

Application of performance metrics to climate models for projecting future river discharge in the Chao Phraya River basin

Satoshi Watanabe¹, Yukiko Hirabayashi¹, Shunji Kotsuki², Naota Hanasaki³, Kenji Tanaka⁴, Cherry May R. Mateo⁵, Masashi Kiguchi⁵, Eiji Ikoma⁶, Shinjiro Kanae⁷ and Taikan Oki⁵

¹Institute of Engineering Innovation, The University of Tokyo, Japan

²Graduate School of Engineering, Kyoto University, Japan

³National Institute for Environmental Studies, Japan

⁴Disaster Prevention Research Institute, Kyoto University, Japan

⁵Institute of Industrial Science, The University of Tokyo, Japan

⁶Earth Observation Data Integration & Fusion Research Initiative, The University of Tokyo, Japan

⁷Department of Civil Engineering, Tokyo Institute of Technology, Japan

Abstract:

Future river discharge in the Chao Phraya River basin was projected based on the performance of multiple General Circulation Models (GCMs). We developed a bias-corrected future climate dataset termed IDD (IMPAC-T Driving Dataset) under which the H08 hydrological model was used to project future river discharge. The IDD enabled us to conduct a projection that considered the spread in projections derived from multiple GCMs. Multiple performance-based projections were obtained using the correlation of monsoon precipitation between GCMs and several observations. The performance-based projections indicated that future river discharge in September increased 60%–90% above that of the retrospective simulation. Our results highlight the importance of appropriate evaluation for the performance of GCMs.

KEYWORDS climate change; river discharge; Chao Phraya River basin; general circulation model (GCM); IMPAC-T Driving Dataset (IDD)

INTRODUCTION

Climate change will alter flood and drought risks and have substantial impacts on society. In the Chao Phraya River basin, a severe flood in 2011 caused extensive damage (Komori *et al.*, 2012). The frequency of such large floods may increase due to ongoing climate change (Hirabayashi *et al.*, 2013). Moreover, an increase in the occurrence of severe droughts has also been projected (Hunukumbura and Tachikawa, 2012). To manage risks and reduce damage, more precise assessment of the impacts of climate change on river discharge throughout the basin is needed.

Several studies projected future river discharge in the Chao Phraya River basin using future climate forcings from general circulation models (GCMs). For example, Ogata *et al.* (2012) simulated future river discharge using the outputs of three GCMs, with their results suggesting that peak discharge may increase within the next three decades.

Champhong *et al.* (2013) evaluated the robustness of projected changes by comparing several GCMs and ensemble simulations. Such future hydrological simulations using outputs from multiple GCMs are important for assessing the uncertainty in projections. In addition, consideration of the spread in GCM projections should be included in the analysis to appropriately evaluate extremes, as there can be significant differences among projections.

The ensemble mean is one method for considering the inter-model spread of projections. A basic approach is to use equal weight (i.e., the arithmetic mean) for all GCMs (hereafter, the simple ensemble method). Weiland *et al.* (2012) applied the simple ensemble method for global assessment of the effects of climate change on hydrological regimes and their accompanying uncertainties. However, other studies (e.g. Nohara *et al.* 2006) have used another approach: the weighted ensemble mean. In order to produce one aggregated model weight, a metric for evaluating GCMs is needed.

Despite efforts to compare the performance of individual GCMs, no definitive metric for characterizing GCM output has been found in previous studies. Gleckler *et al.* (2008) evaluated the reproducibility of the global distribution of 22 variables simulated by multiple CMIP3 GCMs and proposed a metric based on ranking the degree of relative error among GCMs. However, they concluded that their combination of metrics is unlikely to be optimally suited for all applications. Thus, a metric is needed that reflects the underlying constraints in the Chao Phraya basin in order to properly assess the impacts of climate change in that basin. Yokoi *et al.* (2011) pointed out the importance of physical and dynamic constraints, which correlate to performance, in obtaining appropriate groupings to avoid double counting of performance skills among dependent variables.

Evaluation of GCM performance in simulating monsoon behavior is important for projecting future discharge in the Chao Phraya River basin. As noted by Singhrattna *et al.* (2005), the reproducibility of monsoon precipitation is a major constraint for prediction of river discharge in Thailand. It is also important to better understand possible changes in monsoon characteristics. For example, in a previous investigation of monsoon precipitation using CMIP5 GCMs,

Correspondence to: Satoshi Watanabe, Institute of Engineering Innovation, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan. E-mail: stswata@sogo.t.u-tokyo.ac.jp ©2014, Japan Society of Hydrology and Water Resources.

