

Database assessment of CMIP5 and hydrological models to determine flood risk areas

Ponthip Limlahapun¹ and Hiromichi Fukui²

¹ Fellow of the Japan Society for the Promotion of Science hosted by Chubu University, 1200 Matsumoto-cho, Research Center, Kasugai, Aichi, Japan

² Director of the International Digital Earth Applied Science Research Center, Chubu University, 1200 Matsumoto-cho, Research Center, Kasugai, Aichi, Japan

E-mail: thip.limlahapun@gmail.com

Abstract. Solutions for water-related disasters may not be solved with a single scientific method. Based on this premise, we involved logic conceptions, associate sequential result amongst models, and database applications attempting to analyse historical and future scenarios in the context of flooding. The three main models used in this study are (1) the fifth phase of the Coupled Model Intercomparison Project (CMIP5) to derive precipitation; (2) the Integrated Flood Analysis System (IFAS) to extract amount of discharge; and (3) the Hydrologic Engineering Center (HEC) model to generate inundated areas. This research notably focused on integrating data regardless of system-design complexity, and database approaches are significantly flexible, manageable, and well-supported for system data transfer, which makes them suitable for monitoring a flood. The outcome of flood map together with real-time stream data can help local communities identify areas at-risk of flooding in advance.

1. Introduction

Climate change has been a significant research topic in recent years. Although climate change is largely considered harmful to human habitation, it may also deliver positive impacts to some regions. For example, an ADB economics working paper reported that future projections on grain products (corn and soybean) in China will increase significantly by 2030 [1]. However, the negative consequences of climate change seem more likely, as nature and society are affected worldwide, with the negative impacts on population, urbanization, and even industrial and economic developments continuing to increase every year. These changes could influence environmental and social vulnerabilities regardless of geographical frontiers. There is a growing scientific consensus that inconsistent weather patterns (i.e., extreme wet and stormy conditions) has the potential to intensify the hydrological model [2], leading to flood risk expansion, intensification, and significant societal implications for activities of daily living and quality of life. Locating flood prone areas where serious or frequent flooding occurs could make communities more aware of their risks in certain areas and perhaps take steps to prepare ahead and aid



recovery afterwards.

The Coupled Model Intercomparison Project (CMIP) was established by the working group under the World Climate Research Programme (WCRP) since 1995. The CMIP delivers a standard experimental framework for studying the output of the Global Climate Models (GCMs) [3] and supplies infrastructure to support a large number of climate models, including analysis, validation, inter-comparison, and data access for international climate communities to understand, analyze and facilitate model improvement [4]. Recently, more than 60 CMIP5 models [5] distributed from about 30 various agencies have been assembled and publicized to support climate studies and the estimation of future climates [3]. The large volume of CMIP5 model simulations in capacity of terra bytes has required the assessment of data sharing across research community. Taylor et al. (2012) noted that the CMIP5 output will be archived in data nodes distributed at modeling centers and data centers near where the model output is produced. The nodes will be linked together and the model output will be freely accessible through data portals (or gateways) integrated in a way that retains much of the convenience of a single repository [6]. This worldwide climate research has generated large-scale models and shared results to increase the accuracy of various models working with big datasets [7].

Big data refers to a vast number of datasets, heterogeneous, complex, and difficult to store, analyze, process, and visualize [8]. In other words, big data is often defined by its characteristics of increased data capacity; receiving and processing speed; and a variety of data types, sources, and formats as 3V volume, variety, and velocity [9], which require new forms of processing to enable enhanced decision making, insight discovery, and process [10]. Recently, big data has been involved in many studies (i.e., hydrological, natural, environmental, engineering, and computer sciences studies). The term “big data” may appear in social media like Google or Twitter or in scientific formats (like climate model research, distributed sensors, and satellite observation data) that assist in tracking a dynamic situation, observe patterns, or witness near-real-time events [11].

One challenge of this paper is to obtain information by accessing big data (meteorological CMIP5) and integrate hydrological models to identify inundation areas using a database approach. An intelligent alternative in a proper system will provide a faster analytical data transfer, increased usability, consistency, and better sharing. Nonetheless, model selection from a number of available models is an essential requirement intimately tied to the hydrological models structures, input requirements, and results.

2. Methodology

2.1. Data source

CMIP5 model datasets were obtained from the Data Integration and Analysis System (DIAS) provided coordinating support to the University of Tokyo. We acknowledge the World Climate Research Programme’s Working Group on Couple Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1) for producing and making available their model output.

