Bias correction of d4PDF using a moving window method and their uncertainty analysis in estimation and projection of design rainfall depth

Satoshi Watanabe¹, Masafumi Yamada², Shiori Abe³ and Misako Hatono⁴

¹*School of Engineering, the University of Tokyo, Japan*

²*Disaster Prevention Research Institute, Kyoto University, Japan*

³*Mitsui Consultants, co., Ltd., Japan*

⁴*Graduate School of Advanced Science and Engineering, Hiroshima University, Japan*

Abstract:

Design rainfall depth, which is a fundamental index used in river planning, was estimated by rainfall obtained from super-ensemble simulations with bias correction, and the future change under 4 degree warming was projected. The modifications of existing bias correction methods were proposed to resolve the issue of overfitting and gap in size between reference and super-ensemble simulation data. A bias correction approach considering the bias between the historical experiment, the reference data, and the change between the historical and future experiments separately was defined as two-pass bias correction. The two-pass bias correction was performed with a moving window method that calculated moving average for time period and rank-order statistics. The result indicated that the approach proposed in this study estimates the design rainfall depth with a small error compared to that calculated without the moving window. The moving window method effectively resolves the issue of overfitting. The projection indicated that the range of projection among sea-surface temperature (SST) patterns is equivalent to 25% of the design rainfall depth for most basins and 60% for certain specific basins. The results indicate the importance of the appropriate bias correction and the consideration of range among the SST patterns for super-ensemble simulation data.

KEYWORDS design rainfall; bias correction; d4PDF; ensemble climate simulations; moving window

INTRODUCTION

Understanding the characteristics of extreme rainfall under climate change is essential for adapting to the impacts of floods. It is well known that understanding the uncertainty of a projection is critical in climate change impact assessment studies. Multiple ensemble experiments have been found to contribute towards understanding the uncertainty of projections, and have, thus, warranted the widespread adoption of ensemble climate experiments. Downscaling and bias correction are also essential for the assessment of climate change impacts, particularly for

assessments across specific spatial scales.

Design rainfall depth is fundamental for river planning. It is established in all class A river basins in Japan, which are 109 important river basins defined by the law in Japan, and plays a vital role in flood control. Providing knowledge about the design rainfall depth estimated from climate simulations can contribute to the management of flood for climate change adaptation. Multiple ensemble simulation is essential for accurate estimation as the design rainfall is a case of extreme rainfall. The reproducibility of annual maximum basin-averaged rainfall, estimated from multiple ensemble simulations, has been evaluated in all class A river basins in Japan (Tanaka *et al.*, 2019).

Recently, super ensemble experiments that comprise over 1000 years of output have been conducted, and several databases have been published with the results of these experiments. An example of such a database is the database for policy decision-making for future climate change (d4PDF) (Mizuta *et al.*, 2017; Fujita *et al.*, 2019). The d4PDF provides a regional downscaling simulation that focuses on Japan (d4PDF-RCM). It is assumed that the impact assessments of climate change in various fields, including flood management in Japan, considering uncertainty, are conducted using the d4PDF-RCM (Hoshino and Yamada, 2018; Tanaka *et al.*, 2018).

Correcting the persistent bias in climate simulations is an inevitable process in the impact assessment of climate change (Iizumi *et al.*, 2017; Watanabe *et al.*, 2012). This is also true for super ensemble simulations. However, to the best of our knowledge, it is evident that there has been no conclusion in the discussions on the method that contributes to overall climate change uncertainty.

The objectives of this study are: (1) to compare the estimation of design rainfall depth from the d4PDF-RCM with multiple bias corrections and (2) to project the rainfall depth corresponding to the level of return of the design rainfall, considering 4-degree global warming. The novelty of the study is highlighted as it presents a comparison of the applied bias correction method considering multiple assumptions using rainfall obtained from the super ensemble simulations. Moreover, the study also compares and projects design rainfall, which is important for flood management.

