this dissertation. This accomplishment would not have been possible without them. Thank you.

ABBREVIATIONS

ADPC	Asian and Pacific Development Centre
CTLN	Cau-Thuong-Luc Nam river
IOA	Index of agreement
IPCC	Intergovernmental Panel on Climate Change
JJA	Jun, July, August
JMA	Japan Meteorological Agency
JRA-55	Japanese 55-year reanalysis
MAE	Mean absolute error
MLP	Multi-layer perceptron
MLP-BP	Multi-layer perceptron trained using back-propagation learning algorithm
MM5	Fifth-generation mesoscale model
MONRE	Ministry of Natural Resources and Environment
NCEP	National Center for Environmental Prediction
NCEP-FNL	NCEP Final Operational Global Analysis data
NCHMF	Vietnam National Centre for Hydro-meteorological Forecasting
NOAA	National Oceanic and Atmospheric Administration
NOAA OI SST NOAA	Optimum Interpolated 1/4 Degree Daily Sea Surface Temperature Analysis
PSU/NCAR	Pennsylvania State University/National Center for Atmospheric Research
RCM	Regional climate model
RMSE	Root mean square error
RRI	River runoff and inundation model
RRTM	Rapid radiative transfer model
SLP	Single layer perceptron
WRF	Weather research and forecasting

Table of Contents

List of Tables

List of Figures

CHAPTER 1.
INTRODUCTION

1.1. Background and motivation of the research

1.1.1. Climate change

(a) Global warming

Climate change has become a complicated and most destructive environmental issues around the world in the recent decades. The development of climate change exhibited through the variations in increasing trend of temperature which is mainly acknowledged as "Global warming". This warming was resulted from the over consumption of fossil fuel since the Industrial Revolution which lead to the increasing amount of carbon dioxide in the atmosphere. Carbon dioxide and other gases are the main culprit that enhanced greenhouse effect that potentially increase the temperature of the Earth. (Houghton, 2004, Zveryaev and Aleksandrova, 2004).

According to the IPCC, "climate change refers to a statistically significant variation in either the mean state of the climate or in its variability, persisting for an extended period (typically decades or longer). Climate change may be due to natural internal processes or external forcing or to persistent anthropogenic changes in the composition of the atmosphere or in land use" (IPCC, 2001). Therefore, the definition by IPCC pointed out both the concept of anthropogenic and natural component aspects.

Climate change has become one of the most pressing issues in the world today. While civilization has always had to live with, deal with and adapt to environmental and climatic challenges and risks, the challenges posed by climatic change are however believed to be exceed historical experiences. Changes in climate system presenting an unprecedented challenge to the global community at large and local scales. New adaptation measures to this problem are required (Christoplos et al., 2009). With the fourth IPCC report reviewed the linkage between anthropogenic activities and climate change, much of the climate change debate and research has focused on the issue of climate change mitigation (IPCC, 2007). Yet as the report discloses, due to the scope of climate change, mitigation efforts aiming at reducing emissions will not suffice. Adaptation will be necessary as the impacts of warming linked to past emissions cannot be avoided. Human systems are closely tied to climate systems, the globe

local communities will have to adapt to new climatic conditions, which in many regions will entail warmer temperatures, increased climate variability and an intensification of extreme weather

How to use the knowledge from the previous studies and experiences to understand the natural characteristics of the ongoing changes of the world is one of the growing concern recently. The recent efforts of the recent researches are reflected through the synthesis reports of IPCC. When cooperating the issues at the high level of techniques, both detection and attribution have different objectives. Detection of the change in global climate condition is to prove that climate has changed in some statistical range without the need of pointing out any specific reason for that (IPCC, 2007). In the past and current time, there have been many evidences of regional climatic changes that cementing this pressing problem. For instance, a lot of evidences has been found that showed the major parts of the cryosphere components are being generalized reduced in response to climate change. Meanwhile, there are several cases of growth that mainly related to increased snowfall. The reduction of glaciers that we have seen during the recent decades is the largest recession of all time throughout over 5,000 years of history, which is beside of the cover of any normal observation climate change. This problem probably resulted by anthropogenic warming for an certain extent (Jansen et al., 2007). As we can see in the Arctic and the Antarctic, the several-thousand-years-old ice blocks have gradually collapsing as the consequent of warming. For many cases, there has been a significant increasing trend in the shrinkage of the cryosphere during the past century which is proven to be consistent with the enhanced observed warming.

(b) Climate change and rainfall

The most significant impact of climate change to the world is shown in the variation of rainfall and precipitation in all forms. Compare to temperature, rainfall is much more difficult to be forecasted. However, the efforts of scientist in determining this problems have helped us to point out some of the important statements with confident about our future.

Precipitation is a scientific terms related to rainfall, snowfall or other kinds of water, from frozen to liquid matters or even vapor of cloud. Therefore, precipitation happens periodically and it occurrence is heavily depended on the temperature and weather scheme. Weather situation effecting precipitation chance related to the formation of storms, the change of moisture by winds with evaporation effects and how it was accumulated to form clouds.

Precipitation in the form of water vapour is usually condensed in the uprising air that diffused and reduce temperature of the atmosphere (IPCC, 2007).

Warmer atmospheric condition potentially keep more moisture, it is calculated that for every one warmer degree centigrade the water vapour in the earth increase for over 7% (Kevin E. Trenberth 2011). The ways vapour moisture transformed into precipitation is not yet to be fully understand but it is likely to rise 1-2% of rainfall water per a degree of warming (IPCC, 2007). This is a proof to illustrate that if a region is wet already, under global warming it can get even wetter. However, the details of the extent of wetter and its consequences in a regional scale are extremely difficult to be certain. By contrast, the dry regions are likely to become even drier. This symptom is found gradually shift toward the two poles.

Predicting the trend of changing in rainfall is particularly difficult due to the different characteristics of global weather patterns. Since most of climate models agree well in the future warming of global scale, they do not agree well when predicting how the change in temperature effect the variation in distribution, density and intensity of rainfall patterns, especially in detailed level of assumption. However, it is very possible that the warmer climate will enhanced heavy rainfall that produced from less extreme weather events. It could resulted in the longer dry periods and extended floods duration.

As far as we have pointed out, any impact of global warming to local scale rainfall cannot be differentiate from the natural processes. However, in several particular evidences, the signal has started to emerge. In a recent study (Kevin E. Trenberth 2011), anthropogenic climate change can greatly increase the odds of damaging floods occurring in England and Wale during autumn. For England, current findings have pointed out that the intensity of heavy rainfall during winter time might increase and become distinguishable from natural variations from 2020s. The climate models as well as our observation data is always being improved along with the advancement of technologies. The more reliable of climate forecasting results is also likely to be significantly improved in the coming future. For instance, new satellites and more high-resolution models are being established with updated possibilities for understanding and interpreting the flows of water cycles though our climate system.

(c) Climate change and water resources

There is a strong connection between climate system and hydrologic cycle which are both energized by the solar radiation. The climate system is a complicated systems including

the atmosphere, topography surface, oceans and other forms of water bodies, snow and ice patterns of the earth, and also the living species (IPCC, 2007). Climate systems do not stay still but evolves relentlessly over time through the interaction between its inner components. Inside the climate systems, all the components are linked together with internal dynamics and the exchange of external information and other factor effecting the climate or "forcings". The external forcings consist of natural phenomena such as storms, earth quakes, and volcanic eruptions was well as man-made changes that potential effect the atmospheric components and the change in land use. Any fluctuations in those factors can affect the balance between the incoming and outgoing radiation of the earth. It leads to the responses of the climatic systems to such changes, directly or indirectly.

The hydrological cycle is influenced by any impact to the climate system. Hydrologic cycles is considered as the uninterrupted flows of water between oceans, atmosphere, and the land surface. Solar radiation powered the hydrologic cycle started from the surface evaporation from water bodies. Since moisture is floated into the air, it cools down and gradually condenses in the form of clouds. Clouds and moisture is moving around the planet and comeback to the land surface by precipitation. As long as the water touch the soil, it can evaporate back to the air or penetrated through the surface soil layers and reach the ground water layers in the aquifers. Groundwater poured into sea, river, and streams. The part of water remain on the surface of the earth is water runoff which then seeps into lakes, rivers, and stream before finally flow back to the ocean where the hydrologic cycle begins again.

Many clear evidences have been found that climate change have already affected hydrologic cycle. Even when the long term tendency of hydrological indicators are difficult to establish due to the lack of observation variability in both spatial and temporal scale, the insitu observation changes of hydrologic cycle at continent scale are very consistent and associated with the variability of the climate system in the recent decades. (Bates, et al., 2008;.Kundzewicz, et al., 2007):

Global warming could significantly influences the future distribution of water availability around the world as well as the use of water. Water security and shortage is becoming the crucial matters in many pattern of the globe, at both global scale and regional scale, especially in agriculture sectors. Of course, the evidence of increase water variability have also been found in some regions and that changes can potential have exponential losses, for example the magnitude of extreme rainfall or flood. Hence, it can be highlighted that global warming is one of the factors effecting the future variations of water availability and use.

Water stress may become worsen as the results from climate change affecting hydrological condition. Many studies pointed out that water resources related problem will increase with the severe condition of climate change, and other than that, the impacts of other factors—demographic, socio-economic, and technological changes—are even more significant, especially in shorter time space (influences after the 2050s are pretty much dependent on the future population/discharge scenarios used). Depends on how the climate model was designed as well as other factors are embed in the model assessment, the differences between studies results might become larger. When finding the connection between climate change and population growth, Arnell (2004) has estimated that about 0.4 to 1.7 billion of people will be experienced with water stress in the 2020s and it will rise to 1 to 2.7 billion up to 2050s.

1.1.2. Development of climate change in Southeast Asia and in Vietnam

(a) Climate change in Southeast Asia

Recorded climatological data in the Asia and Pacific region has illustrate the increasing trend in the intensity and density of many extreme weather events including tropical cyclones, longer period of dry days, prolonged rainfall spell, snow avalanches, thunderstorms, dust storms, and heat waves. Besides, this regions is also expose to natural disaster which can be seen through the Ocean Tsunami in India in 2004, the 2005 Earthquake in Pakistan, and the landslide calamity in Philippines at late 2006. Those impacts added more risks for the already highly exposed communities who are striving to fight against poverty and find suitable adaptation strategy.

About 91% of the death and 49% of damage worldwide due to natural hazards happens in the Asia/Pacific region during the 20^{th} century. Thus, we can recognize the serious threat caused by climate change especially to poor smallholders and rural people living in the remote mountainous regions or marginal areas. The areas limited access in natural resources, communication and transport facilities often face more trouble in combating climate change.

Particularly, the region of Asia/Pacific has been forecasted with the increased temperature of 0.5 to 2°C at the end of 2030 and up to 1 to 7°C by late 2070. Higher temperature is predicted to be more intense in the dry areas in the North of Pakistan, India, and the western China. Furthermore, rainfall is also expected to accumulate higher regional wide, especially during the summer monsoon period. Besides, total precipitation during winter time seems to decrease in the South and Southeast Asia which indicate the higher chance in increased dry

condition. As the consequence of climate change, this region is believed to be effected by global sea level rise of about 3-16 cm in 2030 and goes up to 7-50 cm in late 2070 in accordance with regional sea level variation. Intense tropical cyclones are expected to experience substantial changes under the effect of large scale climatic driven like the El Niño-Southern Oscillation.

Many factors have pointed out that the Asia/Pacific areas are now highly vulnerable to the impacts of climate changes which may worsen the living condition of millions farmers who are already living in poverty. The major proportion of about 500 million rural poor in Asia are subsistence farmers who mainly depend on precipitation for farming. Influences of various hazards can much vary from the lack of daily food to likely to be harmed by disease, to the reduction in income and worsening livelihoods. Climate change is surely the fact of the modern development problems in the region.

(b) Northern region of Vietnam

Vietnam is divided into three main economic region which are the Northern, Central, and Southern. The Northern Vietnam is the oldest region among the three. It has more than 2000 years of historical culture-social development around the Red River Delta. Vietnamese then started to migrate southward to the Southern region which centralized in the Mekong Delta. Hanoi is the capital of Vietnam and locate in the Northern area, next to Hanoi in the east is Hai Phong city – the second biggest city in the North.

The total area of Northern Vietnam is approximately 116.332 km^2 consists of several sub-regions:

- The Northwestern region of 6 provinces locate on the right bank of the Red River including Lao Cai, Yen Bai, Dien Bien, Hoa Binh, Lai Chau, Son La. Among these provincese, Lao Cai and Yen Bai are also considerred as the sub-region Northeast.
- Northeastern region of 9 provinces including Ha Giang, Cao Bang, Bac Kan, Lang Son, Tuyen Quang, Thai Nguyen, Phu Tho, Bac Giang, Quang Ninh.
- The Red River Delta region of 10 provinces: Bac Ninh, Ha Nam, Ha Noi, Hai Duong, Hai Phong, Hung Yen, Nam Dinh, Ninh Binh, Thai Binh, Vinh Phuc.
- Nothern Region Vietnam located at 23◦23’ North – 8◦27’ West with the total length 1,650 km.

Northern Region is the biggest region in Vietnam:

- North is next to China.
- West is share the boder with Laos.
- East is next to Western Pacific Ocean..

Topography feature

Topography of Northern Vietnam is diverse and complicated. It includes mountain, plain and coastal plain. This place experienced through long time of terrain development and strong weathering condition. The land elevation is lower toward the Northwest-Southeast, that ware illustrated by the direction of main rivers (ADPC, 2003).

The vast land formed by the Red River delta has the total area 14.800 km2, accounts for 4.5% total area of the country. The delta has the triangular shape, with the highest point in Viet Tri city and the bottom side is the eastern coast. The Red River Delta (RRD) is the second biggest delta in Vietnam (The biggest is Mekong River Delta). RRD was formed by Red river

Figure 1.1 Map of Vietnam

and Thai Binh river. The topography of the Red River Delta is flat with height from 0.4 to 12m above sea level, in which about 56% of the area is lower than 2m.

Next to the RRD into the West and Northwest in the Northern mountainous region with the total area about 102,900 km2, accounted for 30.7% of the total area. The terrain includes high and dangerous mountains, from the northern borderline to the western Thanh Hoa province.

In the Northeast is the low mountainous region, next to the Western Pacific Ocean. This region is covered with small islands. In then Gulf of Tokin, there are about 3,000 small islands scattering around the area.

Regional advantages

Location advantages: having the Hanoi city which is the capital of economic, culture and politic, the Northern Vietnam has a great potentiality for developing socio-economic than other cities. In 2017, the total population of the region was about 34.3 million with density of 278 people/km^2. This region consists of the seven most important centrally municipal cities and provinces including Hanoi, Hai Phong, Hung Yen, Ha Tay, Hai Duong, Vinh Phuc, and Bac Ninh. The formation of these dynamic economic zones in the Northern region of Vietnam plays the very important role for development of the country (ADPC, 2003).

- Transportation strength: The door to the South of the contry found in this region. Therefore, this region has various transportation methods to meet the needs of the development including roadway, waterway, seaway, airway, railway (Hai Phong seaport, Noi Bai International airport). Hanoi is placed at the centre of the economic zone and links with others other national economic zones in the region as well as international points around the world. Northern region borders with China which is the largest market in the world and North-East Asian countries.

- Natural resource strength: The region is diversity in ecological environment from flat plain to midland and mountainous areas. Red river delta is most fertile agriculture area in the North that suitable for the adoptation of any comprehensive development plans of agriculture-forestry-fishery. This region is the second most productive agricultural production through the country, second to the Mekong delta in the South. This advantage ensures food security and economic for the vast region. Besides the potential in agriculture development, Northern Vietnam has various type of mineral resources with large capacity of coal (about 98% of natural reserves), kaolin (accounts for 40% of the whole country), and limestone (about 25% of total reserves)

- Advantages of labour resource: The region is abundant in labour sources especially the high education labours with about 26% of university and colleges of the whole country.

Meteorological Condition

The climate system in Vietnam is effected by the South East Asia tropical monsoon system; the average annual rainfall of the whole country is about 1,940 mm. Water is a precious resource and plays a very important role in the development of socio-economic. However, the current condition of surface water in Vietnam is now in danger, threatening by the negative

increasing of wide scale depletion and contamination. About 75% of the total lands is covered by mountains. Thus, precipitation is unevenly distributed between regions and fluctuate over times. The temporal distribution of mean annual rainfall vary in a wide range, in some specific mountainous areas the annual rainfall can get over 4,000 – 5,000 mm, an even up to 8,000 mm in such as Bach Ma Mountain. However, in some areas such as Nha Ho and Binh Thuan the annual rainfall is usually very low at 600 to 800mm (MONRE, 2006).

The climate system in Vietnam shifts upward from humid tropical climate in the Southern provinces to temperature in the Northern region. The Red river deltal is popularized by a tropical monsoonal climate affected by ocean-like climate. The main characteristic of this climate is seasonal and moist subtropical (Pfeiffer, 1984). The two distinguish seasons are wet and dry. The wet (rainy) season starts from Jun to September and dry season from December to March. The average temperature of the whole country is 23◦C.

The annual average precipitation varies from 1,300mm to 1,800mm. The rainy season accounts for over 85% of the annual rainfall (Li et al., 2006). During the year, July has the highest amount of rainfall while December and January are the driest. Even when the temperature between the sub-basin regions was not much different, the regions closer to the coastal zone always had higher rainfall intensity. In some special cases such as the Day Estuary and the lower Red River basins have annual rainfall up to 1860 and 1757 mm, respectively. Meanwhile, the average annual rainfall among other three sub-basins was only 1,600 mm (Luu et al., 2010). The historical observation climate record for 11 years found the peak in August 2006 (450 mm/month). Generally the total annual rainfall in the basin has reduce in recent years, but it happened with much higher intensity. It has resulted in the imbalance in rainfall distribution throughout the year (Luu et al., 2010).

1.1.3. Inundation and flood in Cau Thuong Luc Nam river basin

Vietnam has been known as a rapidly developing country but highly exposed to natural hazards. Inundation and flooding are the two major natural risks that the country have to face. According to the-the Intergovernmental Panel on Climate Change (IPCC), the global surface temperature is expected to rise 1-2°C by the year 2050 and up to about 2–5 °C at the end of 21st century (IPCC 2013). Global warming leads to the increasing trend of harmful disaster worldwide. Furthermore, climate change is forecasted to increase sea level rise accompany with the upward trend in frequency and intensity of floods, globally and in Southeast Asia (IPCC 2014; World Bank 2014). Observing the country's population distribution, population

density, and the economic assets in vulnerable regions, Vietnam has been considered as one of the five most effected countries by climate change worldwide.

Recent findings from IPCC's Fifth Assessment Report (AR5) indicated that climate change has already caused damages in many regions, especially the Southeast Asia (SEA) regions. Vietnam is believed one of the highly climate vulnerable countries in SEA region. Flood is the most destructive disaster in the Northern mountainous region in Vietnam. Since the occurrence of flood events has been increasing in recent decades, it has raised the needs and awareness of flood risk and management. Variations in density and intensity of rainfall directly cause impacts on the generation of floods. Flooding and inundation occur when the watershed system cannot handle the excessive amount of precipitation for either a short or prolonged rainfall events, it results in the unusual high stream flow that exceeds the capacity of a river channel (Vu and Ranzi 2017).

The Cau-Thuong-Luc Nam (CTLN) locates in the northern mountainous region in Vietnam, was formed by three small river basin including Cau river in the west, Thuong river in the middle, and the Luc Nam river in the east. The whole CTLN watershed is surrounded by mountain ranges the north, the terrain gradually lower toward southeast direction with the lower river basin places on the flat plateau. That is the reason why the majority of rivers, especially tertiary rivers have the great gradient.

CTLN river system is the upstream of Thai Binh river. The whole basin is influenced by the tropical region of the northern hemisphere. Under the impact of monsoon winds, this region has two distinctive seasons including the dry season starts from November to next April and the rainy season lasts from May to October. Strong hot south-west wind comes with complicated climatic turbulences like storms or air convergence are the main causes for the hot and humid weather with large amount of annual rainfall.

Annual water volume of the basin contributed by Cau, Thuong, and Luc Nam rivers is 4.5 billion m^3, 4.2 billion m^3 and 2.4 billion m^3 which made the total volume of the system of approximately 11 billion m^3. High rainfall has resulted in annual normal flow of average 22 l/s per km^2. However, this amount of river discharge is still far less than northern east region of the north coastal provinces and some regions in Red River Basin. The flow of the whole CTLN region varies upon rainfall regime which distinguished into 2 seasons, the flood season from Jun to August and low flow season from September to May.

CTLN watershed plays an important role in the development of the Northeastern Vietnam region. It provides domestic and production water source for the vast populated-economic regions including Thai Nguyen, Bac Ninh, Bac Giang, Lang Son, Quang Ninh, and the lower Thai Binh river basin.

In CTLN watershed, the lower river channel capacity is sometimes under accommodating the high discharge of surface flow over the upstream which cause flood. The number of flood occurrence in CTLN watershed is recorded rapidly increasing recently. In the theme of global warming, it is important to prepare the countermeasures to and mitigation methods to cope with the potential risks. It is impossible to establish the action plan without the deep understanding of the characteristics of the flood in accordant with the present and future hydro-meteorological conditions. The traditional simulation method for inundation researches combines both rain-fall-runoff models for river discharge and hydraulic model for water propagation. However, this method not suitable for flat watershed with large inundation area as it requires significant calculation between the river and flood water. This study, therefore, used the Rainfall-Runoff Inundation (RRI) model, which is a fully coupled model of rainfall-runoff model and hydraulic inundation model (Sayama et al. 2012). Besides, the Weather Research and Forecast (WRF) model (Skamarock et al. 2008) was also employed to provide present and future input precipitation for RRI model.

1.1.4. Downscaling methods for climate change

(a) The needs of weather information downscaling

To meet the requirements of management agencies to cope with global warming, a lot of materials such as synthesis report, statistic recrods offer forecasted climate influences at long-term and large scales which are finer than the original projections are prepared. It is necessary to know the group of assumptions embed in the techniques which were adopted to derive this climate details and the drawbacks they made on the results (Trzaska and Schnarr 2014). The most popular tools for projecting climate variation are General Circulation Model (GCMs), which are climatological models that contain numerous physical theories of the earth climate system. These physical theories are generally very famous but difficult to be always fully inserted in the models because of the limitations on computing capability and input data. Therefore, GCM simulation outputs can only be utilized at very coarse scale, which are global or continental scale to research climatological condition at averaged monthly, seasonal, annual,

and even longer time periods. Any data of finer spatial detail less than 100 square kilometres and temporal frequency of lower than monthly data has to be processed by a mathematical process to increase its resolution which is called downscaling.

Downscaling method outputs climatic details at higher resolution, it provides much more detail than the original GCM projections. For implementation of this method, downscaling process requires additional information, calculation, and contain more assumptions, and thus exhibiting more uncertainties and drawbacks of the final projection. A problem that these limitations are usually not made available to end-users (Hunt and Watkiss 2011). Currently, scientific organizations or management offices cannot provide instruction which helps researchers and decision makers for selecting projection model, climatological variable, the approach to downscaling dat, and the sources of data that are suitable for their purposes (Trzaska and Schnarr 2014). Since downscaling methods are still under development of researching organizations, users often have to rely on complex technical report and specialized guidance to understand the model and appropriately apply their results for impacting studies, planning related purposes, or decision-making. The following are important considerations and recommendations to be aware and remember when simulating and describing fine-scale climatological detail on climate variability and its consequences.

- Downscaling techniques considers that the local climate was formed by the combination of large scale climate indicators (such as global, hemispheric, continent, and region) with local variables (including the topographic condition, land cover, surface layers of the earth, water availability). Local variables can only be acquired at the high detail level of assimilation model which cannot be provided by the current GCMs.
- The process of downscaling global climate projections to local scale is a complicated multistep process as can be seen in Figure 1.2. During the process, the additional assumptions and similarity are generated. Confusions and suspicions are made from projections of variation in climate system with their impacts. Those uncertainties originated from various sources and need to be knowledgeable, whether accurately calculated or not.