Received 2 October, 2013
Accepted 27 December, 2013

most of the models projected enhanced global monsoon activity (Hsu *et al.*, 2013). Furthermore advanced onset and delayed retreat of monsoons were projected using GCMs that showed high reproducibility in historical monsoon simulations (Lee and Wang, 2012). Given the above, several metrics to evaluate the performance of each GCM to reproduce monsoon precipitation were proposed to estimate performance-based projection in this study.

The present study was performed to investigate the effects of performance metrics and to estimate the spread of projections derived from the differences in multiple performance metrics. To achieve these objectives, multiple future projections using available GCM outputs were conducted in the Chao Phraya River basin and multiple weighted ensemble means were obtained using the proposed multiple metrics related to monsoon precipitation. Here, we compare the projected results obtained and discuss the characteristics of each projection.

METHODOLOGY

The Chao Phraya River basin

The Chao Phraya River basin has the largest catchment area in Thailand (160,000 km²). Nakhon Sawan is located in the middle of the basin, and has a catchment area of 110,000 km² (68% of the area of the river basin). The climate of the basin can be clearly divided into dry and wet seasons. The precipitation from May to October accounts for almost 90% of the total annual precipitation (Kure and Tebakari, 2012), and precipitation is generally higher in the northern mountainous area compared to other areas (Kotsuki *et al.*, 2014).

Model description

A hydrological simulation was conducted using the H08 model (Hanasaki *et al.*, 2008). This model consists of six modules: land surface hydrology, river routing, crop growth, reservoir operation, environmental flow requirement estimation, and anthropogenic water withdrawal. Only the land surface hydrology and river routing modules were used for the projection in this study because our objective was to investigate the effects of climate change on river discharge, excluding any effects from human activity. The basin boundary and flow direction maps at 5-arc-minute resolution over the Chao Phraya River were developed by manually digitizing a printed map. Details are provided in Supplement Information S1.

Atmospheric forcing data

We have developed a climate-forcing dataset from GCM outputs for use in conducting hydrological simulations. The H08 hydrological model requires daily temperature, precipitation, long- and short-wave radiation, specific humidity, surface pressure, and wind speed data. We used data from the IMPAC-T forcing dataset (IFD) developed by Kotsuki *et al.* (2014) as the atmospheric forcing for the hydrological simulation of the reference period (1980–1999). The IFD is a daily forcing dataset at 5-arc-minute resolution based on in situ observations and reanalysis data.

Future atmospheric forcing (2080–2099), the IMPAC-T

Driving Dataset (IDD), was created by correcting the bias in GCM outputs from the archive of Fifth Coupled Model Intercomparison Project (CMIP5) of the Program for Climate Model Diagnosis and Intercomparison (PCMDI) using the IFD. Nine GCMs were selected according to the availability of daily data for both the reference (1980–1999) and projection (2080–2099) periods. These nine models were bcc-csm1-1-m (hereafter denoted as bc1m), CNRM-CM5 (cnc5), CSIRO-Mk3.6 (cs36), GFDL-ESM2M (ge2m), HadGEM2-ES (hg2e), INM-CM4 (inc4), IPSL-CM5A-LR (ip51), MIROC5 (mir5), and MRI-CGCM3 (mrc3). Details on these GCMs are provided in Supplement Table SI. The GCMs were selected from nine institutions because different versions of GCMs from the same institution may not be independent. Although the IDD includes climate-forcing data for two time periods for the two representative concentration pathway (RCP) scenarios, this study focused on the results obtained using the 2080–2099 dataset under the RCP 8.5 scenario.

Bias correction

Bias correction of GCM output is a key issue for developing forcing datasets from GCM output because the correction method applied has a large impact on the results (Watanabe *et al.*, 2012). Therefore, we developed an advanced bias correction method, in which the trend of variables from the reference to the projection period was preserved. The trend-preserving assumption is reasonable for bias correction and has been used in previous studies (e.g., Hempel *et al.*, 2013). In the bias correction process, monthly variation in GCM data was corrected and then daily variation was corrected using the corrected monthly variation. The trends in GCM output (i.e., changes in mean and standard deviation) from the reference to the projection period were preserved for both monthly and daily variations. Details on the proposed bias correction method are presented in Supplemental Information S2.