Correspondence to: Satoshi Watanabe, School of Engineering, the University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan. E-mail: stswata@hydra.t.u-tokyo.ac.jp

Received 1 June, 2020
Accepted 11 August, 2020
Published online 11 September, 2020

METHOD

Dataset

The rainfall obtained from the d4PDF-RCM was bias-corrected with the in-situ observation dataset of the automated meteorological data acquisition system (AMeDAS) operated by the Japan Meteorological Agency. The bias correction was conducted using historical simulations (HPB) and applied to the projection under 4K simulations with 6 sea surface temperature (SST) patterns (SST1–SST6). While the d4PDF-RCM includes global and regional downscaled products, only the regional product around Japan was used in this study. Only the latter 30 years (1981–2010) of the d4PDF-RCM was used considering the availability of AMeDAS. We used 1500 years (30 years $\times$ 50 ensembles) of precipitation as the total. Not all data of AMeDAS include precipitation data that is sufficient for bias correction owing to issues of missing data. As bias correction is the process used to compare statistical characteristics, it is important to use data over an adequate timeframe for comparison. Therefore, in this study, we used the AMeDAS precipitation data from 1176 observation sites, of which the data at one observation point included the hourly precipitation data collected over 20 years. This represents roughly one site every 20 km.

Design rainfall

Design rainfall is the defined level of rainfall that river planning has been conducted, in order to prevent flooding below this level. In Japan, design rainfall for all class A rivers has been defined considering historical observations of precipitation. The design rainfall for each river basin is defined based on the design return period, which is set depending on the specific conditions of each river. Design rainfall provides fundamental information for planning river management. In Japan, the Ministry of Land, Infrastructure, Transportation, and Tourism defines the value of the rainfall in the design return period in each class A river basin by considering the hydro-meteorological and social characteristics of each basin. As the design rainfall was estimated in the mid-1900s for most of the rivers, these were conducted with limited observation data using extreme statistical values.

The rainfall depth corresponding to the level of return of the design rainfall was estimated from the d4PDF-RCM with bias correction using the AMeDAS data. The estimation was performed with rank-order statistics, indicating that a statistical distribution was not used. It was assumed that the size of the d4PDF-RCM, which is 1500 years for historical simulations, was large enough for the use of rank-order statistics. We evaluated the estimation results for 21 of all 109 class A river basins (Figure 1) considering the following conditions:

1. The area of the basin is more than 1000 km², and more than 90% of the basin is covered by the area where design rainfall is defined;
2. Three or more AMeDAS observation points, adequate for bias correction, are located in the basin.

These limitations are necessary to avoid comparisons between the actual and estimated design rainfall depth using the bias-corrected d4PDF-RCM under different con-

Reference points [mm/DT]

Figure 1. Reference point in the 21 target river basin with the design rainfall depth. The combination of the duration and return periods of each design rainfall has been denoted using different symbols. The unit of figure is mm/DT, whereby precipitation (mm) is divided by design rainfall duration time (DT)

ditions. Bias correction was performed at each AMeDAS observation point, and the Thiessen method was used to estimate the design rainfall depth. Considering the spatial resolution of the d4PDF-RCM, which is nearly 20 km, the estimation of rainfall depth in a small river basin or a basin with limited AMeDAS observation points generates unreasonable results.

It should be noted that our estimation of the design rainfall depth was not consistent with the actual design rainfall depth, even if the bias correction was perfect. This is mainly because of the difference in the observation dataset and availability of the actual method for estimation, which is not fully explained.

Bias correction

Bias correction method can be classified into two types in terms of the definition of bias. The one assumes that bias is constant with the value of the target variable, and the other assumes that bias is constant with occurrence frequency of the value. Because of this assumption, the change of the value from historical to future having the same occurrence frequency is generally not consistent with before and after bias correction in the former definition. It is, on the other hand, consistent in the latter definition, and therefore the bias correction approach with the latter definition is referred to as trend preserving type. The difference and advantage of each method has been discussed in many studies (e.g. Cannon *et al.*, 2015; Watanabe *et al.*, 2014). In line with recent studies (e.g. Lange 2019; Gomez-Garcia *et al.*, 2019), trend preserving type bias correction was used in this study. In the trend preserving type, the probability distribution of pseudo-observations under climate change is

constructed during the process of bias correction. The significant difference between the trend preserving type is the method used to construct this probability distribution of pseudo-observations.