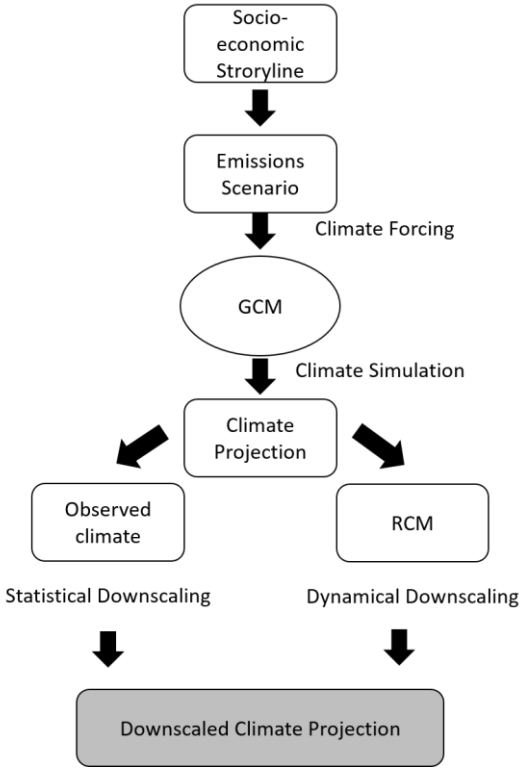

Figure 1.2 Illustration of the components involved in developing global and regional climate projection

Deriving climate information to local detail can be performed spatially and temporally. Popularly, different downscaling approach are fused to achieve the projection data at preferable spatial and temporal levels.

- The two main approaches to combine the climate change information at local scale with large-scale climate indicator are discussed as follows (Trzaska and Schnarr 2014):
    - Dynamical: by incorporating supplement data and physical theories in regional climate models which are similar to GCMs but offer a more detailed resolution but cover only a small part of the global. Such approach has many benefits but very computing expensive and needs large amount of data with an extreme level of knowledge to apply and discussing the results. Dynamical downscaling requirements are too high that often higher than the capability of researching facility in developing countries.
    - Statistical: this approach creates an empirical relationship between large-scale historical climate characteristics between outputs of GCMs and local climate features. Contradict to the dynamical downscaling approach, the statistical downscaling is computational inexpensive and much easier to demonstrate.

They need minimal computing resources but most importantly that they rely explicitly on historical climate in-situ data and the hypothesis that the currently established statistical relationships will continue to hold true in the future. Nonetheless, long term and good quality historical climatological observation data are not always accessible in many developing countries.

For most purposes, different downscaling methods aim to obtain climate detail of the fine scale resolution but the findings from previous researches have shown climate change and its consequences has exhibiting the problems mentioned below:

- Detail information on downscaling results with their uncertainties are usually inappropriately expressed which lead the user to think that the projections are truthful and believable at the resolution achieved. Carefully examining of technical notes is always recommended to deeply understand all the simulation steps and hypothesis that included in the final results.
- Uncertainties included in downscaling output and additionally inherited from the used downscaling approach are usually not shown in detail. This important information is lack of quantification and discussion, that lead the numerical results are considered at face value.
- Verification of downscaling results using observation data are often omitted; it would be better to compare downscaled results to high-resolution observed data to made clear systematic biases as well as the limitations in assumptions.

The above key findings mostly resulted from simple oversight by the authors of this dissertation. However, they are sometime very important and heavily influence the final assessment of the problem. An end user would better be recognizing them and understand the limits of the results.

(b) Dynamical downscaling

Dynamical downscaling is performed by using a regional climate model (RCM) model driven by large-scale GCM outputs to achieve local climate information. RCM models are similar to GCMs but provide finer details which having more regional characteristics. This advantage helps them better capture local topography with the micro-physics processes of local atmospheric (Trzaska and Schnarr 2014). The GCM model contains the feedbacks of the global circulations which will develop in the atmospheric balance following various meteorological

processes. However, a part of this processes need to be approximated because of the very low resolution seen in GCM models. Besides, under the high-resolution of lower than 50 km which cover a smaller region of the earth, RCM models are able to distinguish those smaller-scale processes more reasonably. Meteorological variables such as temperature, pressure, horizontal and vertical wind, heat flux, and humidity which simulated by a GCM are taken as the boundary conditions of the regional model. In which, the meteorological data and physics processes are then utilized to calculate this information to achieve the downscaled information results. The main advantage of regional climate models is that they are able to handle atmospheric equations and variation in land cover accurately. (Trzaska and Schnarr 2014).

Even when many improvements in the ability of RCMs models during these recent decades to better capture the regional climate condition, there still exist a lot of difficulties, drawbacks, uncertainties, and challenges to overcome. As the high resolution grid cells require additional surface information and the RCMs are also embed with more physical processes which is even higher than in GCM models at global scale, dynamical downscaling thus need to perform the enormous amount of calculations. Therefore, RCMs are usually extremely computationally expensive and may consume extensively computing time compare to GCMs (Wilby et al., 2009). RCMs also need a large amount of data, such as land-surface conditions with high frequency of meteorological variables from GCM. In addition, dynamic downscaling requires complicated calibration procedures to generate realistic simulations.

Similar to GCMs, RCMs often have troubles in calculating convective precipitation accurately. This is a big major problem for studying tropical climate zones. RCMs models also have low accuracy in simulating extreme rainfall events which is considered as a systematic bias that may further worsen at the higher resolution of climate projection. Biases correction methods are often required to calibrate the model results to match the observations (Brown et al., 2008). In several cases, small adjustments in the convective schemes would significantly enhance the reproducibility of the simulated rainfall. However, these changes need a lot of expertise while reducing geographic portability. For that reason, this generate a sub-version of the model which was properly calibrated to a specific region but might potentially perform awful in other places.

Applicability of RCM simulated results depends substantially on the quality of the GCM information fed in it. For instance, if the GCM poorly tracks the location of a storm, this will generate errors that continue to exhibited in the RCM simulation (Wilby et al., 2009). Furthermore, different RCMs models are often designed separately so they are having the

different dynamic schemes with wide range of physical parameters. This cause the variation in the downscaling results even if RCMs were driven by the same GCM output.

(c) Statistical downscaling

Statistical downscaling works by generating the empirical relationships between large-scale atmospheric variable of the past with the regional climate condition. When a relationship is created and verified, the large-scale meteorological variables provided by GCMs can be taken to forecast local climate conditions in the future. In another demonstrative way, GCM simulations provide predictors to continue simulate local climate information or predictands (Trzaska and Schnarr 2014). Statistical downscaling holds within it a diversified collection of methods which very different in clarification and applicability.

Statistical downscaling methods are easily to implement since it need a very small computation load. On the other hand, RCM model involving complex computing of physical processes that need significant amount of computational power and time. For that reason, statistical downscaling methods are a replaceable method for dynamical downscaling and sometimes advantageous alternative for researching organizations which cannot access to the computational power and technical knowledge needed for dynamical downscaling (Trzaska and Schnarr 2014). Different to RCMs which can provide downscaled details at a high spatial resolution of about 6-10 kilometres, statistical downscaling can provide up to station-scale climate information and even finer scale.

Even when statistical method is flexible, computationally cheaper, and having of a diverse group of methods, it usually inherits the implicit assumptions as follows:

- The empirical relationship generated from the large scale predictor and local climate condition remain constant over time and continue to hold true in the coming future.
- The predictor always carries the climate change signal.
- Relationship between the predictor and predictand is trong enough for making the realistic relationship
- GCMs generate the predictor with high accuracy

The first mentioned viewpoint is understood as the stationary hypothesis and indicate that the created empirical relationship remains stable which can be brought into the future. The fact that we need to accept that we cannot verify if this relationship will remain under future conditions. The second assumption is the GCM output can accurately represents the studying

climate condition and its simulated results could capture any variations that may be seen in the future. The third assumption believe that the significant of the statistical empirical relationship can be evaluated in advance to assess its validity and applicability. The last assumption is on the accuracy of GCM to reproduce historical climate events and also to simulate the future development of this climate condition. The validation process for predictors must be adopted first to demonstrate the use of GCM model in downscaling the climate characteristic of the selected domain. (Wilby et al., 2013).

1.1.5. Significance of coupling dynamical downscaling and statistical downscaling

Rainfall is one of the most important meteorological phenomena on Earth. It not only provides a vital freshwater source supporting all life forms, but also causes various types of natural disasters such as floods, landslides, storms, and drought. It is important to have a deep understanding of the rainfall formation mechanism to forecast the timing, density, intensity, and trends in a specific region to better manage water resources, maximize the use of water for economic development, and minimize the impacts of extreme events. In many countries, including Vietnam, rainfall is the object of regional planning strategies involving the production and construction sectors. Since the efficiency of water resource management depends on the accuracy and detail of rainfall forecasts, a method to obtain reliable and accurate predictions of rainfall at high spatial resolution is indispensable (Arritt and Rummukainen 2011; Caldwell et al. 2009; Giorgi and Mearns 1991).

Multiple general circulation models (GCMs) have been developed by various research groups to provide future climate predictions using numerical weather simulation. GCMs represent the physical processes and feedbacks for the atmosphere and oceans, which can be used to forecast future climate changes. Although GCM models can make useful predictions about global large-scale climate indicators, their spatial resolution of 100–200 km are too coarse to satisfy the requirements of regional planning. A GCM simplifies the complexities of land-sea distribution, vegetation cover, topography, and terrain. Therefore, downscaling methods, which translate coarse-scale GCM to finer spatial scales, have been developed to use on limited-area domains at higher horizontal resolutions.

Dynamical downscaling works by employing a regional climate model (RCM), which is based on the same principles as a GCM but has higher resolution over a limited area. An RCM uses large-scale atmospheric conditions as determined by a GCM for the lateral boundary

conditions. Higher resolution topography and land-sea distribution are incorporated to generate realistic climate information at a much finer spatial resolution (Seaby et al. 2013). Currently, RCMs are considered the most helpful method for producing climate information at the scales required for actionable strategic planning (Kjellstrom et al. 2016).

Over the years, the applicability of dynamical downscaling has significantly improved owing to the continuous development of computing technology and advances in numerical models. Even though the use of dynamical downscaling has become easier, it continues to be an extremely demanding method that requires considerable computational cost, simulation time, and output storage. Statistical downscaling is an alternative to dynamical downscaling for high-resolution climate downscaling that can overcome the drawbacks of dynamical downscaling methods. Statistical downscaling takes into account the empirical, spatial, and temporal relationships between large-scale climate indicators (predictors) and local-scale climate variables (predictands) and are trained on a historical period. Subsequently, these relationships are presumed to hold in the future, where they can be used to determine future predictands. Statistical downscaling methods are computationally inexpensive and significantly faster than dynamical downscaling, so they can be applied for even higher resolutions, up to station-scale. Since statistical downscaling methods rely on the assumption of an unchanged statistical relationship, they require long historical climate observation data for validation, which is not always available for every region. In contrast, dynamical downscaling operates based on physical realism with complex local processes, which allows it to map important fine-scale variations in climate that otherwise might not be included (Salathé et al. 2008; Pierce et al. 2012; Walton et al. 2017).

While statistical downscaling and dynamical downscaling methods are widely used in climatology research, both face drawbacks that limit their applicability. Recently, the approach of combining dynamical downscaling with statistical downscaling has been explored. Dynamical-statistical downscaling is a blended technique, where an RCM model is initially adopted to downscale the GCM output, followed by the application of statistical formulas to further downscale the RCM output to a higher resolution. Dynamical downscaling methods can utilize the advantages of RCM to provide better predictors for use in statistical downscaling (Guyennon et al. 2013). Berg et al. (2015) demonstrated this promising method by using a hybrid of the Weather Research and Forecasting (WRF) model with the Empirical Orthogonal Function to effectively forecast precipitation changes. In other research, Walton et al. (2015) introduced a new dynamical-statistical downscaling method by coupling WRF with Principal Component Analysis. The statistical-dynamical downscaling

method is another approach for blending techniques where dynamical downscaling is applied after a selected statistical downscale. While statistical-dynamical downscaling is a more complex blending technique, it is computationally less expensive. These methods use a statistical approach to refine the GCM outputs into a few characteristic states, which can be later used with the RCM models (Fuentes and Heimann 2000).

Limited efforts have been made to date to combine dynamical and statistical downscaling methods for precipitation research. In this study, we have introduced a combined dynamical-statistical downscaling technique for rainfall using WRF with an Artificial Neural Network (ANN). The WRF-ANN method aims to downscale high-resolution daily rainfall data for a seasonal length to satisfy the requirements for purposes such as agriculture or water resources planning. This method works by making statistical relationships between moderate- and high-resolution WRF outputs using ANN. The statistical relationships can be used directly to downscale moderate-resolution WRF outputs to fine-resolution rainfall. In this method, we first validated the accuracy of the WRF model to reproduce known climate conditions. Subsequently, the WRF output was downscaled to a finer spatial resolution using ANN. While this method used atmospheric variables from WRF, the relationship between physical and dynamical processes could potentially be included in the ANN. In addition, a bias correction for the ANN input and output (rainfall) was applied to reduce error in the final output. Moreover, the sensitivity of each predictor was also considered to examine their statistical relationships with rainfall.

1.1.6. Objectives

Firstly, the inundation and flood problems in future condition will be addressed in this study, using the combination of weather researching model and hydrological model driven by CMIP5 multi-model dataset.

There is a growing trend in the demands of using high-resolution climatological data for planning and management activities and that thus, raising the need of adopting weather downscaling techniques. The drawbacks of the available methods hinder the further application of these techniques while they are neither too costly nor too limited in reliability. The second aim of this study is to investigate the ability to couple both dynamical downscaling with statistical downscaling for high-resolution rainfall forecasting that can be easily adopted whilst maintaining the accuracy. The case study chose in the Red River Delta in Vietnam.

Finally, giving the key findings, research limitations and suggestion for future works in upgrading the proposed methods and extend its applicability.

1.2. Structure of the dissertation

Chapter 1 gives the brief information about the background and literature reviews of the related researches on climatology and hydrology that have motivated this research. The next section was then followed by the research's goals.

Chapter 2 considers the data and methodologies used in this dissertation. The material and methods to study the possibility of coupling dynamical and statistical downscaling for high resolution rainfall are reviewed. The major focus is placed upon the use of WRF and ANN in climatological information downscaling. For the second major goal of this dissertation, the data and methods required to perform investigation of rainfall runoff and inundation condition of Cau-Thuong-Luc Nam watershed are also illustrated.

The results for the first goal of this study – researching rainfall runoff and inundation condition of Cau-Thuong-Luc Nam watershed under global warming were discussed in Chapter 3.

In Chapter 4, results for the second goal of this study – coupling dynamical and statistical downscaling for high-resolution rainfall forecasting, were shown in detail with discussion.

In the last chapter, Chapter 5, the significant findings of this study were summarized with limitations and recommendation.

CHAPTER 2.
DATA AND METHODOLOGIES

2.1. Data and methodology in researching rainfall runoff and inundation in Cau-Thuong-Luc Nam watershed in Vietnam under global warming

2.1.1. Research data and materials

(a) In-situ daily observation data

The in-situ rainfall data collected by the Vietnam National Centre for Hydro-Meteorological Forecasting (NCHMF) at 56 rain gauges station from 2002-2009 provides a basis for this study. As river discharge data for verification of RRI simulation results, we used monitored data in July and August 2009 from Chu and Gia Bay hydraulic stations which locate in major branches of CTLN river system. Locations of hydro-meteorological stations are shown in Figure 2.1 (b). Northern Vietnam climate is distinguished by the Southeast Asia monsoon system with the hot-rainy season from Jun to August (JJA) then the cold-dry season from December to February (DJF). In this study, we focused on the hydro-meteorological condition of CTLN river system during JJA period.

(b) Coupled Model Inter-Comparison Project (CMIP5) multi-model dataset

For assessment of future global warming, this study based on the global warming experiments by the fifth phase of the Coupled Model Intercomparison Project (CMIP5). Numerous experiments provided by CMIP5 represent for the state-of-the-art international assessment of climate science. The future climate projections in CMIP5 are based on different future greenhouse emission scenarios known as representative concentration pathways (RCPs). Herein, we examined the daily precipitation for mid-21st century from 2060 to 2069. In Vietnam, progress of climate change is taking place faster than expected. Recently, Vietnam government has updated the high emission scenario RCP 8.5 in the national strategy on climate change to propose the action plan for extreme weather events under the worst case scenario (Thuc et al. 2017). In RCP8.5, the radiative forcing of the Earth become 8.5 W/m^2 larger than before the industrial revolution. In this study, we used the output of two Atmosphere Ocean General Circulation Model (AOGCM) projections from NOAA Geophysical Fluid Dynamics Laboratory Climate Model version 3 (GFDL-CM3) (Griffies et al. 2011) and a new version of

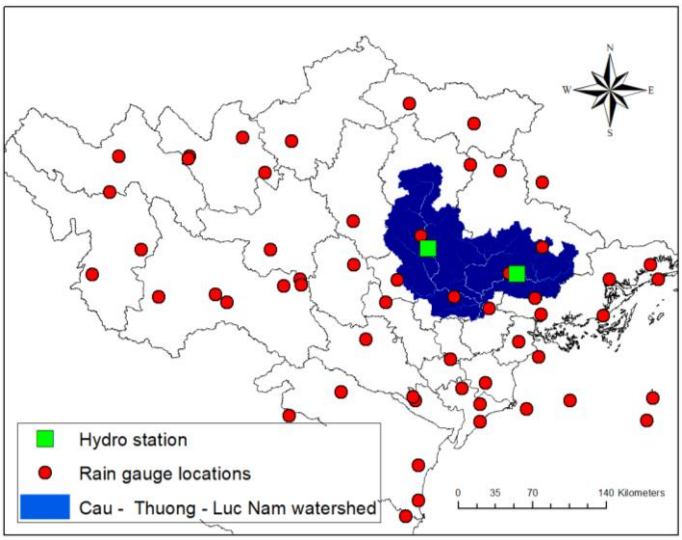

Figure 2.1 Locations of CTLN watershed and hydro-meteorological stations

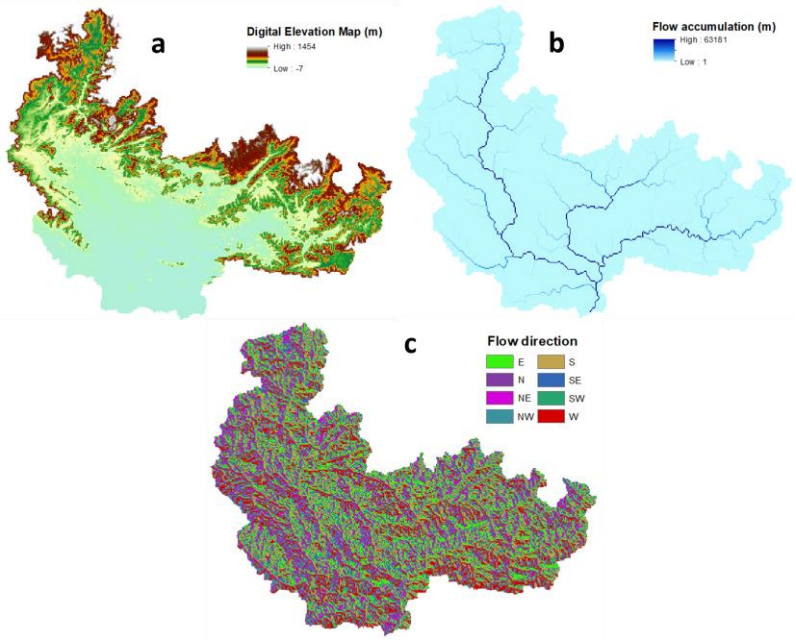

Figure 2.2 Topography and land use inputs for RRI model

Model for Interdisciplinary Research on Climate (MIROC-5) (Watanabe et al. 2010) for preparation of the future conditions.

(c) Initial and boundary conditions for numerical weather simulation

This study uses Japanese 25-year reanalysis (JRA-25) product by The Japan Meteorological Agency (JMA) for reproducing the present climate conditions which are also the base state for future (pseudo global warming) conditions. The global spectral resolution of JRA-25 kept at T106 with 40 vertical layers where 0.4 hPa was set for the height of the top level. Since the JRA-25 only provide the assimilation data from January 1979 to December

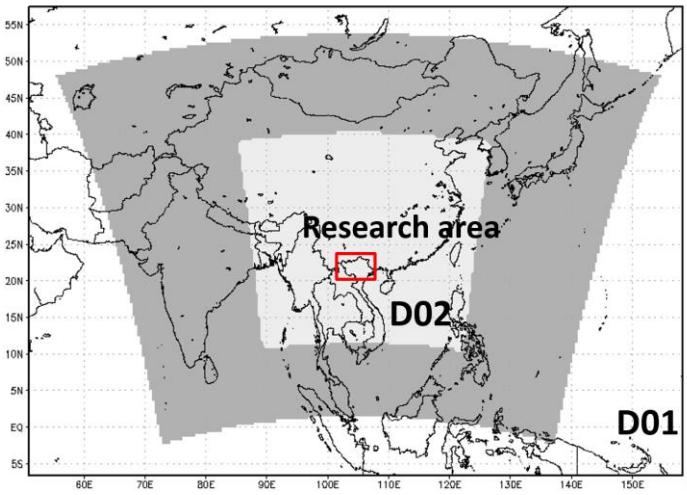

Figure 2.3 The target areas for WRF and ANN models, in which (a) The outer 30 km resolution (D01) and inner 10 km resolution domains (D02) are shown in the grey and white colors, respectively, the red rectangular indicate the locations of the researching area

2004, the reanalysis product from the JMA Climate Data Assimilation System (JCDAS) has been used for the period from January 2005. JCDAS deploys the similar system as JRA-25 which guarantee the homogeneous in quality and accuracy in the application of both datasets. The current climate conditions were obtained by performing dynamical downscaling to the 2000-2009 data. Simulation results for climate condition (or control simulation) is called CTL.

For land-surface boundary conditions for WRF model, we used the NCEP Final Operational Global Analysis data (NCEP FNL) (NCEP 2000). For the lower boundary condition over the ocean, the NOAA Optimum Interpolation 1/4 Degree Daily Sea Surface Temperature Analysis (NOAA OI SST) (Reynolds et al. 2007) was used.

(d) Topography and global land cover data

Input data for RRI model was taken from the hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS). Topography inputs of 15-arc second resolution (approximately 500m) for digital elevation model, flow accumulation, and flow direction was used for RRI model (Figure 2.2 (a, b, c)). Regarding the RRI requirement of land cover data, we used the Moderate Resolution Imaging Spectroradiometer (MODIS) based product of 0.5 km resolution.