Simulation runs

We conducted two hydrological simulations. One was a retrospective simulation in the reference period, while the other was a future simulation in the projection period. The IFD dataset was used for the retrospective simulation and the IDD dataset for the future simulation. For the future simulation, the nine projections were obtained using future atmospheric forcings from the nine GCMs. Both simulations used parameters from the H08 model optimized for the Chao Phraya River basin by Mateo *et al.* (2012).

Observed precipitation data for GCM evaluation

Three observation-based precipitation products were used to evaluate GCM performance pertaining to accurate replication of monsoon precipitation around Thailand, including the Chao Phraya River basin: the Global Precipitation Climatology Project (GPCP) (Xie *et al.*, 2003), Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) (Xie and Arkin, 1997), and Asian Precipitation–Highly Resolved Observed Data Integration Towards Evaluation of the Water Resources (APHRODITE) (Yatagai *et al.*, 2012). GPCP and CMAP are global gauge-satellite precipitation products covering land and sea surfaces, whereas APHRODITE covers land surfaces only

because it is created from a terrestrial rain-gauge observation network.

Application of climate metrics for model ensembles

In this study, we evaluated GCM performance in terms of the reproducibility of climatology. We focused on the performance of climatological spatiotemporal patterns in monsoon precipitation in Southeast Asia for each GCM, because monsoon precipitation is the most dominant climate variable when simulating the hydrological cycle in the Chao Phraya River basin. We focused on precipitation in two regions: a large-scale area around Thailand (0–30°N, 95–110°E) (hereafter, large-scale metric) and a terrestrial area around the Chao Phraya River basin (12–21°N, 96–105°E) (hereafter, terrestrial metric). The target region was somewhat smaller than that used in similar research that evaluated the Asian summer monsoon (20°S–50°N, 40°E–160°E; Sperber *et al.*, 2012) and the summer Eastern Asian Monsoon (10–55°N, 115–145°E; Miyakawa *et al.*, 2013). In both metrics, the 20-year (1980–1999) average pentad mean precipitation for a time-latitude section, averaged over each latitude band in the target region, was calculated based on precipitation modeled by the GCMs before bias correction. This precipitation, modeled during the monsoon season

(days 120–300 of the year), was compared with those of the precipitation products (CMAP, GPCP, and APHRODITE) in this study. Comparisons were conducted using pattern correlation (PC) (Watterson, 1996). The CMAP and GPCP datasets were used as observational data for the large-scale metric, and APHRODITE data were used for the terrestrial metric.

The weighted ensemble mean was then calculated considering each metric, using R^2 weighting. Nohara *et al.* (2006) compared weighting methods for ensemble means for river discharge from multiple GCMs and found that R^2 weighting was an efficient method. The process for applying the R^2 weighting method is somewhat subjective, but our main purpose was to show how the final discharge projection could be altered when climate metrics were applied; we acknowledge that there are other methods for obtaining model ensemble means to project discharge.

RESULTS

Performance of GCMs

Figure 1 shows the spatiotemporal pattern of the pentad precipitation averaged over each latitude band from the 120th