2.2. Approach

2.2.1. Analyzing precipitation

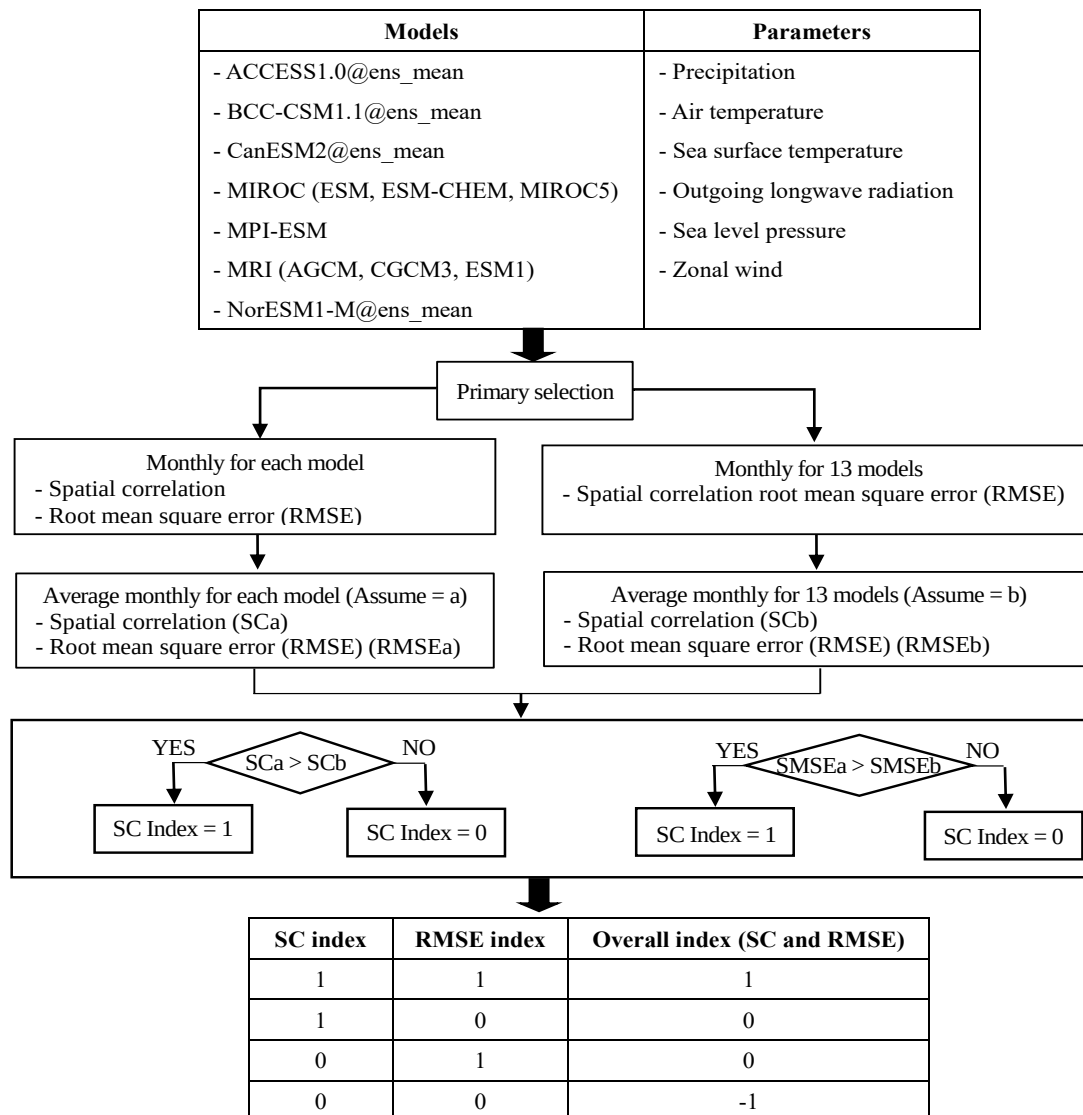
a. Model selection

The model selection approach was evaluated from 20 CMIP5 models based on parameters (precipitation, air temperature, sea surface temperature, sea level pressure and zonal wind). Figure 1 shows model

Table 1. Selected modelling group

Modeling Center (or Group)	Institute ID	Model Name
Beijing Climate Center, China Meteorological Administration	BCC	BCC-CSM1.1 BCC-CSM1.1(m)
Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	MIROC	MIROC-ESM MIROC-ESM-CHEM
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	MIROC4h MIROC5
Max Planck Institute for Meteorology	MPI-M	MPI-ESM
Meteorological Research Institute	MRI	MRI-CGCM3 MRI-ESM1

Source: WCRP [5].