2.1.2. Design of numerical simulation

(a) The Weather Research and Forecasting (WRF) model

Downscaling of present and future climate condition was implemented using the WRF model version 3.4 (Skamarock et al, 2008). A two-level, two-way nesting system for WRF downscaling is shown in Figure 2.3, where D01 and D02 represented for the 30km and 10km grid spacing resolution, respectively. The CTLN watershed locates in the Northern Vietnam where placed the center of the D2 domain. The WRF Double-Moment 5-class scheme (WDM5) microphysics and Grell-Devenyi cumulus parameterization schemes were used to calculate precipitation in the model. Planetary boundary layer parameterization is used from the Mellor-Yamada Nakanishi and Niino Level 2.5 PBL (MYNN2). The used parameterization for the surface layer and land surface are taken from Nakanishi and Niino PBL's surface layer scheme and Noah land surface model, respectively. The new Rapid Radiative Transfer Model (RRTMG) schemes are selected for long wave radiation and shortwave radiation conditions. Detailed WRF configuration setting for researching the inundation and flood condition in CLTN river basin is illustrated in Table 2.1

(b) High-frequency-anomaly Pseudo Global Warming (PGW)

Future climate projections given by Atmospheric Ocean Global Climate Models (AOGCMs) often contain uncertainties which resulted from uncertainties from future climate scenario, the lack of completeness in initial and boundary conditions, the imperfection of climate model, or the imperfect understanding of global and regional climate systems. AOGCM incompleteness is one problem addressed in DDS future weather conditions. The process of reproducing the historical climate conditions using AOGCMs aims to establish the

Table 2.1 Configuration of WRF model used in researching inundation and flood condition in CLTN river basin

Version of model	Version 3.4
Microphysics	WRF Single-Moment 5-class scheme
Cumulus parameterization	Grell-Devenyi schemes
Land model	Eta similarity
Land surface scheme	Nakanishi and Niino PBL's surface layer scheme
Long-wave radiation	Rapid Radiative Transfer Model (RRTM)
Short-wave radiation	Rapid Radiative Transfer Model (RRTM)
Planetary boundary condition	Mellor-Yamada Nakanishi and Niino Level 2.5 PBL
Length of simulation	92 days and 5 days for spin-up

general view of the climate we want to forecast, it is not to recreate the perfect reproduction. Hence, using DDS with initial and boundary conditions given by AOGCMs to evaluate the ability to reproduce the current climate condition is difficult. For that purpose, climatological reanalysis dataset can provide the input forcing for DDS of the past and current climates. For DDS future climate condition, Sato el al (2007) have introduced the pseudo global warming (PGW) method. In this method, the initial and boundary conditions for DDS were generated by combining 6-hourly reanalysis data and climatological monthly mean variations of future condition given by an AOGCM. The resulted PGW forcing data was feed in DDS to obtain final climate condition. Comparison results between dynamical downscaling applied for reanalysis data and AOGCM pointed out the better reproductive of DDS for rainfall when taking the boundary conditions from reanalysis dataset. In this study, PGW conditions for downscaling were processed with future climatological high-frequency variation. (Taniguchi 2016).

Downscaling simulations for present and future climate condition were prepared with pseudo global warming forcing. The lateral boundary conditions for future were constructed by adding projected changes in AOGCM simulations to reanalysis climate. In this study, the high-frequency anomaly pseudo global warming (HF-PGW) methods were applied to prepare the initial forcing with future high-frequency anomalies. The future inter-annual variability and diurnal cycle can differ from the present climate. At first, six-hourly AOGCM and reanalysis dataset (RD) were divided into climatological monthly mean plus the short perturbation terms:

$$AOGCM_P = \overline{AOGCM}_P + AOGCM'_P \quad (10)$$

$$AOGCM_F = \overline{AOGCM}_F + AOGCM'_F \quad (11)$$

$$RD_P = \overline{RD}_P + RD'_P \quad (12)$$

The subscript P and F represent the present (2000-2009) and the future (2060-2069), respectively. Then the bias-corrected six-hourly AOGCM data for the future period, $AOGCM*_F$, was calculated by adding the climatology variation between future and present ($\overline{AOGCM}_F - \overline{AOGCM}_P$) and short term perturbation $\overline{AOGCM}_F'$ to climatological mean reanalysis:

$$AOGCM*_F = \overline{RD}_P + \overline{AOGCM}_F - \overline{AOGCM}_P + AOGCM'_F \quad (13)$$

The inter-annual variation and diurnal circle were both included in HF-PGW conditions which are expected to greatly reduce the bias in AOGCM outputs.

2.1.3. River Runoff and Inundation (RRI) model setup

(a) Introduction of RRI model

Rainfall-Runoff-Inundation (RRI) model was used to simultaneously simulate the conditions of rainfall-runoff and flood inundation for the research domain. RRI is a two-dimensional hydrological model which can simulate rainfall-runoff and flood inundation simultaneously (Sayama et al., 2012, Sayama et al., 2015a, Sayama et al., 2015b). In RRI, slopes and river channel are processed separately. In a grid inside the river channel, both slope and river body are assumed within the same grid cell. The schematic diagram of RRI model is shown in Figure 2.4. The river channel is considered as a single line on the overlying slope cells. The 2-dimension diffusive wave model was adopted to calculate the flow over slope grid cells, while the 1-dimension diffusive wave model was applied for the main channel flow. For the simulation of river runoff and inundation processes, the RRI incorporate both lateral subsurface flow, vertical infiltration flow and surface flow. This boundary condition is especially important to present the hydrological condition of the mountainous areas, it is considered in the form of hydraulic-discharge gradient connection and take into account both saturated subsurface and surface flows. Besides, the vertical infiltration flow is estimated by using the Green-Ampt model. The interaction between the river flow in channel and slope is computed using on various overflowing formulation which depend on water height and current levee conditions.

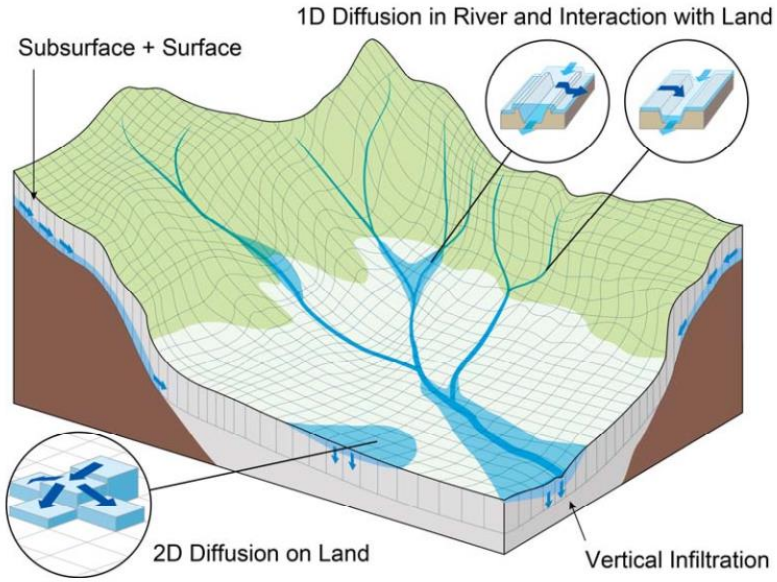

Figure 2.4 Schematic diagram of Rainfall-Runoff-Inundation (RRI) Model

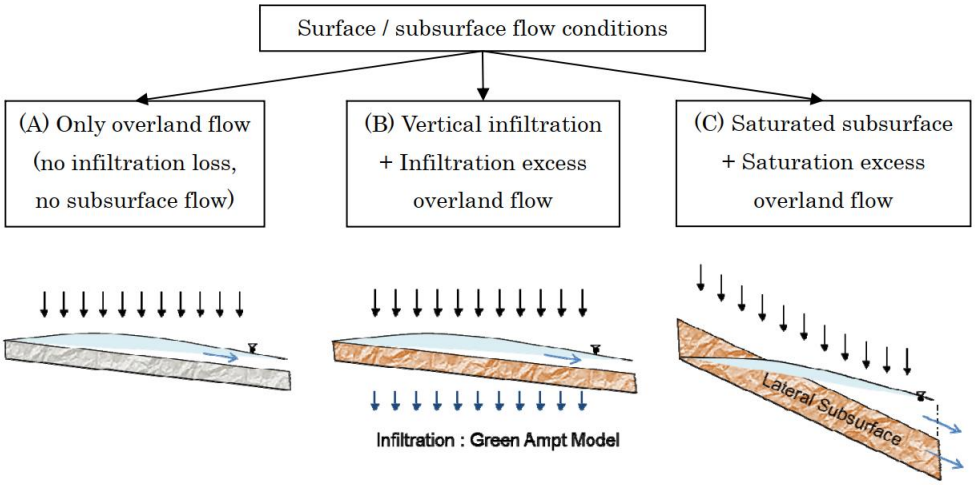

Figure 2.5 Surface and subsurface flow conditions considered in RRI

Governing Equation of RRI model

A method to calculate lateral flows formed on slope grid-cells is characterized as "a storage cell-based inundation model" (e.g. Hunter et al. 2007). The used equations were derived using the following mass balance equation (1) and momentum equation (2) which made for gradually varied unsteady flow.

$$\frac{\partial h}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_x}{\partial y} = r - f \tag{1}$$

$$\frac{\partial q_x}{\partial t} + \frac{\partial u q_x}{\partial x} + \frac{\partial v q_y}{\partial y} = -gh\frac{\partial H}{\partial x} - \frac{\tau_x}{\rho_w} \tag{2}$$

$$\frac{\partial q_y}{\partial t} + \frac{\partial u q_y}{\partial x} + \frac{\partial v q_y}{\partial y} = -gh\frac{\partial H}{\partial y} - \frac{\tau_x}{\rho_w} \tag{3}$$

In which, the water height from the local surface, q_x and q_y are the unit width discharges in x and y directions, u and v are the flow velocities in x and y directions, r is the rainfall intensity, f is the infiltration rate, H is the height of water from the datum, ρ_w is the density of water, g is the gravitational acceleration, and τ_x and τ_y are the shear stresses in x and y directions. The second terms of the right side of (2) and (3) are calculated with the Manning's equation. The model can also consider percolation and groundwater flow governed by Darcy's law. Then, the spatial differentiation in the mass balance equation (1) is discretized by the first-order finite difference method, and the time differentiation is solved by the fifth-order Runge-Kutta formula.

A one-dimensional diffusive wave model is applied to river grid cells. The geometry is assumed to be rectangle, whose shapes are defined by width W, depth D and embankment height H_e. When detailed geometry information is not available, the width and depth are approximated by the following function of upstream contributing area A [km^2].

Table 2.2 Original and Reclassified MODIS land cover types

	MODIS land cover dataset	
No	Original land use types	Reclassified land use type
1	Water	Crop land
2	Evergreen Needleleaf forest	Grass and shrub land
3	Evergreen Broadleaf forest	Mixed forest
4	Deciduous Needleleaf forest	Residental Areas
5	Deciduous Broadleaf forest	Flood plain
6	Mixed forest	
7	Closed shrublands	
8	Open shrublands	
9	Woody savannas	
10	Savannas	
11	Grasslands	
12	Permanent wetlands	
13	Croplands	
14	Urban and built-up	
15	Cropland/Natural vegetation mosaic	
16	Barren or sparsely vegetated	

$$W = C_w A^{Sw} \tag{5}$$

$$D = C_D A^{SD} \tag{6}$$

where C_W, S_W, C_D and S_D are geometry parameters. Here the units of W and D are meters

For each land cover class, decide (A), (B) or (C) in the following figure depending on infiltration and subsurface processes, so that the number of calibration parameters will be limited.

(b) Design of RRI model

Since the original MODIS product includes 16 types of land use which is too detailed to designate all the parameters, similar land cover types were combined into four major categories and also overlaid with river floodplain region. The detailed information about the original and reclassified land use types is shown in Table 2.2, the spatial distribution of reclassified land use types in demonstrated in Figure 2.6.

For better representation the rainfall-runoff-inundation processes of CTLN watershed, the surface/subsurface condition parameterization was enabled with vertical Green-Ampt

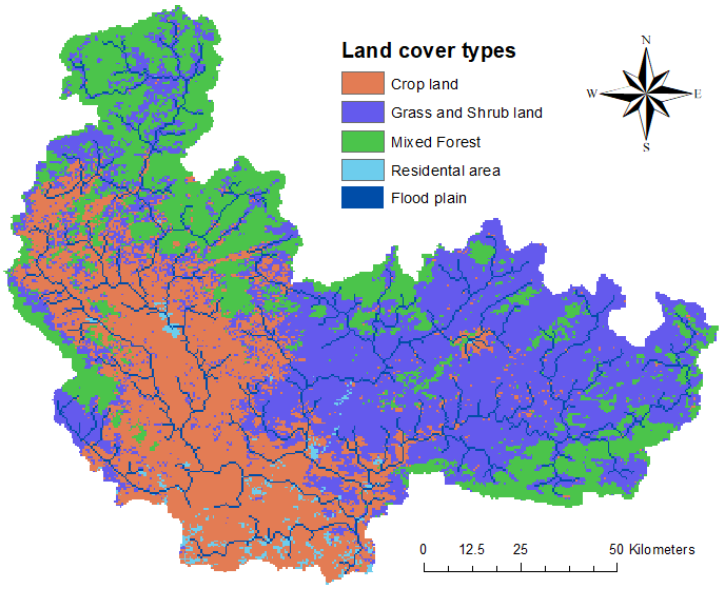

Figure 2.6 Spatial distribution of Reclassified Land cover types

model infiltration flow. Human management activities to control and regulate the river channel was considered by dam model. Evaporation parameter in RRI model was disable due to the lack of detailed monitored evaporation data. Simulation result of RRI for 2009 was used for model verification.

2.1.4. Conclusion

In conclusion, this part has demonstrated the method of using of the Weather Research and Forecasting model and River Runoff and Inundation model for researching river runoff and inundation condition for the Cau-Thuong-Luc Nam watershed under global warming.

The hindcast simulation for historical flood event in Cau-Thuong-Luc Nam watershed in 2009 was performed for the verification of models' configurations. The variation of precipitation during the rainy season in the future was investigated using the numerical weather simulations under the high frequency pseudo-global warming conditions, constructed with the fifth-phase results of Coupled Model Inter-comparison Project multi-model global warming experiments (CMIP5 – RCP8.5). The runoff and inundation conditions of the Cau-Thuong-Luc Nam watershed in the mid-21st Century were investigated by inputting the future rainfall conditions from WRF in to RRI model.

The simulation results for this section were indicated in Chapter 3 of the dissertation.

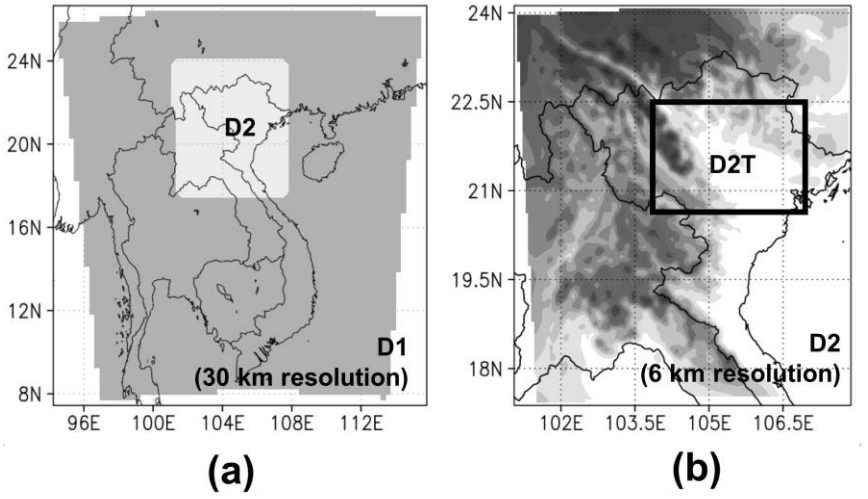

Figure 2.7 Target areas for downscaling with WRF and ANN. (a) The outer (D1) and inner (D2) domains are indicated by gray shade and white, respectively. The spatial resolution was 30 km for D1 and 6 km for D2. (b) The target area for ANN downscaling (D2T) is indicated by a rectangle inside D2.

2.2. Data and research methodology in coupling dynamical and statistical downscaling for high-resolution rainfall forecasting

2.2.1. Domain and data

(a) Study area

The research area used for downscaling by WRF is shown Figure 2.7 (a). The two-domain nesting method is applied with 30 km and 6 km horizontal grid resolutions for outer and inner domains (hereafter, D1 and D2 domains), respectively. While the D1 domain, placed in Southeast Asia, covers the entire Vietnam region, the D2 domain was selected in the Northern part of the country. The D2 domain has a complex topography, including alternating mountain range, midland, lowland, and a small proportion of the East Sea (Figure 2.7 (c). The study for precipitation estimation using ANN was carried out for a smaller rectangular area, which was placed inside the D2 (hereafter, D2T; Figure 2.7 (b)). This target domain was placed inland, defined between the latitudes 20.5 °N–22.5 °N and longitudes 104 °E–107 °E. The selection of D2T covering the large Red River Delta region including the capital Hanoi City. The area defined by D2T is not only the most important municipal areas in Vietnam but it also has the long and reliable climatological records. The weather in northern Vietnam is characterized by the tropical climate system, distinguished by the hot, rainy season from Jun

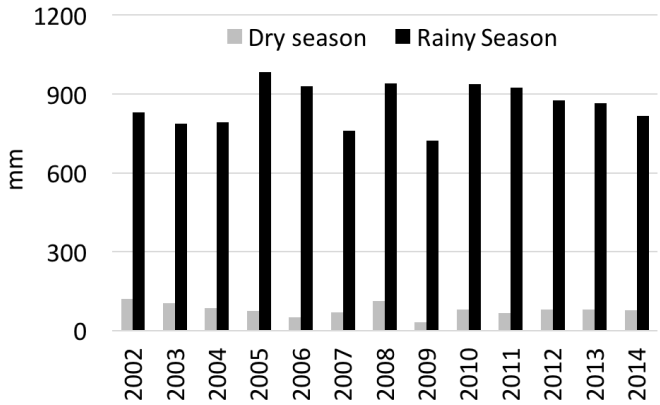

Figure 2.9 Averaged rainfall in JJA and DJF in northern Vietnam from 2002 to 2014

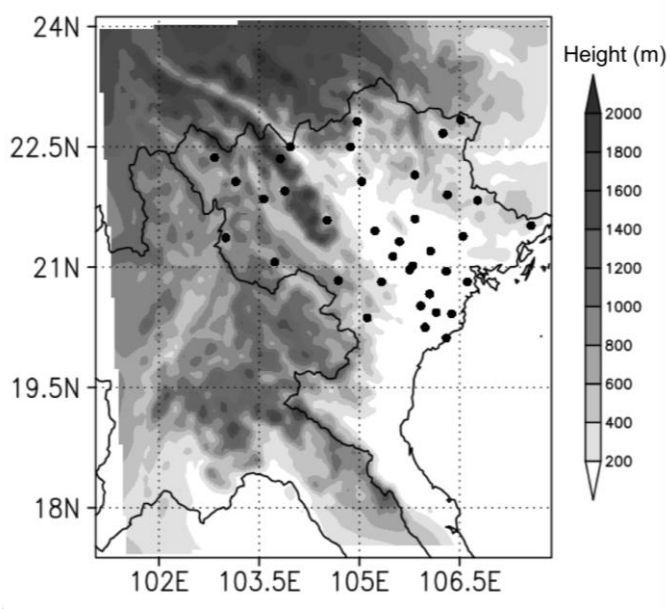

Figure 2.8 Geographical distribution of the 38 rain gauges providing data for this research are indicated by black dots.

to August (JJA), followed by the cold, dry season from December to February (DJF) (Figure 2.9). The average rainfall during ranges from 750 to 1.100mm which accounts for over 70% of the annual precipitation with about. Rainfall is the vital water source for the development of the region and also the cause of many water borne disaster. In this study, we focused on precipitation during the JJA period.

(b) In-situ observation data

The in-situ rainfall data was collected from the Vietnam National Centre for Hydro-meteorological Forecasting (NCHMF), using the rainfall measured by rain gauges. Rain gauge reports are prepared and recorded every six hours, being further processed by the NCHMF for monitoring climate anomalies. The rainfall data for the JJA season covering the years 1996, 1997, 1998, and 2006 provided the basis for this research. Specific criteria were applied to selected weather station data. These criteria include the following: (i) rain gauge stations positioned inside the D2 domain and must use the same monitoring techniques to minimize biases in the recorded data; (ii) a month during JJA is considered as sufficient data if the number of missing days is less than or equal to 5; (iii) a year is considered complete if all the months in JJA satisfy item (ii); (iv) a station that covers the all years of the research period, without missing any year, is considered to have complete data. After screening through these criteria, 38 stations were selected for the validation of WRF output (Figure 2.8).

(c) JRA-55

This study uses the Japanese 55-year Reanalysis (JRA-55, Kobayashi et al. 2015), developed by The Japan Meteorological Agency (JMA), for WRF initial and boundary conditions. Simulation outputs serve as the control run for testing the accuracy of WRF simulations (hereafter, simulations of climate in the past are called CTL). JRA-55 was improved from the JMA's former reanalysis (JRA-25 (Onogi et al. 2007) by deploying a more sophisticated data assimilation scheme to reduce biases in stratospheric temperature, as well as to improve the temporal consistency of temperature analysis. The spectral resolution of the global model projection in JRA-55 was maintained at T319L60 Gaussian grid data (equal to a 55-km horizontal grid) and 60 vertical layers, where 0.1 hPa represents the highest level of the model. The dataset employs the advanced four-dimensional variational data assimilation method, along with the global spectral model, to generate 6-hourly atmospheric variables and forecasting cycles. JRA-55 dataset is the third-generation global atmospheric reanalysis, which covers the period from 1958 to present.

The WRF control simulation is forced by the assimilation data obtained from JRA-55 for the years 1996, 1997, 1998, and 2006. The WRF CTL outputs for all the target years were used for downscaling validation, while the outputs for the first three years from 1996 to 1998 are used as inputs for artificial neural network (ANN) training. The CTL output for the year 2006 was utilized as an independent testing set for ANN.

(d) The sea surface temperature

For the lower boundary conditions of WRF simulation over the ocean, the NOAA Optimum Interpolated 1/4 Degree Daily Sea Surface Temperature Analysis (NOAA OI SST) was used (Reynolds et al. 2007). It is a global-scale reanalysis dataset, constructed by merging observations from various methods, including satellites, ships, and buoys. The complete global sea surface temperature (SST) map was produced by numerical interpolation. This product provides a spatial resolution of 0.25° × 0.25°, with a temporal resolution of one day. NOAA OI SST uses Advanced Very High Resolution Radiometer (AVHRR) infrared satellite SST data, which supports relatively high-resolution observation data. However, it cannot see through clouds. Therefore, since 2012, Microwave Instruments Advanced Microwave Scanning Radiometer (AMSR) has been used along with AVHRR to measure SSTs in most weather conditions.

(e) Land surface condition

For land-surface boundary conditions, we used the NCEP Final Operational Global Analysis data (NCEP FNL; NCEP, 2000). These boundary conditions are prepared at a spatial resolution of 1° × 1° gridded data in the form of reflected solar radiation, surface emissivity, the turbulent exchange of heat-moisture conditions in canopy and soil, and surface exchange momentum with the lower atmosphere. The magnitude of these processes forms the land-surface meteorological function.

2.2.2. Experimental setup

(a) Introduction of the Weather Research and Forecasting (WRF) model

The Weather Research and Forecasting (WRF) Model is a widely adopted mesoscale numerical weather prediction system which is developed for not only atmospheric research but also for practical forecasting purposes. WRF is featured by two dynamical cores, a system to execute data assimilation as well as a software structure that supports simultaneously computing and system extensibility. WRF model offers a huge range of application in terms of meteorology researches with the spatial resolution factor from dozens of meters up to thousands of square kilometres. WRF was initial developed in late 1990's which was born from a

collaborative joint research of the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (represented by the National Centers for Environmental Prediction (NCEP) and the (then) Forecast Systems Laboratory (FSL)), the (then) Air Force Weather Agency (AFWA), the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (FAA).

For deep meteorological researches, WRF model supports simulations using on real climate conditions such as from observations data and also from idealized conditions. WRF provides practical forecasting using a flexible and effectively computing system, while still update recent advancements in physics, numeric, and data assimilation contributed by developers from the vast user's research community. WRF model is utilized in practical simulation at NCEP as well as other meteorological centers around the world. WRF is also adopted in real-time forecasting configurations.

The core components of the WRF modelling system are illustrated in Figure 2.10. The WRF Software Framework (WSF) offers the working structure that connecting the dynamics solvers, physics packages that working with the solvers, programs for initialization, WRF-Var, and WRF-Chem. WSF contains two dynamic solvers: the Advanced Research WRF (ARW) solver developed primarily at NCAR, and the NMM (Non-hydrostatic Mesoscale Model) solver developed at NCEP. We adopted the Advance Research WRF in this study.