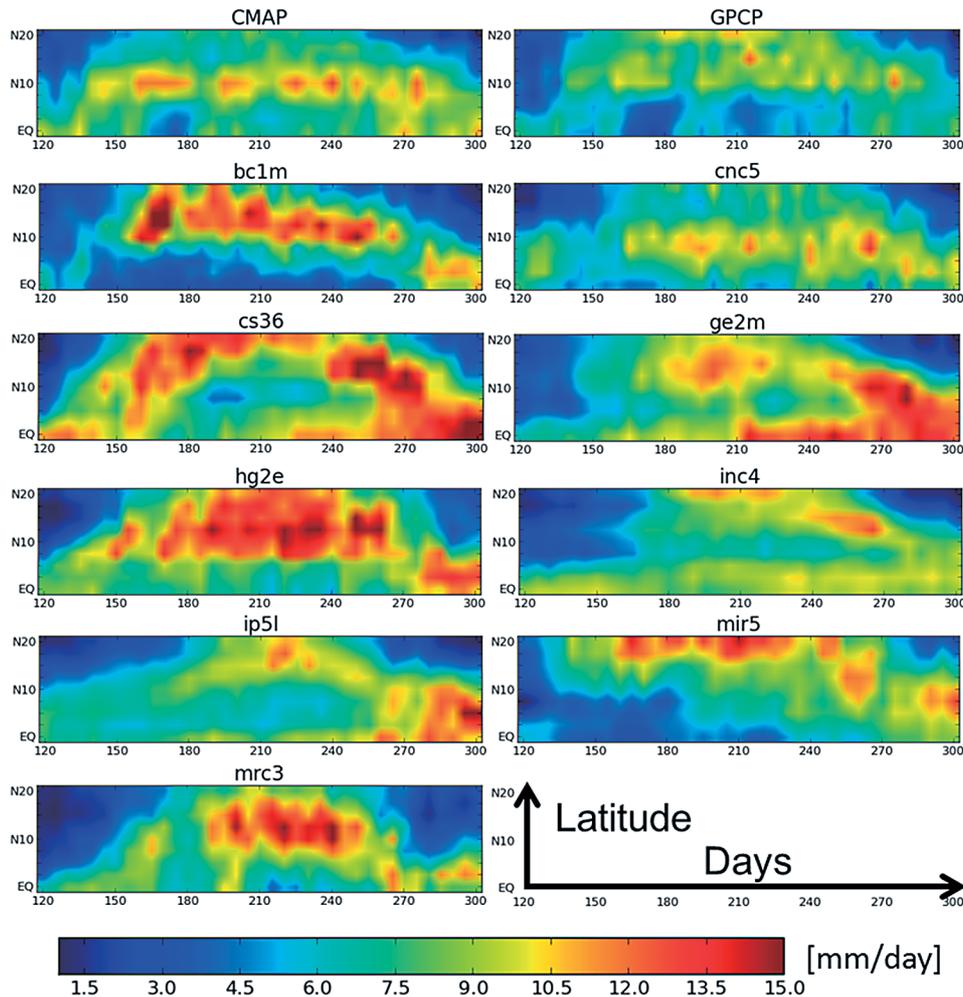


Figure 1. The spatiotemporal pattern in pentad precipitation averaged over each latitude band from day 120 to 300 of year. The unit is [mm/day]

Table I. Pattern correlation (PC) between GCMs and observations. Numbers in brackets indicate the rank of each GCM

	large scale				terrestrial	
	GPCP		CMAP		APHRODITE	
	PC	rank	PC	rank	PC	rank
bc1m	0.54	(5)	0.67	(1)	0.35	(8)
cncm5	0.67	(1)	0.40	(9)	0.56	(5)
cs36	0.49	(7)	0.58	(4)	0.55	(6)
ge2m	0.54	(6)	0.46	(7)	0.62	(2)
hg2e	0.66	(2)	0.63	(3)	0.58	(4)
inc4	0.43	(8)	0.43	(8)	0.62	(2)
ip51	0.60	(4)	0.46	(6)	0.66	(1)
mir5	0.25	(9)	0.66	(2)	0.19	(9)
mrc3	0.61	(3)	0.50	(5)	0.50	(7)

to the 300th day of the year, and Table I shows the PC between the GCM outputs and observations. The difference between CMAP and GPCP was clear around the 10°N latitude band, although it was smaller than the difference among the GCMs. In addition, PC values for GPCP and CMAP for the same GCM were significantly different in some GCMs (e.g., cnc5 and mir5) (Table I). Given these findings, we used the PC values for both GPCP and CMAP to evaluate the GCM performances at the large scale.

The results for the large-scale metric indicated that the timing of monsoon retreat was similar between GCM projections and observational data, but the timing of monsoon onset in some GCMs differed from that observed. These findings imply that precipitation and river discharge around the onset of the monsoon differ among the GCMs. In addition, there were no clear similarities among the results of the large-scale metrics and the terrestrial metric. No GCM showed a result better or worse than others for all metrics. We calculated a weighted ensemble mean using the R^2 score of the PC to obtain a performance-based projection. Hereafter, we define the weighted ensemble mean using the R^2 score of the PC for GPCP, CMAP, and APHRODITE as the ENS_{GPCP} , ENS_{CMAP} , and ENS_{APHRO} , respectively.