**Figure 1.** Model selection procedure

selection methods. Out of these, seven models (MIROC-ESM-CHEM, MIROC5, MPI-ESM-MR-LR, MRI-CGCM3-ESM1) that returned positive overall index were selected for bias correction over Shonai River in Kasugai.

b. Precipitation extraction

The model selection assembled precipitation information and the degree of spread in future climate projections [12] from multiple GCMs via the CMIP5. In this study, the historical data was taken from 1981-2000 based on a number of models. The observed gridded precipitation data ($0.05^\circ \times 0.05^\circ$ grid) was derived and accessed from the DIAS, which used JRA55 and other reanalysis data as reference data for comparison with CMIP5 data provided by the University of Tokyo [13].

2.2.2. Examining discharge with flood analysis system model

The Public Works Research Institute (PWRI), the International Centre for Water Hazard and Risk Management (ICHARM), under the auspices of UNESCO, developed the Integrated Flood Analysis System (IFAS), a satellite rainfall based simulation model for determining rainfall-runoff distribution [14]. The data required to calculate discharge includes elevation, land cover, soil, watershed boundary, and daily rainfall. Digital Elevation Model (DEM) 5-metre grid data was download from Geospatial Information Authority of Japan (GSI). The process of filling DEM; generating water flow; flow accumulation to lower cell and direct to outlet cell; and output as watershed boundary, and daily rainfall. Digital Elevation Model (DEM) 5-metre grid data was download from Geospatial Information Authority of Japan (GSI). IFAS is an executed installation program. Users proceed step-by-step through graphical interfaces provided by the system. To reduce processing time, we adapted the program by running it from a command line. The output from IFAS was graph and table, which could be plotted for a hydrograph (rainfall-discharge) using original IFAS program based on a grid-cell or a MS Excel spreadsheet.

2.2.3. Determining an inundation map with a hydrological model

The Hydrologic Engineering Center River Analysis System (HEC-RAS) is a hydraulic model to execute a steady or unsteady flow calculation [15]. The data preparation was created from HEC-GeoRas on ArcGIS extension. The Hec-GeoRAS interface allows the preparation of geometric data in so-called RAS layers (streamline, river-bank, cross-sectional, and floodplain boundary) for importing and processing simulation results from HEC-RAS [15]. Data requires for creating an import file includes the UTM projected digital terrain model (TIN format) and background Landsat imagery. Once the RAS layers were created, attribute data could be assigned using the GeoRas tools (river and reach name to stream network, channel-left-right to rivers, overbanks, and flow paths). The layers were then ready to written out in spatial data file (SDF) format to an extensible markup language (XML) file, but it is practical to validate data and attributes before importing to the HEC-RAS geospatial data exchange. The final outcome is to create the inundation and velocity spatial distribution of flooding and identify affected areas.

3. Results

3.1. Results of precipitation analysis derived from CMIP5

The historical data from each model was verified similarly to the references in Asian Precipitation-Highly Resolved Observational Data (APHRODITE). Based on the literature and reviewed reports, the length and intensity of rainfall caused flooded along Shonai River in 2000. The Tropical Rainfall

Measuring Mission (TRMM) dataset was also collected and displayed to verify rain-patterned and intensity performed in the areas. The future climate prediction shows a large spread in projections, and the shifting rain pattern is obvious. This result occurs due to variations in the climate sensitivity that determines the influence magnitude of the average global response, and large variations in spatial patterns of change, particularly for precipitation [16].

3.2. Results of discharge analysis using flood analysis model

The precipitation extracted from DIAS system was then formatted for use in IFAS. The data were structured into daily records and simulations were conducted. The result calculated based on the year 2000 data average discharge amount of 2,300 m³/s, showed good agreement with the JMA records of approximately 2,500 m³/s. The same procedure for deriving precipitation data from the CMIP5 model was applied to estimate future flood peak. Additionally, the near-real-time observation rain gauge and water level were collected at 10-minute intervals to observe stream flow. The thresholds were used for three level of alert system (warning, evacuation, and flooding). The verification was crossed checked with the observation data. These results show that the IFAS model could be a useful tool for the overflow estimates and accessible CMIP5 data for prediction of flood inundation process in any area of interest. The TRMM satellite earth observation and near-real-time system assist to monitor rainfall influences to surrounding to global climate and the current situation, accordingly.

3.3. Results of inundation map extracting from hydrological model

The final process to generate an inundation map was derived using HEC, based on the discharge result and other structure (i.e., bridge, land use, and so on). The HEC estimates water flows through system and compute water surfaces and evaluates flood encroachments.

The serial result data input to and output from each process can be transferred among modeling and simulation techniques. This database approach is recognized for its speed and flexibility while also providing the ability to manage with amounts of data. Databases are designed to store data, so data coming into the system must be cautiously considered.