The Advanced Research WRF (ARW) modeling system has been developed and continue to improve for years, the latest release is Version 3 that available for free access in 2008. The ARW is constructed as a flexible, state-of-the-art meteorological simulation system

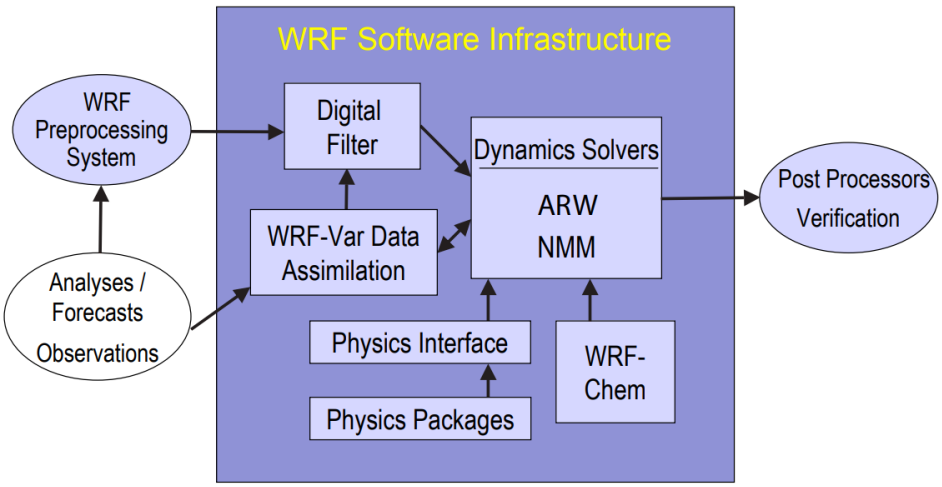

Figure 2.10 The infrastructure of the WRF software

that is easy to use and effectively build into the available parallel computing systems. ARW modules provide a huge range of applications of different spatial scales varies from several meters to tens of thousands kilometres, including:

- Idealized simulations (e.g. LES, convection, baroclinic waves)
- Parameterization research
- Data assimilation research
- Forecast research
- Real-time NWP
- Hurricane research
- Regional climate research
- Coupled-model applications
- Teaching

The detailed components of WRF modeling system is shown in the Figure 2.11. WPS is a program used primarily for real-data simulations. Its functions include 1) defining

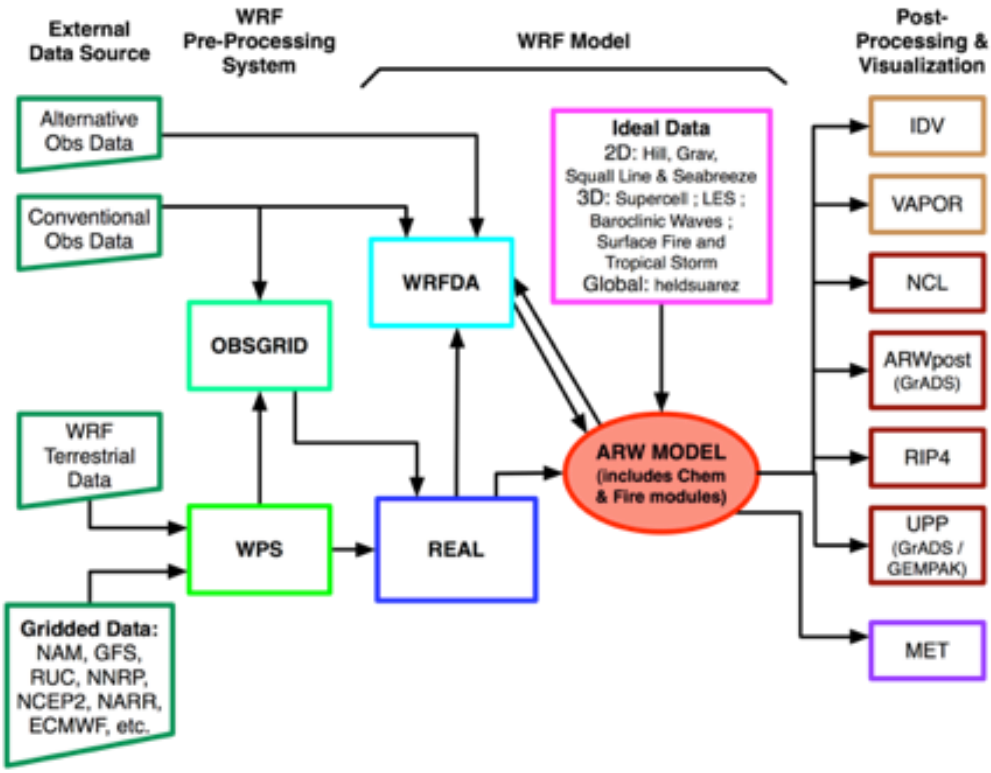

Figure 2.11 WRF system flow chart. In this study, External data is fed into the WPS module which output the domain containing meteorological data then this data is inputted to the ARW model solver.

simulation domains; 2) interpolating terrestrial data (such as terrain, landuse, and soil types) to the simulation domain; and 3) degribbing and interpolating meteorological data from another model to this simulation domain. Its main features include:

- GRIB 1/2 meteorological data from various centers around the world
- USGS 24 category and MODIS 20 category land datasets; USGS GTOPO30 elevation dataset; Global 5-minutes United Nation FAO, and North-America STATSGO 30 sec soil category dataset; 10-min greenness fraction data based on AVHRR and 30-sec greenness fraction data based on 10 years MODIS; MODIS-based leaf-area index; 0.15 degree monthly albedo and snow albedo data; and 1-degree deep soil temperature data; plus a few specialized datasets
- Map projections for 1) polar stereographic, 2) Lambert-Conformal, 3) Mercator and 4) latitude-longitude
- Nesting
- User-interfaces to input other static data as well as met data

ARW Solver is the key component of the modeling system, which is composed of several initialization programs for idealized, and real-data simulations, and the numerical integration program.

There are a various types of graphic tools to view and access the output. In this study, we utilized ARWpost module.

(b) Numerical weather simulation

In this study, the Weather Research and Forecasting (WRF) version 3.6 (Skamarock and Wang 2008) was used. The WRF model is widely used by both, operational and research communities. It represents the up-to-date techniques in mesoscale model development. WRF is a non-hydrostatic model, which was developed to inherit much of the dynamical and physical algorithms similar to the Fifth-Generation Mesoscale Model (MM5), introduced by the Pennsylvania State University/National Center for Atmospheric Research (PSU/NCAR). WRF has a wide range of physical options for parameterization, which can be combined in various ways.

In the context of this study, the selected physical parameterization setting is the optimal combination of schemes used in various studies across Asia. Cloud microphysics was used from the WRF Single-Moment 5-class scheme (Hong et al. 2004). The Kain-Fritsch scheme was used for cumulus parameterization. For surface layer physics, the ETA models based on

Table 2.3 Configuration of WRF model used in coupling DDS and SSD

Version of model	Version 3.6
Microphysics	WRF Single-Moment 5-class scheme
Cumulus parameterization	Kain-Fritsch scheme
Land model	Eta similarity
Land surface scheme	Noah Land Surface Model
Long-wave radiation	Rapid Radiative Transfer Model (RRTM)
Short-wave radiation	Dudhia's scheme
Planetary boundary condition	BouLac scheme
Length of simulation	92 days and 5 days for spin-up

Monin-Obukhov, with Carslon-Boland viscous sub-layer, were used. As a land surface model, the Noah Land Surface Model (Chen and Dudhia 2001) was applied. The Bougeault and Lacarrere (BouLac) (Bougeault and Lacarrere 1989) scheme was used for the planetary boundary layer. Dudhia's scheme (Dudhia 1989; Mlawer et al. 1997) and the Rapid Radiative Transfer Model (RRTM) were selected for short-wave radiation and long-wave radiation conditions, respectively. The spectral nudging option was enabled to include the global-scale effects on a smaller scale, to ensure the simulated result is more consistent with the observation (von Storch et al. 2000). The outline for the model configurations is provided in Table 2.3.

Since this research targets the rainfall season, dynamical downscaling was applied to each JJA period of the research duration. For the initial and boundary conditions, downscaling simulations used JRA-55, NCEP-FNL, and NOAA OI SST datasets.

The goal of the WRF model is to accurately reproduce detailed information about rainfall in the target domain. Here, we evaluate the WRF's ability to reproduce daily rainfall by comparing to surface observations data (see Figure 2.7 (c)) for the spatial distribution of rain gauges). The downscaling experiments were implemented for each JJA period in 1996, 1997, 1998, and 2006. While evaluating the reproducibility of the WRF model, we omitted the results from the first five days, or a spin-up period of 5 days from May 27th to May 31st. After the spin-up period, the 92-day simulation was carried out from the first day of Jun to the end of the JJA period on 18Z, August 31st. The model performance was evaluated by computing a series of statistical measurements for simulated rainfall against the observed rainfall, including the mean absolute error (MAE), (Pearson) correlation (R), root mean square error (RMSE), and the index of agreement (IOA). The statistical measures are defined as follows:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|O_i - S_i| \tag{7}$$

$$R = \left[\frac{\sum_{i=1}^{N}(O_i - \bar{O})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{N}(O_i - \bar{O})}\sqrt{\sum_{i=1}^{M}(S_i - \bar{S})}}\right] \tag{8}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - S_i)^2} \tag{9}$$

$$IO = 1 - \frac{\sum_{i=1}^{N}(O_i - S_i)^2}{\sum_{i=1}^{N}(|O_i - \bar{O}| + |S_i - \bar{O}|)^2} \tag{10}$$

where N is the number of grid observation sites O; $\bar{O}$ and $\bar{S}$ correspond to the average rainfall measured from the rain gauges and simulation result, respectively. The IOA index returns the degree of model prediction error, varying between 0 and 1, with the higher value indicating better agreement between the model predictions and observations, while the lower value indicates weak agreement.

WRF outputs have a higher precision than the observation. Such difference in resolution of rainfall may cause biases in WRF output. Rain gauge sensors currently in use can only detect the accumulated rainfall of more than 0.5 mm per day. In contrast, the frequency of wet days with very low rainfall, for example a resolution of 0.0001 mm per day, can be projected by WRF. This means the WRF rainfall output might not be consistent with the observation, even if its projection perfectly matches with the reality. To reduce the bias in WRF output, while negating the accumulation effect when downscaling with ANN, all rainfall output values less than 0.5 mm (wet-day-threshold) were treated as dry day events (hereafter, DRE). This wet-day-threshold was applied to all the WRF simulated data in this research.

(c) Artificial Neural Network (ANN)

The term Artificial Neural Network (ANN) is a mathematical concept of artificial intelligence that mimics the network of billions of interconnected neurons in the human nervous system. In this, the basic information processing system, like the human brain, is composed of nerve molecules as the core processing unit. In ANN, these core processing units are called neurons, which possess the natural characteristics of processing the information and

storing analyzed data as experimental knowledge for later use. ANN comprises of a number of simple processing elements connected to each other by different weights, according to a specified architecture. During the learning process, ANN gradually adjusts the connecting weights to reproduce the training dataset effectively. ANN has been considered as one of the most effective alternatives to traditional methods for solving the nonlinear time series problems that can yield results in a short period (Liu et al. 2009; Yu et al. 2006).

ANN offers a variety of network architectures suitable to different fields of application. In this study, we adopted the most widely implemented architecture of feed-forward artificial neural network (FFANN) in the climatology field (Abhishek et al. 2012), which is a multi-layer perceptron, trained using the back-propagation learning algorithm (MLP-BP), for downscaling WRF rainfall output from the mother domain to the child domain. When first introduced by Rosenblatt (1958), the simplest ANN model, the single layer perceptron (SLP), offered only a strictly limited transfer function, which was suitable to deal with linear problems. The MLP overcomes many limitations of the SLP by adding one or more hidden layers of neurons between the input and output layers, which includes the capabilities of solving the non-linear problems, based on the non-linear relationships among the neurons (Lippmann 1987).

Figure 2.12 depicts the simplified architecture of an MLP-BP network implemented in this study. The network contains a set of neurons organized in layers from the input layer on the left to the output layer on the right. All the processing neurons are fully connected with other neurons in the following layer, while there is no connection between neurons of the same

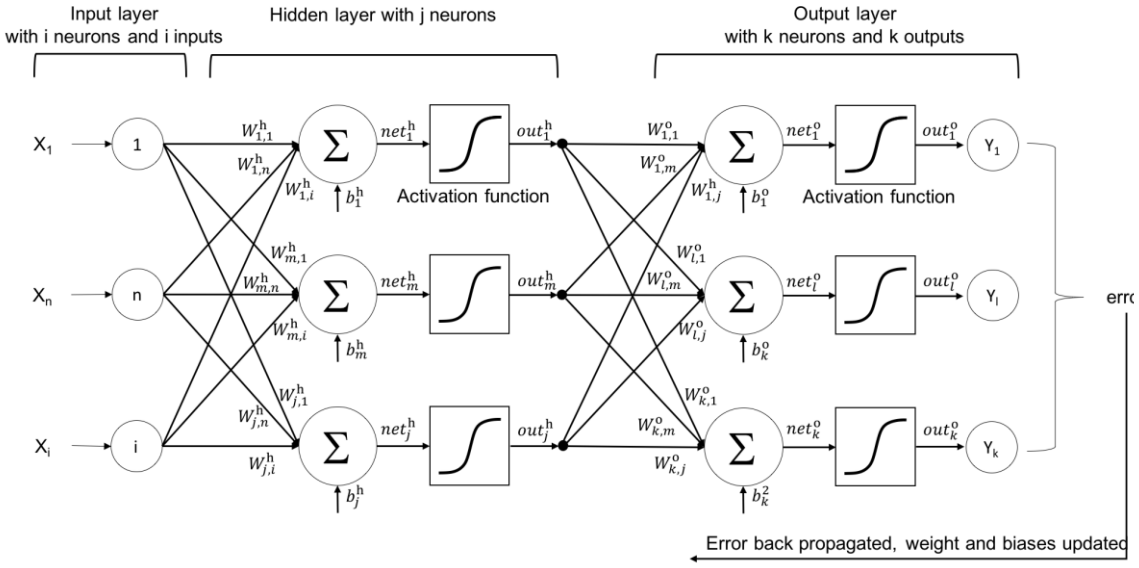

Figure 2.12 Simple multilayer perceptron ANN

layer. The input layer is neither designed for processing data nor generating inputs of its own. It simply stores the input values to be processed in the next layer. After the input layer, one or more processing layers, called the hidden layer, follow. The last layer is the output layer, containing processing neurons to generate a simulated value. The connections between neurons are made with the associated weights. The network illustrated in Figure 2.12 represents a three-layered ANN with an input layer of *i* input neurons (X_1, X_2, …, X_i), one hidden layer with *j* neurons (H_1, H_2, …, H_j), *k* output neurons (Y_1, Y_2, …, Y_k), with connections from the output of one layer to the next layer. The superscript *h* and *o* indicate that the calculations are implemented in the hidden layer and output layer, respectively. The input values calculated by the model for the m^{th} neuron in the hidden layer are the weighted sum of *i* inputs to which the bias value b_m^h is added:

$$net_m^h = \sum_{n=1}^{i} W_{m,n}^h X_n + \mathrm{b}_m^h \; ; = 1, 2, \ldots j \,, \tag{11}$$

where $W_{m,n}^h$ is the associated weight matrix for the connection between the input neurons and the neurons in the hidden layer. Then, the net_m^h vector is entered into a non-linear activation function g( ), which is essential for an ANN model to solve the nonlinear problems, wherein the most useful and widely adopted functions are hyperbolic tangent or logistic sigmoid (Bodri and Čermák 2001). The output of neuron out_m^h in the hidden layer subsequently becomes:

$$out_m^h = \mathrm{g}(net_m^h) \tag{12}$$

The input of the l^{th} neuron in the output layer is calculated as the weighted sum of those activations plus the bias neurons b_l^o:

$$net_l^o = \sum_{m=1}^{j} W_{lm}^o out_m^h + b_l^o \; ; \; l = 1, 2, \ldots k. \tag{13}$$

The same activation function g( ) that was applied to the hidden layer is applied to the input layer. The final network output out_l^o for the l^{th} output of the model is subsequently obtained using the following function:

$$out_l^o = \mathrm{g}(net_l^o) \tag{14}$$

As the goal of training is to minimize the difference between the actual output (desired output) and the simulated output of the network, the network error is computed at this stage.

Table 2.4 Predictor variables considered in the preliminary test

No.	Parameter	Description	Unit	Correlation with RD2T
1	Rd1	Rainfall taken from D1	mm	0.89
2	v10	Vertical wind speed at 10m	$m.s^{-1}$	0.02
3	u10	Horizontal wind speed at 10m	$m.s^{-1}$	0.04
4	hgt	Terrain height	m	0.01
5	t_diff	Temperature different between tk and t2	K	-0.12
6	slp	Sea level pressure	Pa	0.18
7	tk	Temperature at 1400m height	K	0.23
8	t2	Temperature at 2m	K	0.08
9	q2	Specific humidity at 2m	$kg.kg^{-1}$	-0.2
10	psfc	Surface pressure	Pa	0.12
11	vasso	Variance of sub-grid scale orography	m	-0.04
12	pblh	Planetary boundary layer height	m	-0.05
13	tslb	Soil temperature	K	-0.19
14	smois	Soil moisture	$m^3.m^{-3}$	0.21
15	grdflx	Ground heat flux	$W.m^{-2}$	0.24
16	canwat	Canopy water	$kg.m^{-2}$	0.14
17	sfroff	Surface runoff	mm	-0.05

This error is subsequently inserted back into the input layer, where the initial connecting weights and biases are adjusted according to the magnitude of the error. The supervised learning is repeated until the ANN is converged, reaching an error smaller than the threshold. In this study, the connection weights were updated after each training itinerary. ANN was trained by using the Levenberg Marquardt training algorithm, which is proven to be a fast and efficient update rule for medium-sized FFANN (Yu and Wilamowski 2011).

(d) ANN downscaling experiment

The goal of an optimal ANN architecture is to minimize the error between the simulated output and the desired value with the most (possible) compact and simple structure. There are several most essential factors affecting the performance of ANN, like, (1) selection of the predictor, (2) building the number of layers and neuron structure (network structure), and (3) specifying the training algorithm for connecting weights. Input predictors are usually independent variables, believed to have some predictive power over the dependent variable being predicted (predictand). Normally, useful predictors could be selected by looking at the

correlation and cross-correlation between the predictors and predictand. However, the combination of two or more uncorrelated predictors might potentially become a strongly correlated variable (Castellano and Fanelli 2000). In contrast, two or more highly correlated predictors might exacerbate the small change in the model, potentially increasing the error. With regard to the network structure issue, while an insufficient number of hidden neurons might lead to low accuracy in training, an excessive number of hidden neurons tend to add unnecessary training time, with marginal improvement or memorizing, instead of learning (overfitting; Castellano and Fanelli 2000). There is no specific method to find the optimal number of layers and hidden neurons, except for the commonly used trial and error approach (Zhang and Goh 2016). On the last issue related to training algorithm, there are several training functions available to obtain the connection weights, as well as to adjust the weights. The selection of training algorithm is made based on the type of network, input data, and occasionally, computer power.

In this study, CTL outputs of D1 and D2 from 1996 to 1998 were used for ANN training. Since the application of ANN for rainfall downscaling was limited within the D2T region (Figure 2.7 (b)), the high-resolution rainfall output in the rectangular region of D2T was prepared as the predictand variable for ANN, with the predictand rainfall being called RD2T. Predictor variables, on the other hand, were taken from the coarser domain D1. The principle behind selecting the variables to project RD2T is briefly illustrated in Figure 2.13, in which, a grid in D2T (predictand grid) is simulated by the four adjacent grids (predictor grids) in D1.

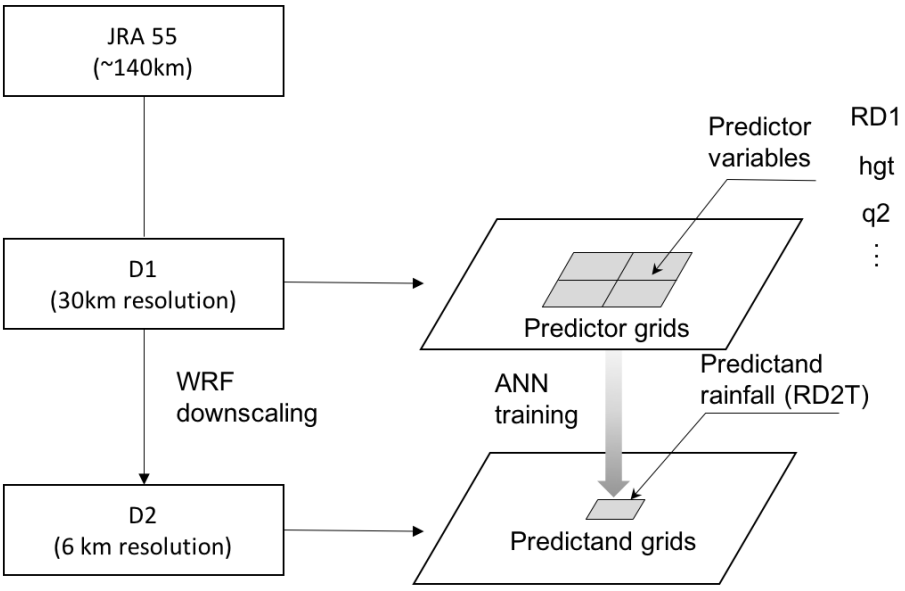

Figure 2.13 Predictor and predictand grid selection principles

Since the predictors considered in this research were present in a large number, we depended on the correlation coefficient between RD2T and the proposed predictors. Correlation coefficients were calculated between RD2T and the mean value of four predictor grids for each variable. Firstly, the effectiveness of combining several highly correlated predictors was examined. The screening would continue with other combinations, including those uncorrelated with the correlated predictors. This selective combination of predictors aims to project data onto a lower dimension space, while retaining as much information as possible by eliminating the correlated information caused by overlapping input instances.

The variables considered in this study are described in Table 2.4, along with their correlation coefficients with RD2T. Predictor variables were subsequently selected and tested using the trial-and-error method, from a simple network of several correlated variables to the larger sets, including the combination of uncorrelated variables. To increase the efficiency of the training process, all the selected variables were normalized using the feature scaling method described in Equation 9 to transform all the values into the range [0:1], where *a(g)* is the original value before normalization occurs, while z_g is the normalized value of *a(g)*. The reason for normalization is to avoid a very high resultant value, when the original data is entered into ANN, which could potentially cause the activation function to exhibit a low performance to resolve the small changes in input data, thus losing sensitivity.

$$z_g = \frac{a(g) - min(a)}{max(a) - min(a)}, \text{where a } = (a_1, \ldots, a_n) \tag{9}$$

In the next step, the number of hidden layers and neurons were determined according to the quantity of variables, gradually increasing the network size from a small neuron number until the desired accuracy is obtained. A common method for ANN training is to separate data into independent training and testing sets. However, it has been proven that better-trained models are not necessarily associated with better estimation capability. An excessively complex network with high parameters/observations ratio might lead to overfitting, which is a model with very low predictive power, even when it was well-trained. A practical way to avoid overfitting, while improving the estimation capability of the network at the same time, is to create a small set of data from the training set for cross-validation. The errors in the training set and validation set are compared during the training stage. As the error in the validation set continues to increase, the training process will be stopped when the best performance of the network is achieved. We adopted the cross-validation approach to train the ANN model in this study. The data used in the training stage was randomized to avoid bias, subsequently being

divided into 3 independent datasets, in which 75% of the data was set as the training set, 15% as the testing set, with the remaining 10% being set aside for cross-validation. During the preliminary stage, numerous ANN models were tested to find the most effective network design for rainfall downscaling. In this section, several distinctive models, whose structures were considered as the most presentable for the trial-and-error tests, are presented, with the details of these models summarized in Table 2.5.