There are three issues in constructing the pseudo observations, namely, temporal resolution, size of the temporal bin, and transfer function. We reviewed each option and compared the bias-corrected estimated design rainfall depth with their combinations.

There are two strategies for bias correction from the perspective of temporal resolution, namely aggregation and disaggregation. The aggregation strategy corrects the fine-scale (e.g. hourly) first, after which they are aggregated to generate a coarse resolution (e.g. daily). Contrastingly, the disaggregate strategy corrects the coarse resolution first and then disaggregates them to generate a fine resolution. Haerter *et al.* (2011) discussed this issue and emphasized the importance of setting the appropriate time resolution based on the objective of bias correction. Generally, the target scale can reflect the characteristics of the reference data well.

The size of the temporal bin is another factor related to time to be considered when performing bias correction. The most popular ones are monthly bias corrections. Seasonal or yearly bias corrections are conducted depending on the target area and the objective.

The transfer function that combines three probability distributions, which are reference data, historical simulation, and future simulation, has many options with minor differences. However, these are classified as parametric or non-parametric. It is necessary to select the statistics to be corrected or size of bins to match the ranks.

The disaggregation approach of time resolution was adopted in this study considering the objective of the study, in which aggregation of hourly data is more important than the hourly data itself. The discussion about the second and third points is presented in the following section, with a proposal for the modification method.

Two-pass bias correction with dual moving window

We reconstruct the bias correction methods considering that the size of the dataset is significantly different between model output and reference data. Figure 2 shows a schematic of our method. As most of the existing methods are appropriate for experiments with a single ensemble member, we proposed the use of a moving window technique. We developed a method that considers the bias between the historical simulations, the reference data, and the change between the historical and future simulations separately, to overcome the gap in size. In this two-pass bias correction, a pseudo-observation was developed by applying the bias and the change to the historical simulations separately. Existing methods apply these to a projection simultaneously, which is the difference between the two-pass approach and existing methods. Two-pass bias correction with dual moving window considers the difference of size between reference and large ensemble experiment data, and enables robust correction against the gap in the size.

In this study, the moving window method was applied to meet two objectives. As it has been used in some bias correction studies to smooth the gap between temporal bins

Figure 2. A schematic of dual moving window and two-pass bias correction

(Smitha *et al.*, 2018), we applied the moving window not only on the temporal gap but also the rank gap. This dual moving window approach is also advantageous when applied to large experiments.

We adopted a 25-day moving window with a five-day slide for the temporal moving window. For example, for the bias correction in the first bin, which was from January 1 to 5, the data from December 22 to January 15 were used for the estimation. For the second bin, which was from January 6 to 10, the data from December 27 to January 20 were used. This temporal moving window technique overcomes the problem in the previous method involving the gap between the bins (the gap between the end of the previous month and the start of the current month).

We applied the rank moving window technique to match the ranks of two different sets of data to improve the robustness. As we applied a non-parametric approach, the gap between the ranks that can generate artifacts is an issue to be solved. To identify the appropriate window size, we applied multiple window size settings for the reference and the historical simulations. In the reference data, the size of the window was 3, 5, 10, and 25. The size for the historical simulations was determined by the ratio of the window size to the total reference dataset depending on the available

AMeDAS dataset. Meanwhile, the window size for comparing the historical and future change was fixed to 0.1% of the entire historical or future data.

The basic strategy mentioned above is similar to that applied by Lange (2019). The major differences are, 1) the two-pass matching; and 2) the non-parametric approach using the moving window technique. We combined all the historical ensemble simulations into one distribution and applied the moving-window technique based on the discussion presented by Chen *et al.* (2019). Additionally, as the aggregated precipitation over several days is more important for the design rainfall depth, we adopted the daily scale bias correction.