4. Conclusion

Empirical relationships between rainfall rates generated from the big data model over the past 20 years and the characteristics of the actual flood have been formulated. The changes in rainfall intensity and its behavior have been forecasted for the 20-year period from 2045 to 2065. The result shows an increase in the amount and duration of rainfall, which may lead to significant future flooding, escalation in the flood peak, and total flood volume. In the context of disaster preparedness, the results may be shocking, but it is objectively apparent that communities may find themselves at considerable risk and such models afford them the opportunity to take specific actions to reduce their risks before it is too late.

A complex issue like being prepared for a natural disaster may not be solvable with a simple solution or the application of one particular model. A selection of appropriate models from a number of available standards is challenging due to the variability of spatial and temporal aspects. This paper proposed alternative ways to approach results to select models, worked backwards and illustrated outcomes we need as a basic notion. Then we craft solutions to the underlying issues, factors, and resources to solve these problems. Three selection models were applied, using one model's output values to serve as another model's input, ultimately producing a delineated inundation map. By developing a database to act as the assembly tool for each component, receive rainfall intensity, transfer parameters of flows in

channel network, and storing forecast datasets that are capable of generating the distributed and predicting inundated areas, it has been found that a database approach is an efficient methodology.

References

- [1] Tun L, Xiaoyun L, Guanghua W, Xian X, and Yongsheng Z 2011 Impacts on climate change on the people's Republic of China's grain output—Regional and crop perspective. *ADB economics working paper series*. **243**.
- [2] Guoyi Z et al. 2011 Quantifying the hydrological responses to climate change in an intact forested small watershed in Southern China. *Global Change Biology*. **17** pp. 3736-3746.
- [3] Guo Y, Dong W J, Ren F M, Zhao Z C, and Huang J B 2013 Surface Air Temperature Simulations over China with CMIP5 and CMIP3. *Advances in climate change research*. **4** No. 3 pp. 145-152.
- [4] David A R and Richard A W Climate Models and Their Evaluation. <https://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter8.pdf>
- [5] WCRP (World Climate Research Programme), CMIP Coupled Model Intercomparison Project, <http://cmip-pcmdi.llnl.gov/index.html?submenuheader=0>
- [6] Taylor K E, Ronald J S, and Gerald A M 2012 An Overview of CMIP5 and the Experiment Design. *American Methodological Society*. pp. 485-498.
- [7] Jon B 2014 Weathering the flood of Big data in climate research. *ESnet News*, [Accessed on March 2, 2016, <https://www.es.net/news-and-publications/esnet-news/2014/weathering-the-flood-of-big-data-in-climate-research/>]
- [8] Akhtar U and Mehdi H 2015 Big Data Mining Based on Computational Intelligence and Fuzzy Clustering. *Handbook of Research on Trends and Future Directions in Big Data and Web Intelligence*. 130-148.
- [9] Siddesh G M, Srinidhi H and Srinivasa K G 2014 Driving Big Data with Hadoop Technologies. *Handbook of Research on Cloud Infrastructures for Big Data Analytics*. 232-262.
- [10] Laney D 2001 3D Data Management. *Controlling Data Volume, Velocity and Variety*.
- [11] Tsou M H 2015 Mapping human dynamics with social media for disaster alerts. *Proceedings of ESRI users conference July 20-24, 2015, San Diego, California, U.S.A.*
- [12] Carol F M, Richard G J 2016 How representative is the spread of climate projections from the 5 CMIP5 GCMs used in ISI-MIP? *Climate Services*. doi: 10.1016/j.cliser.2016.02.001
- [13] DIAS, Data Integration and Analysis System, The University of Tokyo, <http://www.dias.nii.ac.jp/modelvis/cmip5/>
- [14] (The) International Centre for Water Hazard and Risk Management (ICHARM) 2014 IFAS ver.2.0 technical manual. *Public Works Research Institute*. Japan, p.259.
- [15] OPW, The Office of Public Work 2009 The planning system and flood risk management. *Guidelines for planning Authorities*. p.73.
- [16] Guojun G and Robert F A 2015 Spatial Patterns of Global Precipitation Change and Variability during 1901-2010. *Journal of Climate*. **28** No. 11.

Acknowledgement

This project and publication were funded by the Japan Society for the Promotion of Science (JSPS) under a Postdoctoral Fellowships for Overseas Researchers and Grants-in-Aid for Scientific Research (KAKENHI) grant. Any opinions, findings, and conclusions expressed in this publication are those of the authors and do not necessarily reflect the views of the JSPS.