The M1 model series includes 3 models, M1n, M1a, and M1s. The first model (M1n) was the simple MPL network structure, composed of 16 neurons of input layer and one neuron of the output layer, with 3 hidden layers positioned between the input and output layers, containing 10, 5, and 3 neurons, respectively (16-10-5-3-1). The transfer function adopted for neurons was the logistic sigmoid (LS) function. M1n was provided with predictor variables extracted from the 4 predictor grids, including sea level pressure (slp), temperature at 2 m (t2), terrain height (hgt), and planetary boundary layer high (pblh), which brought the total number of predictors to 16. The variables directly extracted from the CTL result will be hereafter called NV. The network output was RD2T, which was the daily precipitation in the D2T domain. The following two models, M1a and M1s, were designed based on the structure of M1n, with different input features and number of neurons. M1a utilized the averaged input of the 4 predictor grids (AV), while M1s utilized the standard deviation of the 4 predictor grids (SV) as inputs. Therefore, both, M1a and M1s, are one-fourth the size of M1n (4 features, as opposed to 12 features). The MLP structure was adjusted with respect to the network size reduction (4-4-4-4-1).

The M2 model series (M2n, a, s, d, e) aimed to further test the differences between the types of input features (NVs, AVs, and SVs) by increasing the complexity of the training set. The models, M2n, M2a, and M2s, adopted the same LS transfer function and input variables as the M1 model series, with the addition of D1 rainfall (Rd1). The features used for each model are NV, AV, and SV, respectively. M2d differed from M1n by replacing the LS transfer function with the hyperbolic tangent sigmoid (HTS) function. M2e was slightly redesigned from M2d by increasing the neurons in the hidden layers (20-20-10-5-1).

The subsequent model, M3n, also used the HTS function and NV features of the M2e model, with different input variables and architecture. The reason behind the changes in the input combination was to examine the contribution of different correlated variables to the network accuracy. Hung et al. (2009) depicted the essence of self-learning algorithm allowing ANN to forecast without significant prior knowledge of all the processes involved. However, a deep

understanding of the interaction between the physical processes is helpful in identifying the appropriate variables to be used in the training stage. Thus, a batch of multiple combinations of physical processes was tested. The M3n model is one representative of the test process, which uses 7 variables including slp, temperature at 1400 m (tk), hgt, pblh, vertical and horizontal wind speed (v10 and u10), and ground heat flux, while excluding R1d—the variable expected to be most closely correlated to RD2T. Since the dimension of the dataset is bigger compared to the M2e model, the numbers of neurons in the network layers of M3n were increased (28-20-10-7-1).

The models in the M4 series, including M4n, M4a, and M4as, have the same input variables, which continued to increase by the rule-of-thumb, adding up to a total of 8 variables, including Rd1, tk, hgt, slp, grdflx, surface pressure (psfc), pblh, and q2. The M4n model adopted the NV variables, adding up to 36 input features. As the network size increased, the network architecture was adjusted to meet the requirements (32-25-20-10-1). The M4a models inherited the AV features of 8 variables, while the M4as included both, AV and SV features, which doubled the input features to 16. The M4 model series used the same HTS transfer function. The M5n models used the same setup as M4n, although it was trained for rainfall events (RE) only. Contrary to DDE, daily precipitation higher than the wet-day threshold of 0.5 mm were considered as RE events.

ANN models developed with the 1996–1998 dataset were applied to the WRF output of 2006. The aim of this application was to study in depth, the stability and applicability in actual WRF-ANN downscaling. Since the testing stage targets the model reproducibility for any weather condition, the model that was designed to train with rainfall events only (RE-ANN), i.e. the M5n model, is expected to experience difficulty in reproducing DDE cases. Therefore, its simulation output will be calibrated to minimize the bias. We tested the correlation of DDE cases in both, D2T and D1, finding a very strong connection between DDE grids at the high-resolution scale with the DDE grids at the coarse scale. Spatially, there are 99.44% DDE grids in D2T, located completely inside the DDE grids in D1. Additionally, any grid in D2T partly overlapping with an DDE grid in D1 also has 98.27% chance of being an DDE grid. This result demonstrates that any grid in the mother domain has a significant similarity in rainfall condition to the child grids within and around it. Therefore, any grid in the RE-ANN model output, with a spatial connection to an DDE grid in D1, will be treated as an DDE grid. This treatment process of the DDE grids in the model output is called RE-ANN calibration. The RE-ANN calibration method was also applied to the M5n model during the preliminary test.

Table 2.5 Distinctive ANN models considered in the preliminary test

Model	Architecture[a)]	Transfer function	Training data	Training set dimension
M1n	16-10-5-3-1	LS[b)]	slp, t2, hgt, pblh	4 NVs[d)] x 4 grids
M1a	4-4-4-4-1	LS	slp, t2, hgt, pblh	4 AVs[e)]
M1s	4-4-4-4-1	LS	slp, t2, hgt, pblh	4 SVs[f)]
M2n	20-12-8-5-1	LS	Rd1, slp, t2, hgt, pblh	5 NVs x 4 grids
M2a	5-5-5-5-1	LS	Rd1, slp, t2, hgt, pblh	5 AVs
M2s	5-5-5-5-1	LS	Rd1, slp, t2, hgt, pblh	5 SVs
M2d	20-12-8-5-1	HTS[c)]	Rd1, slp, t2, hgt, pblh	5 NVs x 4 grids
M2e	20-20-10-5-1	HTS	Rd1, slp, t2, hgt, pblh	5 NVs x 4 grids
M3n	28-20-10-7-1	HTS	slp, tk, hgt, pblh, u10, v10, grdflx	7 NVs x 4 grids
M4n	32-25-20-10-1	HTS	Rd1, tk, hgt, slp, grdflx, psfc, pblh, q2	8 NV x 4 grids
M4a	8-6-5-5-1	HTS	Rd1, tk, hgt, slp, grdflx, psfc, pblh, q2	8 AVs
M4as	16-10-5-3-1	HTS	Rd1, tk, hgt, slp, grdflx, psfc, pblh, q2	8 (AVs+SVs)
M5n	32-25-20-10-1	HTS	Same as M4n but for rainfall events	

[a)] Model architecture indicates the number of neuron in each layer of a 5 layers MLP network, wherein first number is neurons of input, the following 3 numbers are neurons of hidden layers, and the last one is neurons of the output layer.

[b)] LS: Logistic Sigmoid.

[c)] HTS: Hyperbolic Tangent Sigmoid

[d,e,f)] NV, AV and SV correspond to normal variable, averaged variable and standard deviation variable, respectively; in which:

- NVs are directly extracted from the 4-predictor grids, so the actual number of variable are multiplied to 4;

- AVs are the new features created by computing the mean of each variable in the 4-predictor grids;

2.2.3. Conclusion

In conclusion, this part has outlined the general technical principles of the Weather Research and Forecasting model and the Artificial Neural Network method. The detail configuration and setup for both WRF and ANN were given.

The method and steps of experimenting and verification of the coupling WRF-ANN method were presented. A two-levels, two-ways WRF downscaling was implemented for a selected region with 30km and 6km spatial resolution for mother domain and child domain, respectively. It was followed by the ANN training process which utilized WRF output as input data. Then, the results from ANN training stage were applied to an independent dataset to further check the effective of the combined method.

The simulation results of coupling dynamical and statistical downscaling for high-resolution rainfall forecasting were presented in Chapter 3 of this dissertation.

CHAPTER 3.
RESULTS AND DISCUSSION OF RESEARCHING RAINFALL RUNOFF AND INUNDATION IN CAU-THUONG-LUC NAM WATERSHED IN VIETNAM UNDER GLOBAL WARMING

3.1. Introduction

This chapter gives the detailed results and discussion of the research on rainfall runoff and inundation in Cau – Thuong – Luc Nam watershed in Vietnam. WRF model was applied for the present and the future climate condition over the Cau – Thuong – Luc Nam river basin. Rainfall outputs from WRF were feed in RRI model to investigate the hydrological condition of the river during the rainy season. In this chapter, the total rainfall during rainy season in the future was taken into account with the duration and spatial distribution of extreme flood events.

3.2. Simulation result of historical flood inundation

In this section, we verify the accuracy of rainfall downscaling by WRF from 2002-2009 by using the observation data (OBS) from 56 rain gauges. The RRI simulation results has also been examined for the 2009 flood event at Chu and Gia Bay hydro stations.

3.2.1. Reproducibility of WRF model

Figure 3.1 compares the JJA rainfall averaged for 56 rain gauges and corresponding CTL grids from 2002 to 2009. It can be seen from the results that JJA rainfall reproduced by CTL varied from 72.1% to 90.0% of observation data and consistently underestimated. Average spatial correlation coefficient for 8 years (2002-2009) calculated between CTL and OBS was 0.78, indicates the relatively good agreement between simulation results and observation. Regarding the temporal variation of rainfall, correlation coefficients between CTL and OBS averaged for JJA periods in 8 years in 56 observation sites range from 0.51 to 0.9 with an average of 0.71 (details not shown). Verification result indicates the WRF model can be used for rainfall downscaling in CTLN watershed with reasonable accuracy.

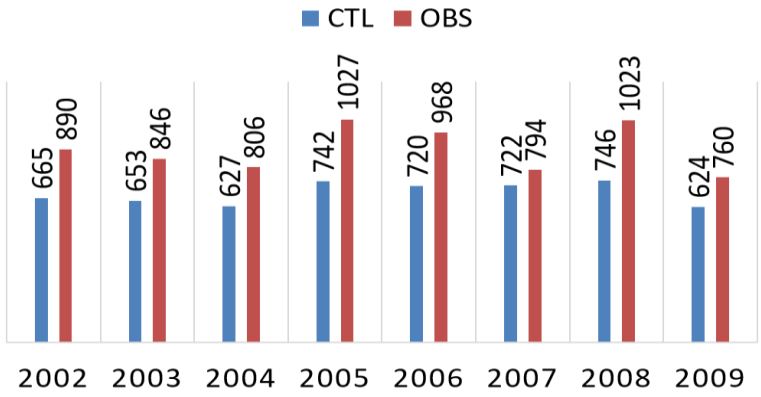

Figure 3.1 JJA rainfall averaged for 56 observation sites and corresponding grids in CTL from 2002 to 2009 (mm)

3.2.2. Simulation of RRI models for historical flood events

Observation rainfall data in July-August 2009 was fed into RRI model to examine the accuracy of the model to represent the flood events. Comparison of river discharge between RRI simulated results (RRI-CTL) and observation data was shown in Figure 3.2. The Nash-Sutcliffe efficiency coefficient (NSE) was used to quantitatively describe the predictive power of RRI model outputs for river discharge. NSE indexes in Gia Bay and Chu stations are calculated 0.77 and 0.73, respectively. NSE results showed the good match between simulated results and OBS in the decreasing trend of river discharge at the end of the rainy season. Since the NSE index in sensitive to extreme value, the high NSE results also suggested that extreme discharge periods were relatively well predicted. Simulation results slightly overestimated the peaks of river discharge in both Gia Bay and Chu stations. However, peaks of large and small

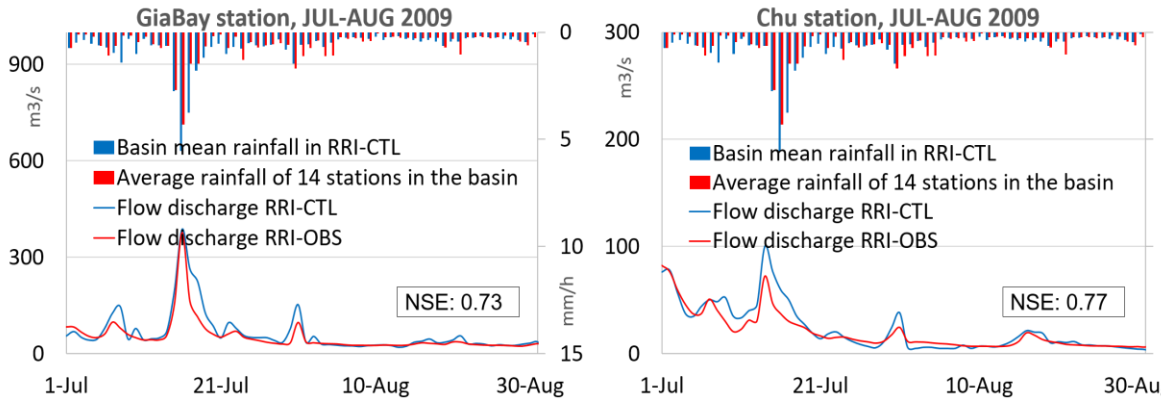

Figure 3.2 Basin mean precipitation and river discharge by RRI model and observation data in Gia Bay and Chu stations during JJA in 2009

floods are correctly reproduced for each stations in RRI simulation. Verification results show the good applicability of adopting RRI model to further investigation of the future hydrological condition of CTLN river system.

3.3. Future flood forecasting

3.3.1. Future trend in rainfall intensity

WRF daily rainfall outputs during JJA for 10-years periods of CTL (2000-2009), GFDL-CM3 and MIROC-5 (2060-2069) was used as inputs for RRI model to examine the variation of rainfall distribution over the CTLN river basin. Spatial distribution of 10-years average JJA rainfall clearly shows the future higher intensity of rainfall throughout the

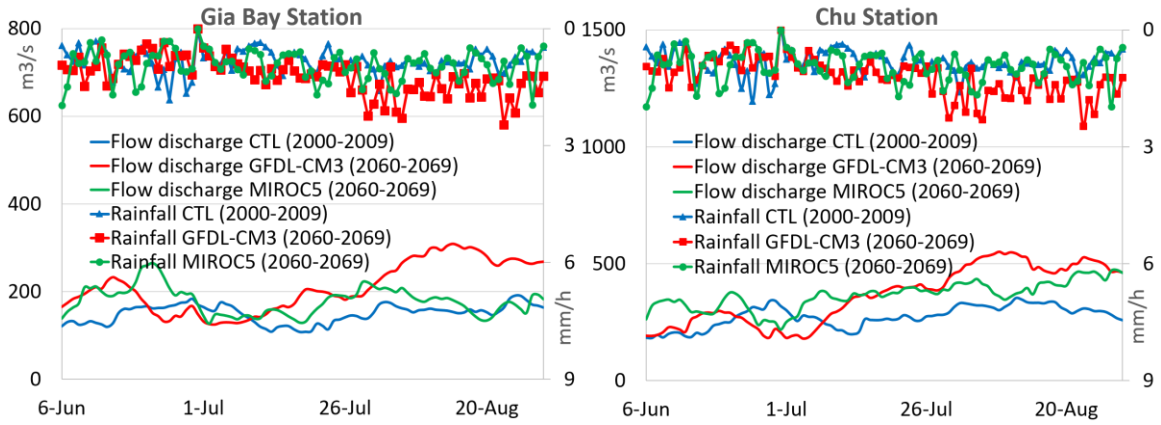

Figure 3.3 Daily precipitation and river discharge for present and future condition at Chu and Gia Bay stations

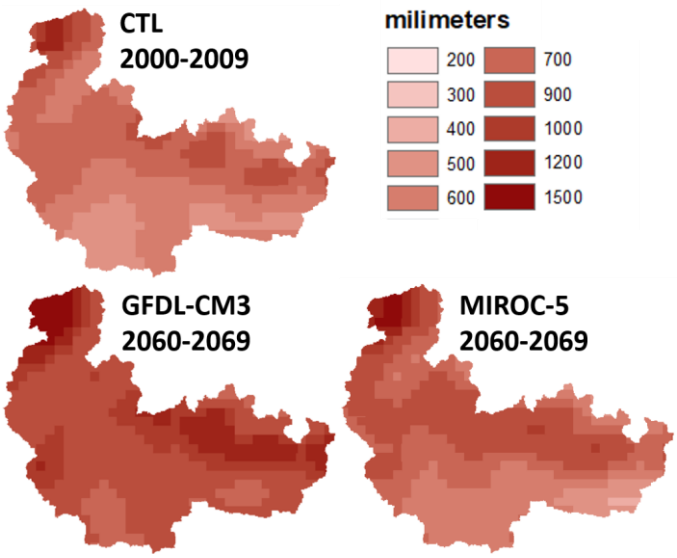

Figure 3.4 Spatial distribution of 10-years average JJA rainfall

watershed (Figure 3.4). Average JJA rainfall indicated by CTL, GFDL-CM3, and MIROC-5 were 607 mm, 872 mm, and 682 mm, respectively. While MIROC-5 predicted small rainfall variation (9% increment), there was a significant gap between GFDL-CM3 and CTL (44% increment). Rainfall forecasted by GFDL-CM3 was 28% greater than MIROC-5. Tran and Taniguchi (2014) used 19 CMIP5 models ensemble and found the same increasing trend of rainfall over the northern of Vietnam at the second half of 21st century. The same results also found by experts of MONRE (MONRE 2012). Future scenario forecasted by both GFDL-CM3 and MIROC-5 indicated the similar rainfall distribution patterns with present where the major amount of JJA rainfall concentrated on the upper river basin, especially the far north and eastern corner. The featured distribution of rainfall pattern in CTLN watershed results from its typical topographic condition.

3.3.2. Future trend of flood and river runoff

Future precipitation was predicted to increase during mid-21st century. Increasing river discharge is found in Figure 3.3. RRI simulated river runoff in CTL (2000-2009) were comparable with GFDL-CM3 and MIROC-5 (2060-2069) in the first period of the rainy season with no significant difference in average river discharge. Differences caused by higher rainfall intensity in the future scenario was found mostly in the latter half of the rainy season when future river discharge was projected larger than CTL. In the late JJA period, GFDL-CM3 model showed average river discharge will significantly increase, almost double as in CTL in both Chu and Gia Bay station. Meanwhile, MIROC-5 exhibited lower average river discharge than GFDL-CM3 as expected while future precipitation in GFDL-CM3 is projected much higher than in MIROC-5. Future river discharge in MIROC-5 is slightly larger than in CTL at during the peak runoff period and equal to CTL at the end of rainy season. In Gia Bay station which

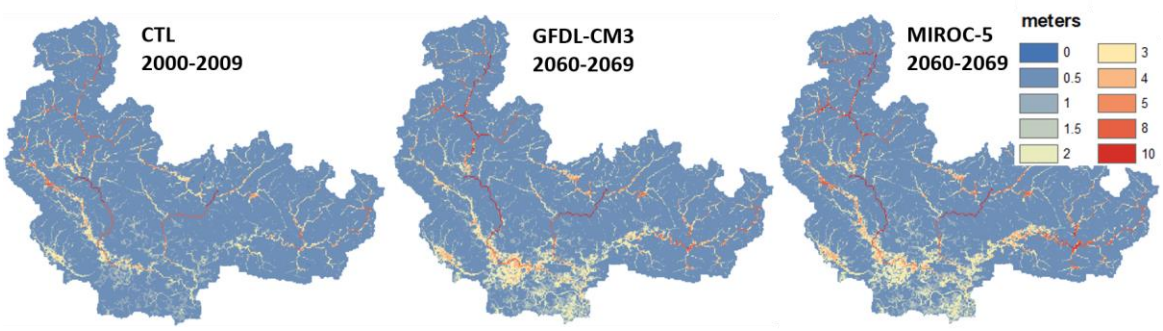

Figure 3.5 Maximum level of inundation depth in CTLN watershed during JJA

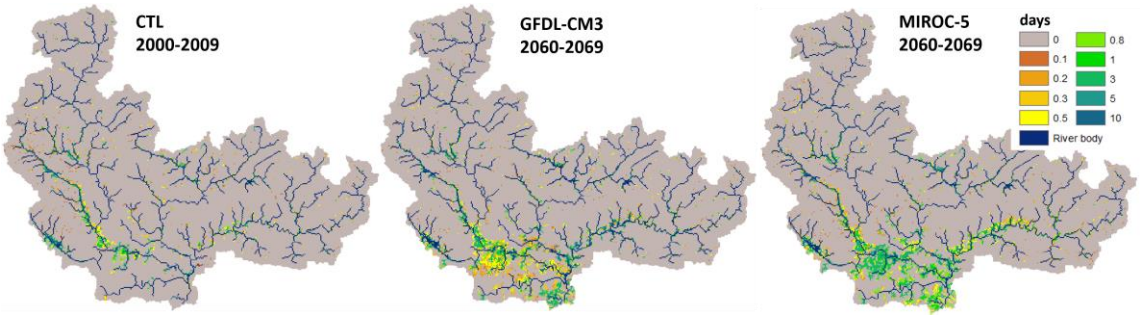

Figure 3.6 Average period of inundation depth of over 100 cm

locates in the largest branch of CTLN watershed, MIROC-5 show clear higher average

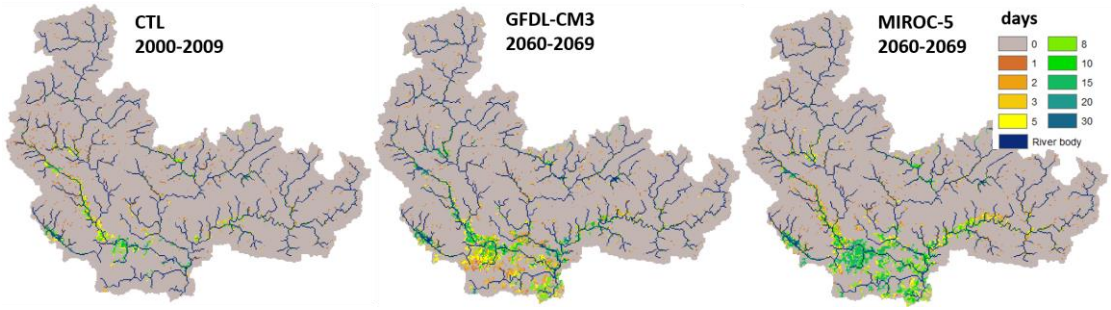

Figure 3.7 Maximum period of inundation depth of over 100 cm

discharge than in CTL. In Chu station, MIROC-5 exhibited higher discharge at late July-early August and similar to CTL at the end of JJA.

Figure 3.5 illustrates the maximum inundation depth by CTL and future projections. The distribution map for maximum depth was prepared by selecting the highest value of depth for every grid in the CTLN watershed during a 10-year period. Variation of maximum inundation depth represents the changes in the worse situation might happen. Both GFDL-CM3 and MIROC-5 show the huge difference in maximum inundation level than in CTL. The worst CTL situation indicates the inundation level of about 0.5 to 1.5 meters depth at the areas far from the river channel; closer to the river channel, the influence level becomes greater of 1.5 to 2 meters depth and up to 4-5 meters at some few locations (excluding river channel). Projection by GFDL-CM3 and MIROC-5 both indicate the more severe inundation in the lower river basin, about 0.5-1 meters higher than in CTL. Extreme flood (inundation depth > 3 meters) occur in more locations with extended inundation radius. Closer to main river channels at the downstream, the common inundation depth caused by the worst future flood situation in GFDL-CM3 and MIROC-5 are 2-3 meters, much higher than CTL. Compare between GFDL-CM3 and MIROC-5, the worst inundation situation in MIROC-5 was more intense for the most locations. Luc Nam river in the east branch watershed was unaffected by heavy inundation in

CTL but under severe flood in future projection. Increase in the mean JJA rainfall and river discharge is larger in GFDL-CM3 than in MIROC-5. However, the maximum inundation depth is larger in MIROC-5. These result suggests the amplitude of climate variability in MIROC-5 is stronger than in GFDL-CM3. Since the comparison between CTL and future river runoff only exhibits the clear differences on from late July/early August, it is projected that the severe future flood conditions tend to occur in the late JJA period.