RESULTS AND DISCUSSIONS

Estimation of design rainfall depth

The results indicated that the moving window bias correction reduced the bias in the design rainfall depth estimated using the d4PDF-RCM. Figure 3 illustrates the characteristics of the results obtained using each method. The window using 3, 5, and 10 reference data has shown to reduce the error compared to the non-corrected (NC) data, without using the moving window (ALL) and the correction of only the mean value (MEAN). Although the window using three reference data is better than the others in

Figure 3. Difference between estimated and actual design rainfall depths. The ratio of the difference between the estimated and actual values is divided by the design rainfall depth and is shown in each river (a). The number in figure (a) corresponds to that in Figure 1. The number of basins belonging to each error range is shown as a cumulative bar chart (b). NC, MEAN, and ALL denote the results of estimation using the d4PDF-RCM without bias correction, with bias correction only for mean value, and with bias correction using all reference values independently, which corresponds to that of the bias correction without moving window for rank matching. Window 3 to 25 denote the results with a bias correction using moving window with the range from 3 to 25, respectively

many basins, the appropriate window size depends on the basins. However, the window using 25 reference data is not better than other window sizes, suggesting that a larger window tends to fail to estimate extreme values. It should be noted that the number of available AMeDAS datasets is different among basins, which is one of the possible factors determining the appropriate window size.

The result of using the MEAN strategy, which corrects only the mean value, shows a larger bias ratio than that using NC, indicating that the bias correction only for mean value increases the error for most basins while reducing it for basins 1 and 2. Similarly, window bias correction using 25 reference data reduced the error better than that using fewer references for basin 1 and 2. However, the error remained large in some basins. These results suggest two possibilities.

1. The d4PDF-RCM simulates extreme precipitation well if only the mean bias is removed.
2. The quality of the reference data is poor owing to reasons such as the presence of many missing values, and the inadequate number of observation sites for the estimation of design rainfall depth.

The comparison of results between the bias correction with and without a moving window suggests the issue of overfitting and highlights the effectiveness of the moving window in resolving it. The result of bias correction using the moving window always reduced more error than without using it except in basins 1 and 2. This indicates that the use of the traditional bias correction, such as quantile-based mapping with non-parametric matching without a moving window, is not appropriate for super ensemble experiments.

Projection of design level rainfall depth

The results of the projections considering 4 degrees of global warming indicate that the range of projection among the SST patterns is large. Figure 4 illustrates the future extreme rainfall depth in the design return period. The ratios of the increase are found to be large in the northern and southern regions. However, the ratio of increase in other regions is not clear.

The range of projection among the SST patterns was clearer compared to the design rainfall depth and mean future change (Figure 5). The range of the design rainfall depth among SST patterns was equivalent to 25% of the design rainfall depth for most basins and 60% for certain basins. It was equivalent to or greater than the mean future change.

The issue of the projection range among the SST patterns is the same or more important compared to the issues encountered using bias correction methods for the projection of future extreme rainfall. Figure 6 shows that the range of projection among the SST patterns is more than the difference in the bias correction with and without a moving window.

SUMMARY AND CONCLUSIONS

This study projected the extreme rainfall depth in the design return period by considering the estimation of the design rainfall depth using super ensemble simulations with bias correction. A comparison of the bias correction

Figure 4. Ratio of change in the extreme rainfall in the design return period under 4-degree warming with 6 different SST patterns (SST1–SST6)

method, with and without a moving window, and the number of reference data in the window was conducted. The results indicate that the methods used in previous studies, in which bias is corrected using a non-parametric quantile mapping technique without a moving window, are not appropriate for super ensemble simulations due to the issue of overfitting. The projection of future extreme rainfall indicates that the range of projection among the SST patterns is large, considering the range of the mean future change and the difference between the bias correction methods. It is necessary to understand the difference in the projections among the SST patterns to understand the change in extreme rainfall under future warming conditions. The use of performance metrics (Watanabe *et al.*, 2014) may contribute to improving this issue. Additionally, the framework of this study, which highlights that all ensembles in historical simulations are combined, is a potential for consideration in future studies.