Ensemble projection considering climate metrics

Figures 2 and 3 show the projected monthly river discharge averaged over the future period (2080–2099) as simulated by the H08 model, and the rate of change in river discharge from the reference to the projection period. There was a large difference among ensembles in projected maximum monthly river discharge; the difference in September was almost 30% in the retrospective simulation. The ENS_{APHRO} projection was generally smaller than the other projections, and a shift in maximum monthly river discharge from the reference to the projection period appeared in all projections. River discharge in April increased in the simple and ENS_{CMAP} projections whereas it decreased in the ENS_{GPCP} and ENS_{APHRO} projections.

Each GCM projection had large spread. The highest rate of change was over 300% more than that of the retrospective simulation. Another remarkable result was that one of the GCMs projected increased discharge around the peak season. It should be noted that the effects of such extreme projections are weakened in the ensemble projection.

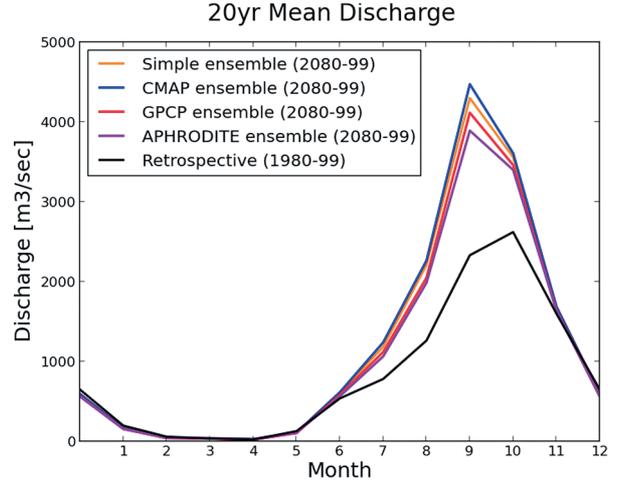


Figure 2. Comparisons between simple ensemble and performance-based projections of 20-year mean monthly river discharge at Nakhon Sawan. Black line indicates retrospective simulation by IFD; yellow line indicates projection of the simple ensemble; blue, red, and purple lines indicate performance-based projections for CMAP, GPCP, and APHRODITE, respectively

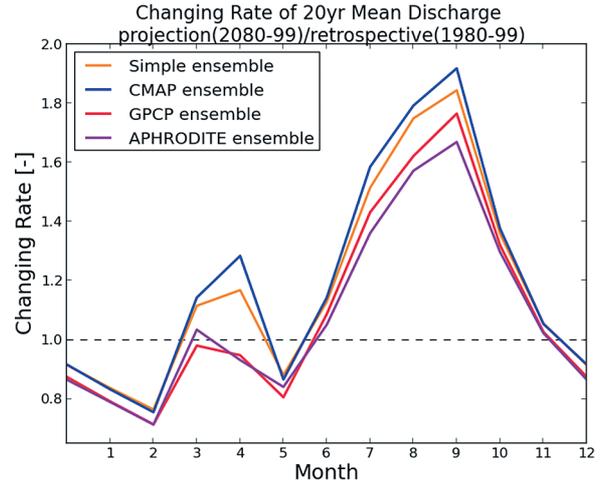


Figure 3. Comparison of the rate of change in river discharge from the reference simulation to the future projection

DISCUSSION

We found that the performance of GCMs differed according to the observations used, which indicates that differences in observations should not be neglected during the assessment of projections. Interestingly, the projected river discharge of ENS_{CMAP} was higher than that of the simple ensemble, whereas those of ENS_{GPCP} and ENS_{APHRO} were lower. This is because the ensemble weighting of ENS_{CMAP} is larger for the GCMs that project higher river discharge than others, and vice versa. The results indicated that the weighting of ENS_{GPCP} and ENS_{APHRO} is lower for the GCMs that project higher river discharge.

Extreme projection in river discharge with some GCMs (e.g., the projection of bc1m was 360% higher than that of

the retrospective simulation in September) could result in part from the bias correction for precipitation. The bias correction method that we employed corrected GCM data separately for each month; thus, corrected data would tend to show unreasonable values if a GCM was biased in its timing of peak monthly precipitation. For example, if the maximum monthly discharge projected by a GCM differs from observational data, the changed projection in the maximum month would be applied to another month. Because the difference from the reference to the projection period in the peak month is often larger than that in other months, the corrected data tended to generate unrealistic results. It is important to evaluate the efficiency of bias correction to appropriately project future changes. We used a terrestrial metric to evaluate the ability to reproduce spatiotemporal patterns in relation to bias correction by considering only terrestrial precipitation.