3.3.3. Duration of extreme flood events

To obtain the overview of future flood condition in CTLN watershed, we further assessed the influences of the heavy flood by calculating the duration of inundation depth of over 100cm (IDO100) during JJA. The maximum duration of IDO100 were calculated for every grid in 10-year periods in present (2000-2009) and future condition (2060-2069).

Figure 3.6 compares the average duration of IDO100 between CTL, GFDL-CM3, and MIROC-5. According to CTL results, some areas adjacent to the left downstream river branch and some small areas in the right branch are frequently affected by the heavy flood with an average period of 0.5 – 0.8 days. The lower river basin region is quite safe without significant impacts by a long flood. In contrast, Both GFDL-CM3 and MIROC-5 projected more severe flood trend not only at near river channel but also in the lower river basin. The major part of the lower river basin will frequently be affected by the flood with larger flood radius and prolonged period. GFDL-CM3 projected the annual IDO100 will be longer, 0.3 to 0.8 days at the buffer zone, and up to 1 days at the places close to the river channel. In MIROC-5 models, the situation was projected even more severe when average duration of IDO100 at the lower river basin might be prolonged to 3 days.

The results of the maximum period of IDO100 for all cases shown in Figure 3.7 also indicated the same findings as in Figure 3.6 when projection from both GFDL-CM3 and MIROC-5 were worse than in CTL. The heavy flood might occur with IDO100 from 3-8 days in GFDL-CM3 and from 8-10 days in MIROC-5. Again, worst flood indicated in MIROC-5 is stronger and larger extent than in GFDL-CM3 by a significant gap. While precipitation in GFDL-CM3 is larger than in MIROC-5, both maximum inundation depth and duration of IDO100 in MIROC-5 are projected more severe in GFDL-CM3 suggested the harsh climate conditions in MIROC-5, especially at the end of JJA. The heavy rainfall at the end of JJA is believed the cause of this phenomena in MIROC-5.

Compared to CTL, future climate models projected higher flood impacts on the lower CTLN basin at mid-21st century. Since agriculture fields and residential areas are mostly placed at lower river basin (Figure 2.6), the result of this study indicated the increasing risks for the social-economic development of the region.

3.4. Chapter Summaries

This chapter has pointed out the significance of researching flood and inundation conditions over the CTLN watershed under global warming. Since the frequency and intensity of flood occurrence in CTLN watershed have been gradually increasing. There is an urgent need to establish the countermeasures for this key economic region based on the deep understanding of the hydro-meteorological characteristics of the watershed.

The reproducibility of WRF model for the CTLN regions was assessed by comparing the simulation results with observation daily rainfall from 2002 to 2009. Verification process indicated that the selected WRF setting can be applied with reasonable accuracy. Besides, the verification results for RRI model using observation rainfall and CLTN river discharge in 2009 have also shown the good reproducibility.

The future climate conditions for the region was given by the MIROC5 and GFDL-CM3 using RCP8.5 scenario. Simulation results indicated an increase of total rainfall over CTLN river basin. Consequently, the basin region was projected to be under the influence of more severe extreme flood events, especially in the lower river basin areas.

CHAPTER 4.
RESULTS AND DISCUSSION OF COUPLING DYNAMICAL AND STATISTICAL DOWNSCALING FOR HIGH-RESOLUTION RAINFALL FORECASTING: CASE STUDY OF THE RED RIVER DELTA, VIETNAM

4.1. Introduction

This chapter begins with the literature review and significance of the study on coupling dynamical and statistical downscaling in climatological research. It was followed by the detail verification and testing results of the combined WRF-ANN method. Firstly, dynamical downscaling by WRF was performed for weather information over the Red River Delta (RRD) and verified with observation data. Secondly, WRF outputs for 30km and 6km resolution were used as predictors and predictand inputs for the preliminary training stage with ANN. Thirdly, an independent dataset was used to further testing the applicability of the WRF-ANN for high-resolution rainfall downscaling. Finally, the comprehensive assessment for the method was finished by adding the sensitivity analysis for predictors and calculating the benefit of computational cost reduction.

4.2. Results of coupling dynamical and statistical downscaling for high-resolution rainfall

4.2.1. Dynamical Downscaling Experiment

This section evaluates WRF's ability to reproduce weather conditions in D2T. Table 4.2 shows evaluation statistics for the spatial distribution of mean rainfall from CTL for D1 and D2 as compared with observed rainfall at 38 rain gauges in 1996, 1997, 1998, and 2006, along with the statistical measures R and IOA. According to Table 4.2, the JJA rainfall in CTL was noticeably underestimated in both D1 and D2, as illustrated in the average accumulated values of 784 mm in D1 and 799 mm in D2, versus 1107 mm in the observation data. Furthermore, the rainfall projections at all observation locations were lower than the observed values. The summarized statistic indicated that simulated rainfall in D2 was slightly closer to the observation data than it was in D1. Both spatial correlation and IOA between observation data and CTL results exhibited slightly better values in D2 than in D1. This finding indicates that, in addition to the spatial resolution advantages, D2 can better resolve the finer resolution

Table 4.2 Statistical measures for WRF simulated rainfall over JJA periods

Year	Mean Obs (mm)	D1				D2			
		Mean CTL[a] (mm)	CTL/ Obs	R	IOA	Mean CTL (mm)	CTL/ Obs	R	IOA
1996	1078	679	0.63	0.77	0.65	701	0.65	0.78	0.66
1997	983	764	0.75	0.73	0.75	772	0.78	0.74	0.77
1998	1167	903	0.77	0.71	0.75	893	0.77	0.71	0.74
2006	1198	974	0.81	0.71	0.7	983	0.82	0.71	0.7
Average	1107	784	0.74	0.73	0.71	799	0.76	0.74	0.72

[a] Mean CTL was calculated by averaging the simulated rainfall from 38 locations in CTL corresponding to the locations of rain gauges

IOA: Index of agreement; Obs: Observation

R: Spatial Correlation (Pearson)

Table 4.1 Temporal correlation, RMSE, and MAE between CTL and Observation daily rainfall averaged for 38 locations for the JJA period in 1996, 1997, 1998, and 2006

	D2	RRD	NMR
Number of rain gauge	38	12	14
R	0.63	0.7	0.5
RMSE (mm)	14.53	9.12	17.98
MAE (mm)	4.67	3.52	7.21

than can the D1. In D2, the ratio of the precipitation average for D2T between CTL and the observed value was consistently lower than 1, ranging from 0.65 to 0.82, with an average of 0.72. Both spatial correlation and IOA between D2 and the observations were relatively high, at 0.66 to 0.77 and 0.71 to 0.78, respectively. This indicates reasonable accuracy for the CTL in reproducing the spatial distribution of JJA rainfall. Regarding the underestimated rainfall in the CTL output, several studies (Bukovsky and Karoly 2011; DeMott et al. 2007) have experienced the same problem. A common insufficiency of climate models is that they often underestimate extreme precipitation events, while overestimating the occurrence of light precipitation events. In this study, we focused on the rainfall season of a tropical country, where intense rainfall is a regular occurrence. Thus, it is not surprising to find that rainfall is underestimated in our simulation.

In addition to the spatial distribution of JJA accumulated rainfall, the temporal variations of daily rainfall between CTL and observations for D2 were examined at 38 locations

using the metrics of temporal R, MAE, and RMSE (Table 4.1). The CTL results for JJA during the research period suggest a moderate temporal agreement with observations, in which the average correlation coefficient for all 38 locations was 0.63. However, there was substantial variation in the correlation coefficients across the study site. While the temporal correlations for daily rainfall in midland and lowland areas, i.e., the Red River Delta, were high (average 0.7), the CTL output for the high mountainous regions was in low temporal agreement with the observations (average 0.5), especially in the areas between alternating high mountain ranges (as in the western part of D2T, see Figure 2.7 (c)). This finding highlights the limitation of WRF in resolving micro-climatological conditions and covering complex topography effectively. The same conclusion was arrived at by Li et al. (2016) in their research on the influence of topography on precipitation distribution.

The average MAE for the testing locations was 4.67 mm, while the average RMSE was significantly higher at 14.53 mm. Since the RMSE tends to amplify large biases, the large gap between the two values reflected the underestimation of heavy and extreme rainfall cases in the CTL, which partly resulted in underestimation of the accumulated rainfall mentioned above. The average correlation coefficient for all 38 locations indicates a moderate agreement, but there was large variation in correlation coefficients between the locations (detail not shown). JJA accounts for over 70% of the annual rainfall, which begins in June, peaks in late July, and decreases through August. The correlation coefficient results indicate that the proposed WRF setup can reproduce seasonal variation in rainfall relatively well, especially for the lowland region where D2T is located.

To examine the significance of the calibration method for DDE (omitting values less than 0.5 mm per day), we directly compared DDE during the JJA period from the CTL results, and CTL calibrated results, and observed values for each observation location. The maximum, minimum, and average DDE percentages in D2T at the 38 locations are presented in Table 4.3. The summarized results clearly show miscalculation of DDE by CTL. The percentage of DDE in the CTL results ranged from 4.5% to 8.6% of the total grid cells in D2, which were 4 to 6 times lower than the actual data across all years. The application of a wet-day-threshold showed a good result in eliminating the biases between simulated and observed data, with a large improvement in the calibrated CTL results. Even when the ranges of maximum and minimum percentage DDE did not perfectly match the observations, the average DDE results for JJA in the calibrated CTL were very close to the observed average. However, quantitative assessment of the reduction in total rainfall owing to DDE calibration indicates that the mean total rainfall

Table 4.3 Comparison of the percentage of DDE in JJA among 38 rain gauge locations

Year	Observation (%)			CTL (%)			Calibrated CTL[d] (%)		
	Max[a]	Min[b]	Average[c]	Max	Min	Average	Max	Min	Average
1996	51.2	18	32.5	8.6	4.5	5.6	43.1	17.5	26.49
1997	56.5	18.5	28.6	7.3	4.5	5.3	38.7	12.5	26.53
1998	56.4	18.6	31.4	6.5	4.8	5.1	45.1	16.3	30.27
2006	52.1	16.2	32.5	6.6	4.5	6.3	45.9	10.4	28.18

[a, b, c] Max, Min, and Average correspond to the maximum, the minimum, and the averages of DDE cases during JJA among 38 rain gauge locations, respectively.

[d] Calibrated CTL: the CTL daily rainfall was calibrated with wet-day-threshold

decrease is 0.28% for D1 (detailed result not shown here). Thus, the calibration helps WRF better capture DDEs, with a negligible effect on total rainfall. WRF simulation results for D2T after calibration were expected to be a good predictand for the ANN training stage.

4.2.2. Results of the ANN preliminary training stage

This section describes the training stage results of rainfall downscaling, using the MLP-based ANN on different model configurations as mentioned in Table 2.5. To compare the predicted model outputs with the desired output, various statistical measurements were adopted. Their results are presented in Table 4.4, while regression plots for the testing set are shown in Figure 4.1. Results of the training stage show substantial variations in performance among the models. The training results improved with regard to higher model complexity; however, there was clear consistency in network performance, with most models exhibiting similar correlation coefficients in the training and test sets. The cross-validation method was proven to be effective in detecting the best generalization point and stopping the training process before the model shifted to over-learning.

The simplest designs in the M1 model series provided the worst results (very low accuracy and large RMSE) and were unable to forecast DDE cases. Figure 4.1 indicates that the M1n, M1a, and M1s models heavily underestimated RD2T, which might be attributed to the low predictive power of the combination of variables sea level pressure (slp), temperature at 2m (t2), geographical height (hgt), and planetary boundary layer height (pblh). However, the R^2 coefficient indexes suggest that M1n was a better model than M1a and M1s. While the M1n

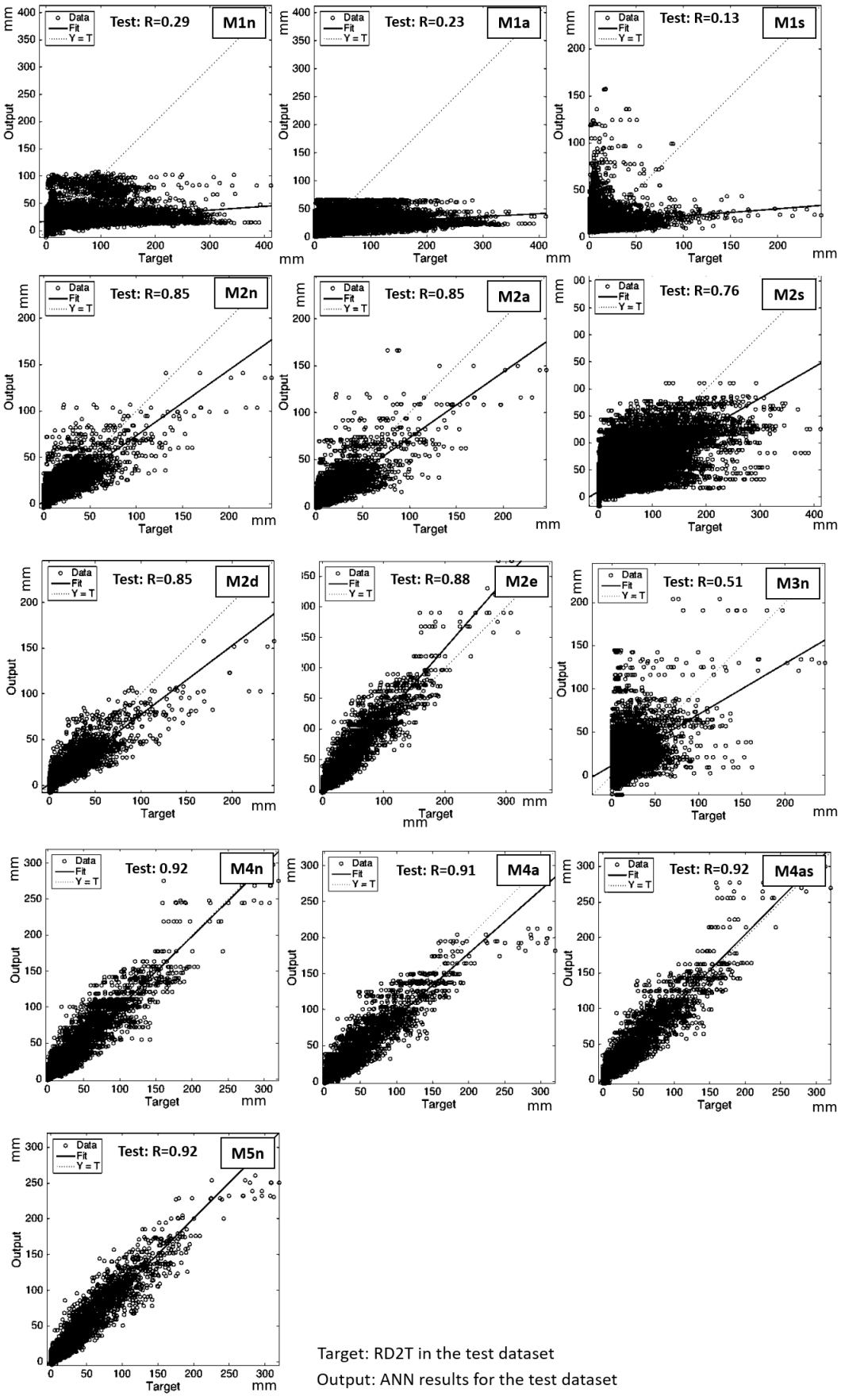

Figure 4.1 Correlation coefficients for the ANN model test sets

Table 4.4 RMSE and R^2 for training and test sets of different ANN model configurations

	M1n	M1a	M1s	M2n	M2a	M2s	M2d	M2e	M3n	M4n	M4a	M4as	M5n	RD2T
R^2 of Training set	0.32	0.25	0.19	0.91	0.92	0.85	0.91	0.91	0.62	0.92	0.92	0.93	0.92	
R^2 of Test set	0.29	0.23	0.13	0.85	0.85	0.76	0.85	0.88	0.51	0.92	0.91	0.92	0.92	
RMSE for all dataset (mm/day)	32.53	35.53	35.94	13.24	12.95	18.49	14.52	12.93	28.47	9.44	8.67	8.94	10.42	
DDE in full-dataset (%)	0.01	0.01	0.04	10.34	3.24	21.94	11.14	11.84	2.32	14.53	11.59	20.53	27.43	27.74

setting used the NV input features, which might be better than AV and SV, the accuracy levels of the M1 model series were too low to determine any differences. The M2 model series showed significantly better fitting and bias results compared to the M1 series. Except for M2s, with R^2 for the test data of 0.76, the other M2 models yielded at least 0.85 for R^2 in the test set. The large improvement in the M2 series was achieved by incorporating the rainfall in D1 (RD1), which is highly correlated to RD2T (Table 2.4). The sudden drop in the predictive power of the M3n model also indicates the significance of RD1 in model design, since M3n eliminated the RD1 variable in the training stage. With the same settings, the M2n and M2a models showed significantly better correlation to RD2T than did M2s, thus indicating a lack of signal strength in SV features. Although M2s appears to be weaker, its simulated DDE percentage was 21.94%, which was close to that of RD2T (27.74%). The ability of M2s to resolve DDE was significantly better than those of M2n and M2a, whose DDE percentages were 10.34% and 3.24%, respectively. The reduction of DDE percentage in M2a reflected the drawback of the AV features, compared to the NV features, since the average value reduces the variation signal in the predictors. On the other hand, M2d and M2e exhibited slightly better results than did M2n, especially with regard to the reproducibility in DDE cases (11.14% and 11.84%). The large number of neurons and the high rescaling range—[-1:1] of HTS to [0:1] of LS—made it more convenient for the network to detect very small rainfall values.

The M4 series and M5n were clearly better than the other models, as their predictions were close to that of RD2T, yielding significantly lower RMSE. Adding more predictor variables have proven to be helpful in increasing model accuracy. There was variation in the behaviors between the models in the M4 series, despite their skillful results. Even though the RMSE of M4a was lowest at 8.67 mm/day, its ability to map small rainfall signals was significantly lower than those of M4n and M4as. M4a indicated 11.59% DDE, while M4n and M4as indicated 14.53% and 20.53%, respectively. Compared to the DDE percentage of 27.74% in RD2T, M4as was clearly the best at forecasting small rainfall values. The study results indicate that a combination of AV and SV features might be better in detecting DDE. The higher RMSE values in M4n and M2n, as compared to the other M2 models, might be attributed to the interaction of highly-correlated inputs among the NV features. The same behavior was also demonstrated by Wendemuth et al. (1993), who found that the combination of correlated inputs potentially adds more weight not only to the predictive information, but also to the biases. M5n discarded rainfall event information during the training stage, which understandably resulted in higher RMSE than the M4 model series. The RE-ANN calibration method was

employed for the M5 model to successfully map a total of 27.43% DDE grid cells, which was smaller than that of RD2T by a small margin.

4.2.3. Results of WRF-ANN downscaling of an independent dataset

The second testing stage in this study aimed to further assess the applicability of coupling ANN to WRF output for high-resolution rainfall downscaling and compare it with the interpolated data using a bilinear interpolation method. We selected the ANN architectures that demonstrated the most promising global approximation abilities during the training stage (the M4 model series and M5n) to apply for an independent dataset for the year 2006. In addition, we also used bilinear interpolation to downscale RD1T from 30 km to 6 km (denoted BIP-

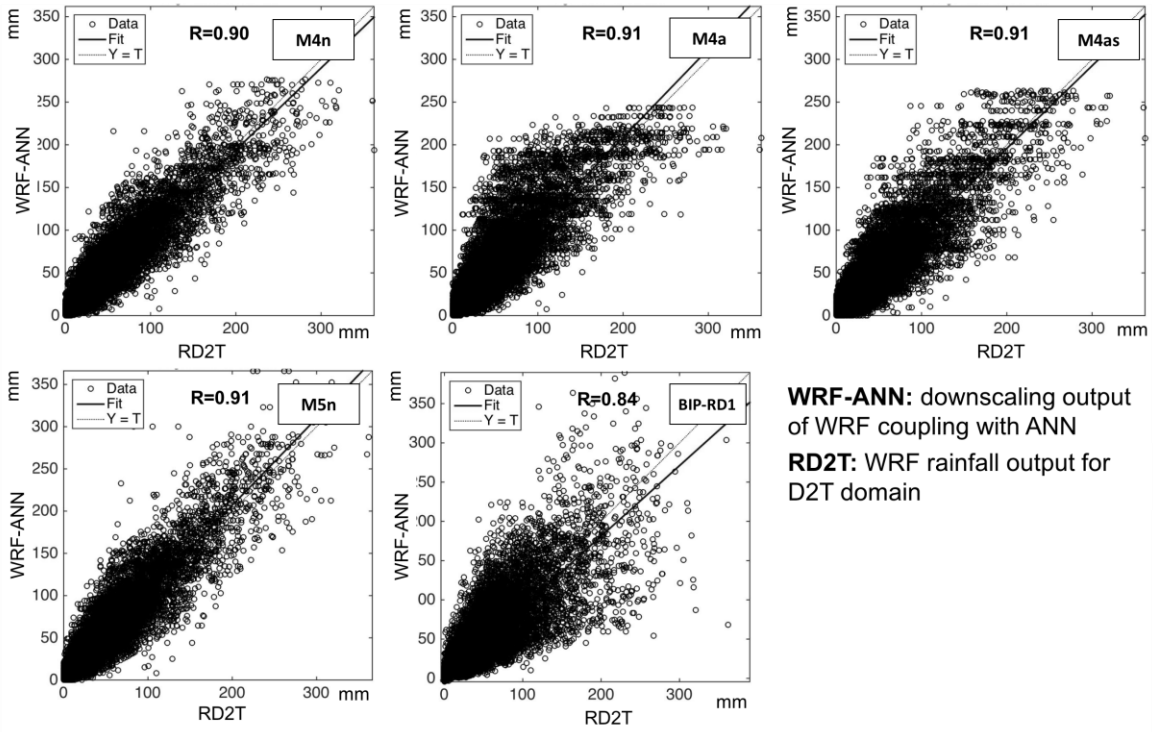

Figure 4.2 Regression plots for target and forecasted rainfall in 2006

Table 4.5 The second stage testing results of ANN models

	M4n	M4a	M4as	M5n	BIP-RD1	2006-RD2T[a)]
Spatial correlation for 2006-dataset	0.9	0.91	0.91	0.91	0.84	
RMSE for 2006-dataset (mm/day)	14.21	11.27	10.24	12.42	16.93	
DDE in 2006-dataset (%)	9.54	15.94	19.48	23.84	17.43	23.78

[a)]2006-RD2T: RD2T of the 2006 dataset

RD1). The summarized results of the tests are presented in Table 4.5, while regression plots for the target and forecast rainfall are plotted in Figure 4.2. All models in the second stage continued to show predictive consistency in performance with the 2006 dataset. Differences in correlation coefficient metrics were observed, although they were insignificant. The correlation coefficients (> .9) for simulation outputs in this stage were comparable to results from the preliminary stage, indicating good model reproducibility. However, in the 2006 dataset, the simulation results also exhibited more prediction errors, as can be seen in the RMSE. The unexpected reduction in model stability may be due to imperfect model design or a lack of representative information in the training dataset (Sánchez Lasheras et al. 2010). Sometimes, the incomplete nature of model development may also contribute to the problem (Tu 1996). The models that adopted NV variables, including M4n and M5n, were observed to have higher biases than those that adopted AV and SV variables, including M4a and M4as. Highly correlated NV inputs seemed to yield more error than their generalized features. Both M4a and M4as proved better than M4n at predicting the DDE percentage, with 15.94% and 19.48%, respectively, compared to 9.54%. Between the two, the M4as model, which inherited the predictive power of both SV and AV features, outperformed M4a in every measure. However, M5n is the model that delivered the best forecast of DDE percentage, at 23.84%, which was within 0.1% of the 2006-RD2T of 23.78%. Since M5n was designed with the same setting as M4n, the RE-ANN calibration method was proven effective in locating DDE cases. Results of the bilinear interpolation method, BIP-RD1, on the other hand, showed noticeably lower spatial

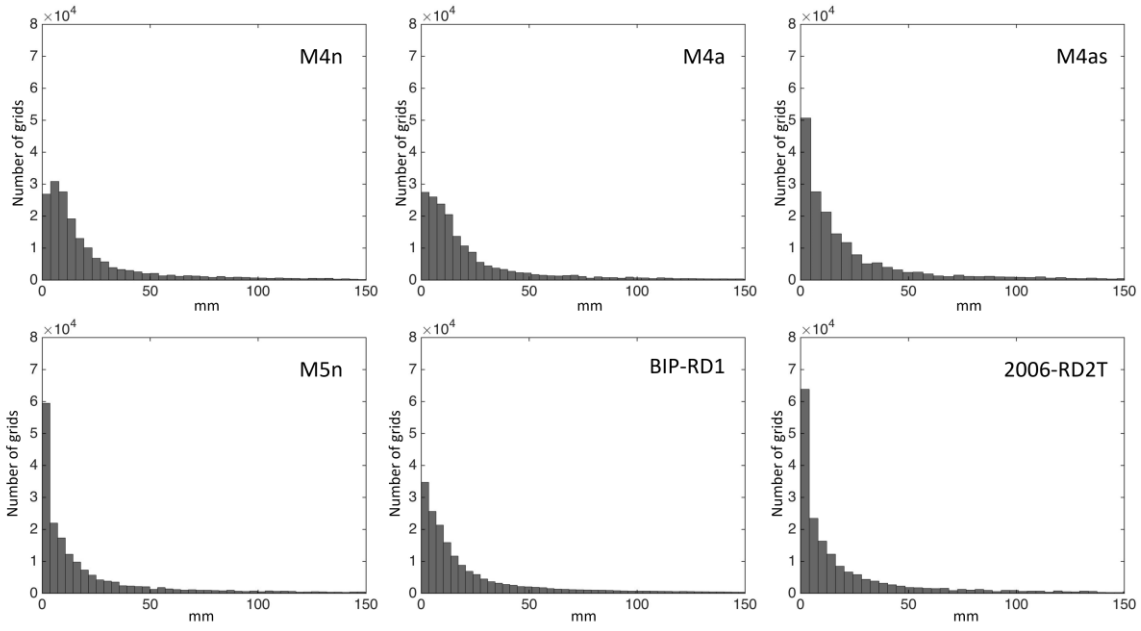

Figure 4.3 Histogram plot of JJA rainfall (mm) in 2006

correlation coefficients and higher RMSE values than did the ANN downscaling. DDE percentage determined by BIP-RD1 was 17%, much lower than the observed value of 24% in 2006-RD2T. The bilinear interpolation method generates estimated values between grid points. It is a simple and fast method, but lacks important embedded dynamical processes that are contained in the WRF models. The ANN method, on the other hand, performs downscaling by creating statistical relationships between high-and intermediate-resolution WRF outputs. ANN incorporates the dynamical processes given by WRF during the training processes. This added value provided by ANN helped to capture fine-scale variations in the downscaling results. It is therefore reasonable to find that downscaling with ANN outperformed the bilinear interpolation method.

Comparisons between the ANN models and target data with regard to the frequency of dry days, wet days, and extreme rainfall events is shown in Figure 4.3. The rainfall frequency illustrated by all models was similar to that of RD2T, wherein the dry day and low rainfall (less than 20 mm) cases accounted for most of the days during JJA. Regarding the distribution of

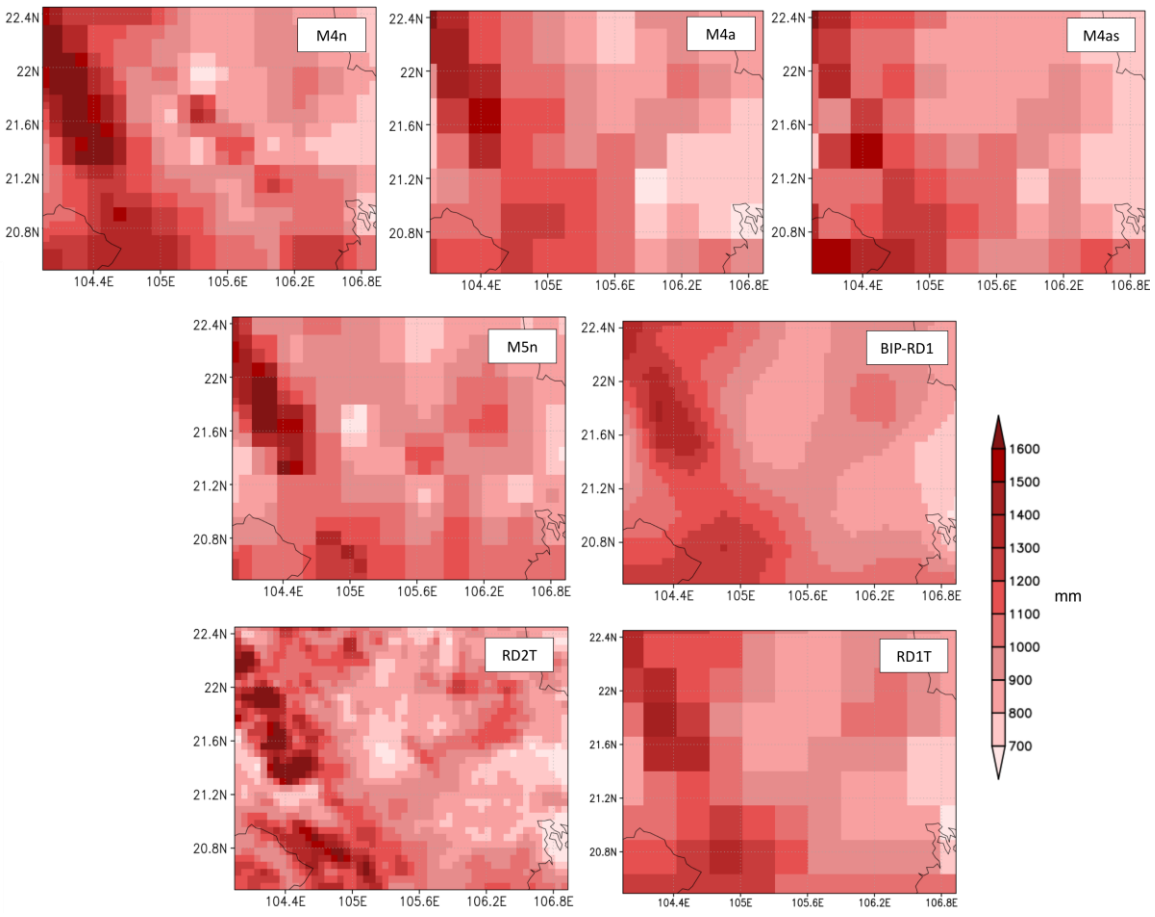

Figure 4.4 Spatial distribution of cumulative rainfall (mm) in JJA of 2006

very low rainfall cases (less than 5 mm) and extreme rainfall cases (higher than 100 mm), the M4n and M4a models showed weakness in their underestimation of low rainfall cases. These two models failed not only in resolving the DDE cases as illustrated in, but also in projecting small rainfall values. BIP-RD1 exhibited better DDE percentages than did the M4n and M4a models, but these were still much lower than the observed values. Meanwhile, the M4as and M5n were nearly identical, with rainfall frequencies in these two models being similar to the observed values. Both M4as and M5n showed significant improvement over M4n and M4a in locating small rainfall ranges.

The distribution maps for cumulative JJA rainfall in 2006 by M4n, M4a, M4as, M5n, and BIP-RD1 are depicted in Figure 4.4, in comparison with RD2T (2006-RD2T) and RD1 (2006-RD1). Owing to the high correlation with 2006-RD2T, ANN downscaled the rainfall in all models, clearly demonstrating a good pattern-correlation. The highest rainfall areas were accurately located in the southwestern corner of D2T, and rainfall gradually decreased towards the northeast. While the spatial correlations of cumulative rainfall were similar among the models, the rainfall distribution results indicate an absolute strength of NV input features over AV and SV features, as pertains to the downscaled detail. We can explicitly recognize the smoother transition of rainfall withdrawal from higher to lower rainfall areas in the M4n and M5n models than is demonstrated in the M4a and M4as models. Compared to the 2006-RD2T distribution pattern, the rainfall transition patterns in M4n and M5n showed a loss in detail; even so, its resolution was sufficiently high to distinguish minor changes. The essence of the WRF-ANN downscaling method was the use of four D1 grid cells to predict one spatially-overlapped grid cell in D2. When the resolution of D2 was too high for comparison with D1, it was unavoidable that some adjacent cells in D2 would have the same predictor values. This problem results in predicted values repeating for some cells. Less detail was expected in M4n and M5n than in 2006-RD2T, since increasing resolution from 30 km to 6 km is a large jump. As expected, both the M4n and M5n models showed significantly higher resolution than that of the 2006-RD1. In contrast, M4a and M4as had significantly lower resolution and coarse rainfall patterns. The differences between M4a and M4as were too small to indicate any advantages from combining both AV and SV features for prediction. Even with their higher resolution, neither M4a nor M4as demonstrated better changes in the minor rainfall pattern. In this test, simulation results suggest that generalized features might be more effective in bias control. However, this approach loses essential information for examining the spatial distribution of precipitation, which leads to similar generalized results. On the other hand,

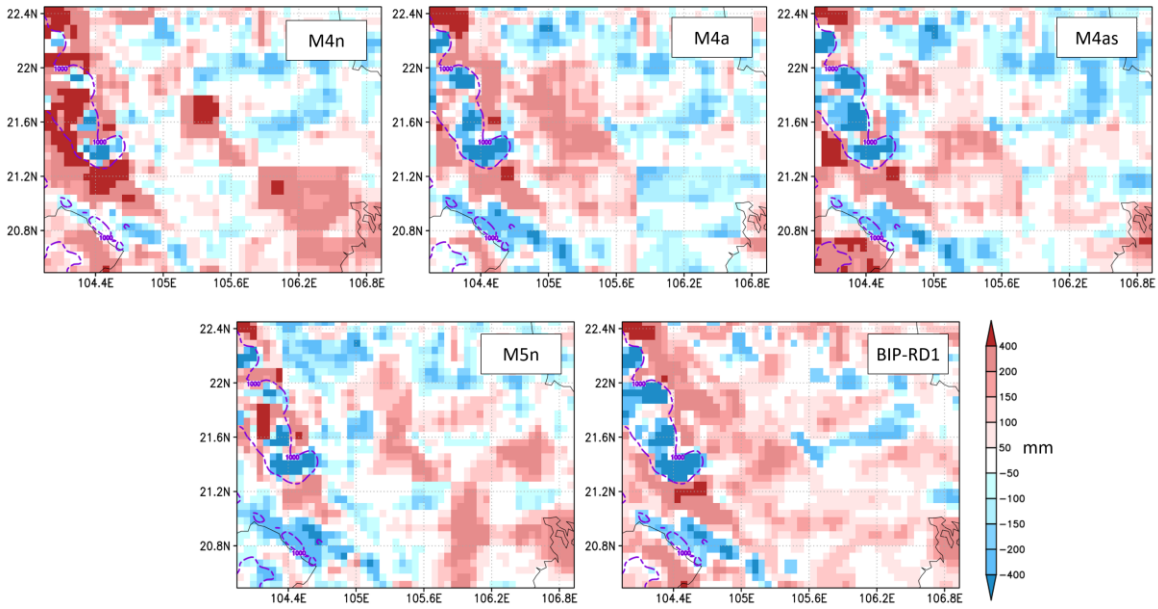

Figure 4.5 Differences between simulations in cumulative rainfall (mm) in JJA of 2006 results and RD2T. The purple contour dash lines indicate the areas with terrain height of over 1.000m

while the original NV predictor exhibited a larger bias, it better mapped the variability in rainfall. BIP-RD1 showed larger increase in spatial resolution than did RD1T, but failed to generate the wide range of rainfall variation present in RD2T. It tended to overestimate rainfall in light rainfall grid cells and underestimate rainfall in heavy rainfall grid cells.

The differences between cumulative JJA rainfall simulated by M4n, M4a, M4as, M5n, and BIP-RD1with RD2T are indicated in Figure 4.5. All models exhibited larger estimation errors in the northwestern part of D2T, especially in the high terrain and surrounding area. However, these large errors were not a surprise because this area accounts for the highest JJA rainfall (Figure 4.4). The BIP-RD1 model showed slightly larger error than the other models, while both M5n and M4n overestimated the total JJA cumulative rainfall, with M4n having the larger overestimation, as reflected in its RMSE. Since M5n neglected very small rainfall values during calibration, it potentially avoided bias intensification by small rainfall values during training. Moreover, in a comparative study on software estimation efforts, Nassif et al. (2012) also found an overestimation tendency by MLP-ANN, especially for an MLP trained with a complicated range of inputs. The model behavior suggests that small rainfall values, which accounted for 10% to 40% of the dataset, were difficult to reproduce by ANN. However, they can be addressed using the RE-ANN calibration methods.

Table 4.6 Performance statistics for ANN sensitivity analysis for 1996-1998 dataset

	M5n	W/o Rd1	W/o tk	W/o hgt	W/o slp	W/o grdflx	W/o psfc	W/o pblh	W/o q2
R^2 of Training set	0.92	0.48	0.88	0.92	0.88	0.9	0.91	0.91	0.86
R^2 of Test set	0.92	0.47	0.87	0.91	0.86	0.89	0.89	0.88	0.85
RMSE for all dataset (mm/day)	10.42	24.52	15.29	10.82	15.24	13.48	11.37	10.98	17.43

W/o : Without

4.2.4. Predictor sensitivity analysis

To obtain a comprehensive view on the applicability of coupling WRF and ANN to downscaling, the influence of each input predictor on the output should be investigated. Normally, variables with a higher correlation to the predictand are expected to be more helpful in forecasting. However, an unusual combination of correlated or uncorrelated variables might also be useful. In this study, we considered 17 variables (Table 2.4), gradually fine-tuning their combination through the trial and error method. Although we could not cover all possible combinations, our best effort so far—as used in the M5n model—demonstrated promising results. Sensitivity analysis was conducted for each variable input for the M5n model to examine their significance to the ANN outputs. The sensitivity analysis method used in this study was introduced by Hung et al. (2009), in which each input parameter in the M5n model was alternately removed from the ensemble, subsequently comparing the performance statistics with the original. Since the M5n model utilized eight variables, including RD1, atmospheric temperature at 1400 mm (tk), hgt, slp, ground heat flux (grdflx), surface pressure (psfc), pblh, and humidity at 2 m (q2), there were eight models included in the sensitivity test. The results of the sensitivity test are presented in Table 4.6.

As can be seen in Table 4.6, RD1 has the largest impact on the predictand. Excluding RD1 substantially reduced network performance. Meanwhile, the model indicated the second largest impact by q2, while the third and fourth most important parameters were tk and slp, whose results were very similar to each other. Among the remaining variables, grdflx differed from the others by a higher RMSE, achieving the fifth position. The remaining variables pblh, psfc, and hgt were the least important, since the models trained without them were comparable to the original M5n model.

Apart from input variables, it is also important to consider sensitivity to the treatment methods used for the input variables, which classify the variables into NV, AV, and SV features.

While NV features tend to yield more error, they can resolve the spatial variability of rainfall. Meanwhile, generalized features such as AV and SV can better control the bias in prediction values, but have lower effective resolution. NV features were concluded to be the best fit for making WRF-ANN models.

4.2.5. Computational Cost

The expected results when adopting WRF-ANN over WRF include a comparable downscaling quality with reduced computational load and time. Since the step of downscaling from 30-km to 6-km resolution using ANN gives results instantly, the advantage of using WRF-ANN methods was measured by comparing the time consumption needed by WRF to downscale rainfall to 30 km or to 6 km. Our measured results indicate that WRF downscaling to 6-km resolution took 9.3 times longer than downscaling to 30-km resolution. Rainfall downscaling using the WRF-ANN method can therefore save up to approximately 89% of the computational cost, as compared to downscaling using WRF alone.

4.3. Chapter Summaries

This chapter has indicated the advantages and disadvantages of both dynamical and statistical downscaling. Since dynamical downscaling is excessively required computational power and time consuming, it has incorporate the physical processes of the global which generally can better simulate the weather conditions of complicated different conditions and regions. On the other hand, statistical downscaling works based on the empirical statistical relationship between large scale predictors and small scale predictand which can give results instantly but limited in accuracy. The combination of both methods is expected to be the reliable and fast way for climate researchers. This proposed method, however, is quite new without many studies published so far. This dissertation has partly contributed for the improvement of the new approach on downscaling weather information.

The WRF-ANN method has taken into account both WRF and ANN models. WRF downscaling was applied for the Red River Delta in Vietnam for 30km and 6km resolution for mother and child domain, respectively. The configuration setting for WRF used in this study showed the reasonable accuracy in reproducing historical climate data in the period 1996-1998 and 2006.

WRF outputs for 1996-1998 was used as inputs for ANN preliminary training stage where climate variables from 30km resolution domain was taken as predictors and rainfall from 6km resolution was taken as predictand. Various ANN settings were tested using the traditional trials and errors approach during the training stage. The RE-ANN calibration method was developed to increase the performance of the model. The M5n model was be best model tested during the preliminary stage that was a MLP-BG network of three hidden layers using the hyperbolic tangent sigmoid activation function.

WRF-ANN results of the preliminary training was used for an independent dataset of 2006. Application of WRF-ANN method for the 2006 produced high-resolution rainfall patterns that are highly correlated with WRF (r = 0.91) and have low RMSE (12 mm/day). Despite of the loss in spatial resolution, the WRF-ANN downscaling results have showed the significant improvement compare to the WRF-30km resolution data and had the good agreement with WRF-6km resolution data.

The predictor sensitivity analysis and computational cost sections have illustrated the significant of each variables and the benefit of WRF-ANN method.

CHAPTER 5.
CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

5.1.1. Research problems and overview of the thesis aims

Since Vietnam is considered among the five countries heavily influenced climate change. This projection has been authenticated through the increasing trend of climate related disaster in Vietnam, especially the northern region. The number of flood and inundation occurrence in Cau-Thuong-Luc Nam river basin has been increased considerably since the beginning of the 21st century. For the better support the decision maker to prepare the response plan for the coming condition. We proposed our first aim of this study is to researching the inundation and flood conditions of the Cau-Thuong-Luc Nam river basin under climate change.

Decision makers need increasingly demands for local scale climate information to address the risk posed by projected climate changes and their related impacts. To meet the desire to derive climate projections at scales that can support decision makers, a process termed downscaling has been developed. Downscaling consists of a variety of methods, each with their own merits and limitations. Since dynamical downscaling is reliable but require significantly amount of computational power and time consuming, statistical downscaling is generally very fast but having a lot of constraint on downscaling results. The blended methods that use both dynamical downscaling and statistical downscaling is quite new without many researches have done so far. The second aim of this study is to investigate the ability to couple dynamical and statistical downscaling for high-resolution rainfall forecasting.

In general, the objectives of the study have been achieved: for the first target of the study, inundation and flood conditions over the Cau-Thuong-Luc Nam river basin under global warming have been researched; the coupling dynamical downscaling with statistical downscaling for high-resolution rainfall forecasting - case study of the Red River Delta in Vietnam was also successfully developed.

- Firstly, researching inundation and flood conditions of Cau – Thuong – Luc Nam river basin under global warming was conducted. The WRF model provide the high-resolution future rainfall data that can be put into RRI model to investigate the coming hydrological conditions.

- Secondly, for the purpose of researching the ability to couple dynamical downscaling and statistical downscaling, the WRF and ANN models were used. We selected the Red River Delta as the domain for the research. The essential of the coupling method is to use the intermediate downscaling output from WRF to feed into ANN for high-resolution downscaling.

5.1.2. Key findings

(a) Coupling WRF and ANN for high-resolution rainfall forecasting

The possibility of coupling WRF and ANN for high-resolution rainfall downscaling was investigated with a case study from the Red River Delta in Vietnam. The evaluation shows that the WRF modeling system can reproduce temporal variation in the JJA daily rainfall reasonably well, but underestimates the total precipitation. Owing to the higher precision of WRF, the region appears to have more drizzle, resulting in significantly fewer dry days than were observed. However, by implementing a wet-day threshold of 0.5 mm, we were able to correct this issue.

The best performing ANN model, M5n, produced high-resolution rainfall patterns that are highly correlated with WRF (r = 0.91) and have low RMSE (12 mm/day). High-resolution rainfall in each grid cell was downscaled by taking the climatological variables from the four grid cells in the coarse domain. The M5n model was configured as an MLP-BG network with three hidden layers using the hyperbolic tangent sigmoid activation function. The optimal predictors for M5n were rainfall in D1 (RD1), atmospheric temperature at 1400 mm (tk), geographical height (hgt), sea level pressure (slp), ground heat flux (grdflx), surface pressure (psfc), planetary boundary layer height (pblh), and humidity at 2 m (q2). In addition to having high accuracy, applying WRF-ANN is also expected to reduce computational costs. Running 30-km WRF and using ANN to downscale to 6 km is 89% less expensive than running nested 30-km and 6-km WRF simulations. We developed a calibration method (RE-ANN) to help ANN better capture dry days. This method treats a grid cell in D2T as dry if it was touching a dry grid cell in D1. This improved our simulation of dry days with ANN. The network trained for RE events and calibrated with the RE-ANN calibration method delivered the best prediction for our study area and period. Statistical relationships created by ANN can be used to directly downscale climate information from 30-km WRF output to a 6-km grid with reasonable accuracy. The application of ANN with WRF was effective for rapidly downscaling daily-basic rainfall data in a season at low computational cost.

(b) Inundation and flood in Cau-Thuong-Luc Nam river basin under climate change

In this study, the present and future rain-fall-runoff and inundation conditions of CTLN river basin during JJA were also examined using the combination of WRF and RRI models. While simulation results of WRF for the present precipitation showed reasonable accuracy for temporal and spatial distribution, reproductive RRI results for CTLN river discharge came relatively close to OBS. Generally, both WRF and RRI models are capable of deploying further assessment on the future river basin condition.

The future downscaling results by GFDM-CM3 and MIROC-5, indicated heavier rainfall conditions in the mid-21st century and consequently cause more severe inundation conditions. At the first half of JJA, there is no significant difference in average river runoff conditions between CTL and future projections, heavy rainfall and inundation were expected to increase in the second half of JJA. In both GFDL-CM3 and MIROC-5, the impacted areas due to flood will increase in both spatial and temporal extent, intensity, and density. MIROC-5 model forecasted the extreme flood might occur in late JJA. Future inundation condition will affect mostly the agricultural and residential areas in the lower CTLN river basin.

5.2. Limitations of the study

Inundation and flood in Cau-Thuon-Luc Nam river basin under climate change

The current outcome of the study have not yet determine the influence level of inundation and flood for the specific aspects of socio-economic development. Future climate condition given by only MIROC5 and GFDL-CM3 models might be the good representative inputs for the research of global warming. However, the results from two models cannot be considered as the reliable sources for any conclusion why there is always a significant variation between global climate models. A set of models or an ensemble of models is the alternative approach has been widely use to give the more consistent forecasting results although this method is very demanding. In this study, using RCP8.5 scenario – the high emission storyline, is suggested by the Vietnam ministry of Natural Resources and Environment. Comparison between RCP4.5 and RCP8.5 results is still needed to have the complete overview of the problems.

Coupling WRF and ANN for high-resolution rainfall forecasting

The target of this dissertation to research the possibility of coupling dynamical downscaling with statistical downscaling has been conducted successfully when the coupling WRF-ANN method can give comparable results to downscaling by WRF with 89% less time consuming. However, the spatial distribution of rainfall output by WRF-ANN is still lower than the output by WRF due to the big jump in resolution increasing. This has reflected the limitation of the current method to apply to more ambitious research works. Hence, to increase the applicability of the method, the current algorithm of selecting variables for ANN training must be revised. On the other hand, the outcome of this research so far has not determined the accuracy of WRF-ANN downscaling method for high and extreme rainfall events. Such objectives are necessary to be included to enhance the credibility of the method.