PC is not an optimal metric for evaluating the efficiency of bias correction because both the pattern correlation and the absolute difference are important when correcting bias. Therefore, we introduced a metric for evaluating the GCMs in relation to spatiotemporal patterns and absolute error (Details are provided in Supplement Information S3) and checked the relationship between reproducibility and characteristics of the future projection. This indicated that the three GCMs with lowest reproducibility corresponded to the GCMs with the lowest and the two highest projections, which implies that there is a relationship between reproducibility and the efficiency of bias correction. However, further research is required to confirm this hypothesis.

SUMMARY

A performance-based river projection was conducted for the Chao Phraya River basin that considered the ability of GCMs to reproduce monsoon precipitation. We developed future forcing data by applying a new, advanced bias correction method. Two types of metrics were applied to investigate the reproducibility of each GCM, and the results from these metrics enabled us to construct performance-based projections. These projections showed a 60–90% increase in discharge in September in relation to reference simulations. Differences between the projections and observational data were examined to evaluate model performance; the projections were within 30% of reference simulations. The results indicated that the performance-based projection of the river discharge in the Chao Phraya River had non-negligible uncertainty derived from the differences in observational data. In addition, it was implied that the uncertainty of bias correction also had an impact on the results of the projection.

ACKNOWLEDGEMENTS

This work was supported by the Program for Risk Information on Climate Change from the Ministry of Education, Culture, Sports, Science, and Technology-Japan (MEXT); the Integrated Study Project on Hydro-Meteorological Prediction and Adaptation to Climate

Change in Thailand (IMPAC-T) from Science and Technology Research Partnership for Sustainable Development, JICA-JST, Japan; CREST project from Japan Science and Technology Agency; Funding Program for Next Generation World-Leading Researchers (NEXT Program) from the Cabinet Office.

SUPPLEMENTS

Detail explanations of river and administrative boundary maps (Supplement information S1), the bias correction method (Supplement information S2), a performance metric using normalized cross correlations (Supplement information S3), and the list of GCMs used in this study (Supplement Table SI) are provided in the supplementary materials. The catchment areas for major river gauging stations, areas within administrative boundaries, the basin boundary and flow direction in the Chao Phraya River Basin, an administrative boundary map, and the relationship between normalized cross correlation (NCC) and the future projection of river discharge are shown in Table SII, SIII, Figure S1, S2 and S3, respectively.

REFERENCES

- Champhong A, Komori D, Kiguchi M, Sukhannaphan T, Oki T. 2013. Future projection of mean river discharge climatology for the Chao Phraya River basin. *Hydrological Research Letters* **7**: 36–41. doi: 10.3178/HRL.7.36.
- Gleckler PJ, Taylor KE, Doutriaux C. 2008. Performance metrics for climate models. *Journal of Geophysical Research* **113**: D06104. doi: 10.1029/2007JD008972.
- Hanasaki N, Kanae S, Oki T, Masuda K, Motoya K, Shirakawa N, Shen Y, Tanaka K. 2008. An integrated model for the assessment of global water resources – Part 1: Model description and input meteorological forcing. *Hydrology and Earth System Sciences* **12**: 1007–1025. doi: 10.5194/hess-12-1007-2008.
- Hempel S, Frieler K, Warszawski L, Schewe J, Piontek F. 2013. A trend-preserving bias correction – the ISI-MIP approach. *Earth System Dynamics* **4**: 219–236. doi: 10.5194/esd-4-219-2013.
- Hirabayashi Y, Mahendran R, Koirala S, Konoshima L, Yamazaki D, Watanabe S, Kim H, Kanae S. 2013. Global flood risk under climate change. *Nature Climate Change* **3**: 816–821. doi: 10.1038/NCLIMATE1911.
- Hsu PC, Li T, Murakami H, Kitoh A. 2013. Future change of the global monsoon revealed from 19 CMIP5 models. *Journal of Geophysical Research: Atmospheres* **118**: 1247–1260. doi: 10.1002/jgrd.50145.
- Hunukumbura PB, Tachikawa Y. 2012. River discharge projection under climate change in the Chao Phraya River basin. *Journal of the Meteorological Society of Japan Ser. II* **90A**: 137–150. doi: 10.2151/jmsj.2012-A07.
- Komori D, Nakamura S, Kiguchi M, Nishijima A, Yamazaki D, Suzuki S, Kawasaki A, Oki K, Oki T. 2012. Characteristics of the 2011 Chao Phraya River flood in Central Thailand. *Hydrological Research Letters* **6**: 41–46. doi: 10.3178/HRL.6.41.
- Kotsuki S, Tanaka K, Watanabe S. 2014. Projected hydrological changes and their consistency under future climate in the Chao Phraya River Basin using multi-model and multi-scenario of CMIP5 dataset. *Hydrological Research Letters* **8**:

- 27–32. doi: 10.3178/hrl.8.27.
- Kure S, Tebakari T. 2012. Hydrological impact of regional climate change in the Chao Phraya River basin, Thailand. *Hydrological Research Letters* **6**: 53–58. doi: 10.3178/hrl.6.53.
- Lee JY, Wang B. 2014. Future change of global monsoon in the CMIP5. *Climate Dynamics* **42**: 101–119. doi: 10.1007/s00382-012-1564-0.
- Mateo CMR, Hanasaki N, Komori D, Tanaka K, Yoshimura K, Kiguchi M, Oki T. 2012. Optimization of H08 Land Surface Model Parameters in the Chao Phraya River Basin to Reproduce the 2011 River Discharge. *Annual Conference of Japan Society of Hydrology and Water Resources*, 2012 Hiroshima, Japan; 78–79.
- Miyakawa T, Yokoi S, Hirota N, Hirahara S, Inoue T, Takayabu YN. 2013. An iterative selection method: A summer eastern asian metric for CMIP3 climate models as an example. *Journal of the Meteorological Society of Japan* **91**: 305–321. doi: 10.2151/jmsj.2013-305.
- Nohara D, Kitoh A, Hosaka M, Oki T. 2006. Impact of climate change on river discharge projected by multimodel ensemble. *Journal of Hydrometeorology* **7**: 1076–1089. doi: 10.1175/JHM531.1.
- Ogata T, Valeriano OCS, Yoshimura C, Liengcharernsit W, Hirabayashi Y. 2012. Past and future hydrological simulations of Chao Phraya River basin. *Journal of Japan Society of Civil Engineers Ser. B1* **68**: I_97–I_102.
- Singhtrattna N, Rajagopalan B, Kumar KK, Clark M. 2005. Interannual and interdecadal variability of Thailand summer monsoon season. *Journal of Climate* **18**: 1697–1708. doi: 10.1175/JCLI3364.1.
- Sperber KR, Annamalai H, Kang IS, Kitoh A, Moise A, Turner A, Wang B, Zhou T. 2013. The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. *Climate Dynamics* **41**: 2711–2744. doi: 10.1007/s00382-012-1607-6.
- Watanabe S, Kanae S, Seto S, Yeh PJF, Hirabayashi Y, Oki T. 2012. Intercomparison of bias-correction methods for monthly temperature and precipitation simulated by multiple climate models. *Journal of Geophysical Research* **117**: D23114. doi: 10.1029/2012JD018192.
- Watterson IG. 1996. Non-dimensional measures of climate model performance. *International Journal of Climatology* **16**: 379–391. doi: 10.1002/(SICI)1097-0088(199604)16:4<379::AID-JOC18>3.0.CO;2-U.
- Xie P, Arkin PA. 1997. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society* **78**: 2539–2558. doi: 10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.
- Xie P, Janowiak JE, Arkin PA, Adler RF, Gruber A, Ferraro RR, Huffman GJ, Curtis S. 2003. GPCP Pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *Journal of Climate* **16**: 2197–2214. doi: 10.1175/2769.1.
- Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of American Meteorological Society* **93**: 1401–1415. doi: 10.1175/BAMS-D-11-00122.1.
- Yokoi S, Takayabu YN, Nishii K, Nakamura H, Endo H, Ichikawa H, Inoue T, Kimoto M, Kosaka Y, Miyasaka T, Oshima K, Sato N, Tsushima Y, Watanabe M. 2011. Application of cluster analysis to climate model performance metrics. *Journal of Applied Meteorology and Climatology* **50**: 1666–1675. doi: 10.1175/2011JAMC2643.1.