5.3. Future research and some recommendation

Inundation and flood in Cau-Thuon-Luc Nam river basin under climate change

This study suggests further assessments on the impacts of the future flood to agriculture and environment as well as the needs of study on an adaptive management plan. For the reliable forecasting results, it is necessary to include a model ensemble in the assessment since model ensemble can minimize the bias caused by single model. Detailed research on the frequency and intensity of rainfall is also important to accurately assess the impact of extreme weather events.

Coupling WRF and ANN for high-resolution rainfall forecasting

To further improve predictive skill of the WRF-ANN model, an additional analysis of the model biases will be required, e.g., sources of overestimated cumulative rainfall during JJA. Such analysis will require more detailed and extensive comparison of the various model configurations and predictor combinations in ANN. Using the coupling methods, we plan to extend the applicability of WRF-ANN to an ensemble of climate models, in which the principal components of the model ensemble can be considered as inputs for ANN downscaling. This approach will potentially help facilitate the use of ensemble model prediction, without the need for excessive time and computational power. Additionally, we plan to experiment with even higher resolution (finer than 6 km) downscaling using WRF-ANN.

APENDIXES

Appendix 1: *Weather stations used in this dissertation*

ID	Station name	Commune	District	Province	Lon	Lat	Time in Operation
1.	Điện Biên	Bản Bôn Na	Mường Thanh	Lai Châu	10300	2122	01/06/58
2.	Lai Châu	Thị Xã	Lai Châu	Lai Châu	10309	2204	01/04/56
3.	Mường Tè	Bản Nậm Củm	Mườngtè	Lai Châu	10250	2222	01/02/61
4.	Pha Đin	Đỉnh Pha Đin	Tuần Giáo	Lai Châu	10331	2134	01/01/64
5.	Sìn Hồ	Thị Trấn	Sìn Hồ	Lai Châu	10314	2222	01/03/61
6.	Tam Đường	Bản Giáng	Phong Thổ	Lai Châu	10329	2225	01/05/73
7.	Tuần Giáo	Quai Long	Tuần Giáo	Lai Châu	10325	2135	01/01/61
8.	Bắc Yên	Thị Trấn	Bắc Yên	Sơn La	10425	2115	01/06/73
9.	Cò Nòi	Cò Nòi	Mai Sơn	Sơn La	10409	2108	01/01/64
10.	Mộc Châu	Thị Trấn	Mộc Châu	Sơn La	10441	2050	01/06/61
11.	Phù Yên	Thị Trấn	Phù Yên	Sơn La	10438	2116	01/01/61
12.	Quỳnh Nhai	Thị Trấn	Quỳnh Nhai	Sơn La	10334	2151	01/02/61
13.	Sơn La	Thị Xã	Sơn La	Sơn La	10354	2120	01/12/60
14.	Sông Mã	Thị Trấn	Sông Mã	Sơn La	10344	2104	01/01/62
15.	Vạn Yên	Taân Phong	Phù Yên	Sơn La	10444	2003	01/01/00
16.	Yên Châu	Thị Trấn	Yên Châu	Sơn La	10418	2103	01/01/61
17.	Chi Nê	Lạc Long	Lạc Thuỷ	Hoà Bình	10547	2029	01/01/73
18.	Hoà Bình	Thị Xã	Hoà Bình	Hoà Bình	10520	2049	01/09/55

ID	Station name	Commune	District	Province	Lon	Lat	Time in Operation
19.	Kim Bôi	Thị Trấn	Kim Bôi	Hoà Bình	10532	2040	01/11/62
20.	Lạc Sơn	T.T Vụ Bản	Lạc Sơn	Hoà Bình	10527	2027	01/02/61
21.	Mai Châu	Phố Vàng	Mai Châu	Hoà Bình	10503	2039	01/01/61
22.	Bắc Mê	Thị Trấn	Bắc Mê	Hà Giang	10522	2244	01/01/64
23.	Bắc Quang	Thị Trấn	Bắc Quang	Hà Giang	10452	2230	01/06/61
24.	H-su Phì	Phố Huyện	Hoàng Su Phì	Hà Giang	10441	2245	01/07/56
25.	Hà Giang	Khu Đoàn Kết	Vị Xuyên	Hà Giang	10458	2249	01/07/56
26.	Bắc Hà	Thị Trấn	Bắc Hà	Lào Cai	10417	2232	01/02/61
27.	Phố Ràng	Thị Trấn	Bảo Yên	Lào Cai	10428	2214	01/06/74
28.	Sa Pa	Thị Trấn	Sa Pa	Lào Cai	10349	2221	01/10/57
29.	Than Uyên	Mường Can	Than Uyên	Lào Cai	10353	2157	01/02/61
30.	Lục Yên	Thị Trấn	Lục Yên	Yên Bái	10443	2206	01/01/61
31.	Mù Căng Chải	Kim Noi	Mù Căng Chải	Yên Bái	10403	2152	01/05/62
32.	Văn Chấn	Khu II	T.T Nghĩa Lộ	Yên Bái	10431	2135	01/01/61
33.	Yên Bái	Thị Xã	Yên Bái	Yên Bái	10452	2142	01/09/55
34.	Chiêm Hoá	Vĩnh Lộc	Chiêm Hoá	Tuyên Quang	10516	2209	01/01/61
35.	Hàm Yên	Cây Số 41	Hàm Yên	Tuyên Quang	10502	2204	01/01/61
36.	Tuyên Quang	Thị Xã	Tuyên Quang	Tuyên Quang	10513	2149	01/01/60
37.	Bắc Cạn	Thị Xã	Bắc Cạn	Bắc Cạn	10550	2209	01/08/56
38.	Chợ Rã	Thị Trấn	Chợ Rã	Bắc Cạn	10543	2227	01/06/61

ID	Station name	Commune	District	Province	Lon	Lat	Time in Operation
39.	Ngân Sơn	Thị Trấn	Ngân Sơn	Bắc Cạn	10559	2226	01/05/61
40.	Định Hoá	T.T Chợ Chu	Định Hoá	Thái Nguyên	10538	2155	01/02/61
41.	Thái Nguyên	Thị Xã	Thái Nguyên	Thái Nguyên	10550	2136	01/10/58
42.	Minh Đài	Minh Đài	Thanh Sơn	Phú Thọ	10503	2110	01/06/72
43.	Phú Hộ	Phù Ninh	Phong Châu	Phú Thọ	10514	2127	01/05/62
44.	Việt Trì	P.Tiên Cát	T.P Việt Trì	Phú Thọ	10525	2118	01/12/60
45.	Tam Đảo	Tam Đảo	Vĩnh Yên	Vĩnh Phúc	10539	2128	01/12/61
46.	Vĩnh Yên	Thị Xã	Vĩnh Yên	Vĩnh Phúc	10536	2119	01/01/60
47.	Bảo Lạc	Thị Trấn	Bảo Lạc	Cao Bằng	10540	2257	01/01/61
48.	Cao Bằng	Thị Xã	Cao Bằng	Cao Bằng	10615	1140	01/08/56
49.	Nguyên Bình	Thị Trấn	Nguyên Bình	Cao Bằng	10557	2239	01/01/61
50.	Trùng Khánh	Thị Trấn	Trùng Khánh	Cao Bằng	10631	2250	01/01/61
51.	Bắc Sơn	Hữu Vĩnh	Bắc Sơn	Lạng Sơn	10619	2154	01/12/62
52.	Đình Lập	Thị Trấn	Đình Lập	Lạng Sơn	10706	2132	01/01/63
53.	Hữu Lũng	Sơn Hà	Hữu Lũng	Lạng Sơn	10621	2130	01/02/61
54.	Lạng Sơn	Mai Phá	Lạng Sơn	Lạng Sơn	10646	2150	01/10/55
55.	Thất Khê	T.T Thất Kh	ê Tràng Định	Lạng Sơn	10628	2215	01/01/60
56.	Bắc Giang	Thị Xã	Bắc Giang	Bắc Giang	10613	2218	01/01/60
57.	Hiệp Hoà	Thị Trấn	Hiệp Hoà	Bắc Giang	10558	2121	01/01/71
58.	Lục Ngạn	Thị Trấn	Lục Ngạn	Bắc Giang	10633	2123	01/01/61

ID	Station name	Commune	District	Province	Lon	Lat	Time in Operation
59.	Sơn Động	Thị Trấn	Sơn Động	Bắc Giang	10651	2120	01/01/61
60.	Bãi Cháy	Thị Xã	Bãi Cháy	Quảng Ninh	10704	2058	01/01/60
61.	Cô Tô	Đảo Cô Tô	Cô Tô	Quảng Ninh	10746	2059	01/10/58
62.	Cửa Ông	Thị Trấn	Cửa Ông	Quảng Ninh	10721	2101	01/01/60
63.	Quảng Hà	Thị Xã	Móng Cái	Quảng Ninh	10745	2127	01/03/79
64.	Tiên Yên	Thị Xã	Tiên Yên	Quảng Ninh	10724	2120	01/02/56
65.	Uông Bí	Yên Chung	Uông Bí	Quảng Ninh	10645	2102	01/01/65

REFRENCES

1. Abhishek K, Singh MP, Ghosh S, Anand A (2012). Weather Forecasting Model using Artificial Neural Network. Procedia Technology, 4, 311-318. doi: 10.1016/j.protcy.2012.05.047
2. ADPC, A. D. P. C.-. 2003. Climate Change and Development in Vietnam: Agriculture and Adaptation for the Mekong Delta Region. Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH.
3. Arnell, N.W., 2004: Climate change and global water resources: SRES emissions and socio economic scenarios. Global Environmen. Chang., 14, 31–52.
4. Arritt, R. W, Rummukainen, M (2011). Challenges in Regional-Scale Climate Modeling. Bulletin of the American Meteorological Society, 92(3), 365-368. doi: 10.1175/2010bams2971.1
5. Bates, B.C., Z.W. Kundzewicz, S. Wu and J.P. Palutikof, Eds., 2008: Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 210 pp
6. Berg N, Hall A, Sun F, Capps S, Walton D, Langenbrunner B, Neelin D (2015). Twenty-First-Century Precipitation Changes over the Los Angeles Region*. Journal of Climate, 28(2), 401-421. doi: 10.1175/jcli-d-14-00316.1
7. Bodri L, Čermák V (2001). Neural Network Prediction of Monthly Precipitation: Application to Summer Flood Occurrence in Two Regions of Central Europe. Studia Geophysica et Geodaetica, 45(2), 155-167. doi: 10.1023/a:1021864227759
8. Bougeault P, Lacarrere P (1989). Parameterization of Orography-Induced Turbulence in a Mesobeta--Scale Model. Monthly Weather Review, 117(8), 1872-1890. doi: 10.1175/1520-0493(1989)117<1872:pooiti>2.0.co;2
9. Brown, C., Greene, A. M., Block, P., & Giannini, A. (2008). Review of downscaling methodologies for Africa climate applications. IRI Technical Report 08-05: IRI Downscaling Report, International Research Institute for Climate and Society, Columbia University.
10. Bukovsky, MS, Karoly DJ (2011). A Regional Modeling Study of Climate Change Impacts on Warm-Season Precipitation in the Central United States*. Journal of Climate, 24(7), 1985-2002. doi: 10.1175/2010jcli3447.1

11. Caldwell P, Chin H-N S, Bader DC, Bala G (2009). Evaluation of a WRF dynamical downscaling simulation over California. Climatic Change, 95(3-4), 499-521. doi: 10.1007/s10584-009-9583-5
12. Castellano G, Fanelli AM (2000). Variable selection using neural-network models. Neurocomputing, 31(1–4), 1-13. doi: https://doi.org/10.1016/S0925-2312(99)00146-0
13. Chen F, Dudhia J (2001). Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. Monthly Weather Review, 129(4), 569-585. doi: 10.1175/1520-0493(2001)129<0569:caalsh>2.0.co;2
14. Christopolos, I., Anderson, S., Arnold, M., Galaz, V., Hedger, M., Klkein, R. J. T. & Le Goulven, K. 2009. The human dimension of climate adaptation: the importance of local and institutional issues.
15. DeMott CA, Randall DA, Khairoutdinov M (2007). Convective Precipitation Variability as a Tool for General Circulation Model Analysis. Journal of Climate, 20(1), 91-112. doi: 10.1175/jcli3991.1
16. Dudhia J (1989). Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. Journal of the Atmospheric Sciences, 46(20), 3077-3107. doi: 10.1175/1520-0469(1989)046<3077:nsocod>2.0.co;2
17. Fuentes U, Heimann D (2000). An Improved Statistical-Dynamical Downscaling Scheme and its Application to the Alpine Precipitation Climatology. Theoretical and Applied Climatology, 65(3), 119-135. doi: 10.1007/s007040070038
18. Giorgi F, Mearns LO (1991). Approaches to the simulation of regional climate change: A review. Reviews of Geophysics, 29(2), 191-216. doi: 10.1029/90RG02636
19. Griffies, S.M., et al., The GFDL CM3 Coupled Climate Model: Characteristics of the Ocean and Sea Ice Simulations. Journal of Climate. 24(13): p. 3520-3544, 2011.
20. Guyennon N, Romano E, Portoghese I, Salerno F, Calmanti S, Petrangeli AB, Tartari G, Copetti D (2013). Benefits from using combined dynamical-statistical downscaling approaches – lessons from a case study in the Mediterranean region. Hydrol. Earth Syst. Sci, 17(2), 705-720. doi: 10.5194/hess-17-705-2013
21. Hong S-Y, Dudhia J, Chen S-H (2004). A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation. Monthly Weather Review, 132(1), 103-120. doi: 10.1175/1520-0493(2004)132<0103:aratim>2.0.co;2

22. Houghton, J. 2004. Global warming, The complete briefing, Cambridge, UK, Cambridge University Press.
23. Hung NQ, Babel MS, Weesakul S, Tripathi NK (2009). An artificial neural network model for rainfall forecasting in Bangkok, Thailand. Hydrol. Earth Syst. Sci, 13(8), 1413-1425. doi: 10.5194/hess-13-1413-2009
24. Hunt, A., & Watkiss, P. (2011). Climate change impacts and adaptation in cities: A review of the literature. Climatic Change, 104(1), 13–49. doi: 10.1007/s10584-010-9975-6
25. IPCC 2001. Climate Change 2001: Impacts, Adaptation and Vulnerability. In: IPCC (ed.).
26. IPCC 2007. Climate Change 2007: Impacts, Adaptation and Vulnerability. In: IPCC (ed.).
27. IPCC, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 1535, 2013.
28. Jansen, E, Overpeck, J., Briffa, K. R., Duplessy, J.-C., Joos, F., Masson Delmotte, V., Olago, D. O., Otoo Bliesner, B., Peltier, W. R., Rahmstorf, S., Ramesh, R., Raynaud, D., Rind, D. H., Solomina, O., Villalba, R. & Zhang. A. D. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, S. Solomon. In: S. Solomon, D. Q., M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor And H.L. Miller, EDS (ed.) Cambridge University Press, Cambridge,.
29. Kjellstrom E, Barring L, Nikulin G, Nilsson C, Persson G, Strandberg G (2016). Production and use of regional climate model projections - A Swedish perspective on building climate services. Clim Serv, 2-3, 15-29. doi: 10.1016/j.cliser.2016.06.004
30. Kobayashi S, Ota Y, Harada Y, Ebita A, Moriya M, Onoda H, Onogi K, Kamahori H, Kobayashi C, Endo H, Miyaoka K, Takahashi K (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. Journal of the Meteorological Society of Japan. Ser. II, 93(1), 5-48. doi: 10.2151/jmsj.2015-001
31. Li Y, Lu G, Wu Z, He H, Shi J, Ma Y, Weng S (2016). Evaluation of Optimized WRF Precipitation Forecast over a Complex Topography Region during Flood Season. Atmosphere, 7(11), 145.

32. LI, Z., Saito, Y., Matsumoto, E., Wang, Y., Tanabe, S. & Lan Vu Q. 2006. Climate change and human impact on the Song Hong (Red River) Delta, Vietnam, during the Holocene. Quaternary International, 144, 4-28.
33. Luu, T. N. M., Garnier, J., Billen, G., Orange, D., N Mery, J., LE, T. P. Q., Tran, H. T. & Le, L. A. 2010. Hydrological regime and water budget of the Red River Delta (Northern Vietnam). Journal of Asian Earth Sciences, 37, 219-228.
34. Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. Journal of Geophysical Research: Atmospheres, 102(D14), 16663-16682. doi: 10.1029/97JD00237
35. MONRE 2012. Climate change, sea level rise scenarios for Vietnam. In: ENVIRONMENT, M. O. N. R. A. (ed.). Hanoi.
36. MONRE, Climate change, sea level rise scenarios for Vietnam. Ministry of Natural Resources and Environment. Vietnam. 2012.
37. MONRE, M. O. N. R. A. E.-. 2006. The State of water environment in 3 river basins of Cau, Nhue – Day and Dong Nai river system. Vietnam environmental report. Vietnam.
38. MONRE, M. O. N. R. A. E.-. 2010. Overview of Vietnam Environment. Vietnam environmental report
39. Nassif AB, Capretz LF, Ho D (2012, 12-15 Dec. 2012). Estimating Software Effort Using an ANN Model Based on Use Case Points. Paper presented at the 2012 11th International Conference on Machine Learning and Applications.
40. National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2000. NCEP FNL Operational Model Global Tropospheric Analyses, Continuing from July 1999. Accessed 12. June. 2014 http://dx.doi.org/10.5065/D6M043C6.
41. NCEP (2000). NCEP FNL Operational model global tropospheric analyses, continuing from July 1999 (Publication no. http://dx.doi.org/10.5065/D6M043C6). from Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory
42. Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, Kadokura S, Wada K, Kato K, Oyama R, Ose T, Mannoji N, Taira R (2007). The JRA-25 Reanalysis. Journal of the Meteorological Society of Japan. Ser. II, 85(3), 369-432. doi: 10.2151/jmsj.85.369

43. Pfeiffer, E. W. 1984. The Conservation of Nature in Viet Nam. Environmental Conservation, 11, 217-221.
44. Pierce DW, Das T, Cayan DR, Maurer EP, Miller NL, Bao Y, Kanamitsu M, Yoshimura K, Snyder Mark A, Sloan Lisa C, Franco G, Tyree M (2012). Probabilistic estimates of future changes in California temperature and precipitation using statistical and dynamical downscaling. Climate Dynamics, 40(3-4), 839-856. doi: 10.1007/s00382-012-1337-9
45. Reynolds RW, Smith TM, Liu C, Chelton DB, Casey KS, Schlax MG (2007). Daily High-Resolution-Blended Analyses for Sea Surface Temperature. Journal of Climate, 20(22), 5473-5496. doi: 10.1175/2007jcli1824.1
46. Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., Schlax, M.G., 2007. Daily high-resolution-blended analyses for sea surface temperature. J. Clim. 20 (22), 5473–5496, http://dx.doi.org/10.1175/2007JCLI1824.1.
47. Salathé Jr EP, Steed R, Mass CF, Zahn PH (2008). A High-Resolution Climate Model for the U.S. Pacific Northwest: Mesoscale Feedbacks and Local Responses to Climate Change. J. Climate, 21, 5708–5726, doi:10.1175/2008JCLI2090.1.
48. Sánchez Lasheras F, Vilán Vilán J. A, García Nieto PJ, del Coz Díaz JJ (2010). The use of design of experiments to improve a neural network model in order to predict the thickness of the chromium layer in a hard chromium plating process. Mathematical and Computer Modelling, 52(7–8), 1169-1176. doi: https://doi.org/10.1016/j.mcm.2010.03.007
49. Sayama, T., et al., Rainfall–runoff–inundation analysis of the 2010 Pakistan flood in the Kabul River basin. Hydrological Sciences Journal. 57(2): p. 298-312, 2012.
50. Seaby LP, Refsgaard JC, Sonnenborg TO, Stisen S, Christensen JH, Jensen KH (2013). Assessment of robustness and significance of climate change signals for an ensemble of distribution-based scaled climate projections. Journal of Hydrology, 486, 479-493. doi: 10.1016/j.jhydrol.2013.02.015
51. Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Huang XY, Wang W, Powers JG (2008). A Description of the Advanced Research WRF Version 3. Technical Note, NCAR/TN-475 + STR.
52. Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Huang, X.Y., and W. Wang, Powers, J.G, A Description of the Advanced Research WRF Version 3. Technical Note. NCAR/TN-475 + STR, 2008.

53. Storch Hv, Langenberg H, Feser F (2000). A Spectral Nudging Technique for Dynamical Downscaling Purposes. Monthly Weather Review, 128(10), 3664-3673. doi: 10.1175/1520-0493(2000)128<3664:asntfd>2.0.co;2
54. Taniguchi, K., Future changes in precipitation and water resources for Kanto Region in Japan after application of pseudo global warming method and dynamical downscaling. Journal of Hydrology: Regional Studies. 8: p. 287-303, 2016.
55. Thuc, T, et al., Estimating Sea Level Rise for Vietnam East Sea. Environmental Science: Climatology. 1(1):p. 73-77, 2017.
56. Tran Anh, Q. and K. Taniguchi, Variations of precipitation and water resources in the Northern part of Vietnam under climate change. Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering). 70(4): p. I_211-I_216. 2014
57. Trenberth, K. E. & Hoar T. J. 1997. El Niño and climate change. Geophysical Research Letters, 24, 3057-3060.
58. Trzaska, S., & Schnarr, E. (2014). A review of downscaling methods for climate change projections (Technical report, United States Agency for International Development). Burlington, Vermont: Tetra Tech ARD
59. Tu JV (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. Journal of Clinical Epidemiology, 49(11), 1225-1231. doi: http://dx.doi.org/10.1016/S0895-4356(96)00002-9
60. Vu, T.T. and R. Ranzi, Flood risk assessment and coping capacity of floods in central Vietnam. Journal of Hydro-environment Research. 14: p. 44-60, 2017.
61. Walton DB, Sun F, Hall A, Capps S (2015). A Hybrid Dynamical–Statistical Downscaling Technique. Part I: Development and Validation of the Technique. Journal of Climate, 28(12), 4597-4617. doi: 10.1175/jcli-d-14-00196.1
62. Walton DB., Hall A, Berg N, Schwartz M, Sun F (2017). Incorporating Snow Albedo Feedback into Downscaled Temperature and Snow Cover Projections for California's Sierra Nevada. J. Climate, 30(4), 1417-1438, doi:10.1175/JCLI-D-16-0168.1.
63. Watanabe, M., et al., Improved Climate Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity. Journal of Climate. 23(23): p. 6312-6335, 2010.
64. Wendemuth A, Opper M, Kinzel W (1993). The effect of correlations in neural networks. Journal of Physics A: Mathematical and General, 26(13), 3165.
65. Wilby, R. L., & Dawson, C. W. (2013). The Statistical DownScaling Model: insights from one decade of application. International Journal of Climatology 33(7), 1707-1719. Wilby, R. L., Troni, J., Biot, Y., Tedd, L., Hewitson, B. C., Smith, D. M., & Sutton, R.

T. (2009). “A review of climate risk information for adaptation and development planning.” International Journal of Climatology 29(9), 1193-1215

66. Yu H.,Wilamowski BM. (2011). Intelligent Systems. In J. D. Irwin (Ed.), Intelligent Systems (pp. Pages 1–16): CRC Press
67. Zhang W, Goh ATC (2016). Multivariate adaptive regression splines and neural network models for prediction of pile drivability. Geoscience Frontiers, 7(1), 45-52. doi: 10.1016/j.gsf.2014.10.003
68. Zveryaev, I. I. & Aleksandrova, M. P. 2004. Differences in rainfall variability in the South and Southeast Asian summer monsoons. INTERNATIONAL JOURNAL OF CLIMATOLOGY, 24, 1091-1107.