ACKNOWLEDGMENTS

This study was supported by the Academic-Industry Collaboration Program “Water cycle data integrator”; Grants-in-Aid for Scientific Research (18K13834, 18H01543, 18KK0117, 18J11683, 18J00585) from Japan Society for the Promotion of Science (JSPS); the Program for Integrated Research Program for Advancing Climate Models

Figure 5. Range of projection among the SST patterns compared with design rainfall depth (a) and projection means future change (b). The number corresponds with the basin shown in Figure 1. The dotted line denotes the ratio between the compared values. The number of each plot and unit correspond to those in Figure 1

Figure 6. Range of projection among the SST patterns compared with the difference of bias correction with and without the moving window using the window 3. The unit corresponds to those in Figure 1

(JPMXD0717935498) and the Data Integration and Analysis System (DIAS) project from the Ministry of Education, Culture, Sports, Science, and Technology-Japan (MEXT); The d4PDF was produced under the SOUSEI program and provided by the DIAS.

REFERENCES

- Cannon AJ, Sobie SR, Murdock TQ. 2015. Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *Journal of Climate* **28**: 6938–6959. DOI: 10.1175/JCLI-D-14-00754.1.
- Chen J, Brissette FP, Zhang XJ, Chen H, Guo S, Zhao Y. 2019. Bias correcting climate model multi-member ensembles to assess climate change impacts on hydrology. *Climatic Change* **153**: 361–377. DOI: 10.1007/s10584-019-02393-x.
- Fujita M, Mizuta R, Ishii M, Endo H, Sato T, Okada Y, Kawazoe S, Sugimoto S, Ishihara K, Watanabe S. 2019. Precipitation changes in a climate with 2-K surface warming from large ensemble simulations using 60-km global and 20-km regional atmospheric models. *Geophysical Research Letters* **46**: 435–442. DOI: 10.1029/2018GL079885.
- Gomez-Garcia M, Matsumura A, Ogawada D, Takahashi K. 2019. Time scale decomposition of climate and correction of variability using synthetic samples of stable distributions. *Water Resources Research* **55**: 3632–3658. DOI: 10.1029/2018WR023053.
- Haerter JO, Hagemann S, Moseley C, Piani C. 2011. Climate model bias correction and the role of timescales. *Hydrology and Earth System Sciences* **15**: 1065–1079. DOI: 10.5194/hess-15-1065-2011.
- Hoshino T, Yamada JT. 2018. Analysis of annual maximum precipitation over first-class river domains in Japan using a large-ensemble dataset (d4PDF). *Journal of Japan Society of Civil Engineering, Ser. B1 (Hydraulic Engineering)* **74**: I_187–I_192 (in Japanese with English summary). DOI: 10.2208/jscejhe.74.I_187.
- Iizumi T, Takikawa H, Hirabayashi Y, Hanasaki N, Nishimori M. 2017. Contributions of different bias-correction methods and reference meteorological forcing data sets to uncertainty in projected temperature and precipitation extremes. *Journal of Geophysical Research Atmosphere* **122**: 7800–7819. DOI: 10.1002/2017JD026613.
- Lange S. 2019. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development* **12**: 3055–3070. DOI: 10.5194/gmd-12-3055-2019.
- Mizuta R, Murata A, Ishii M, Shiogama H, Hibino K, Mori N, Arakawa O, Imada Y, Yoshida K, Aoyagi T, Kawase H, Mori M, Okada Y, Shimura T, Nagatomo T, Ikeda M, Endo H, Nosaka M, Arai M, Takahashi C, Tanaka K, Takemi T, Tachikawa Y, Temur K, Kamae Y, Watanabe M, Sasaki H, Kitoh A, Takayabu I, Nakakita E, Kimoto M. 2017. Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bulletin of the American Meteorological Society* **98**: 1383–1398. DOI: 10.1175/BAMS-D-16-0099.1.
- Smitha PS, Narasimhan B, Sudheer KP, Annamalai H. 2018. An improved bias correction method of daily rainfall data using a sliding window technique for climate change impact assessment. *Journal of Hydrology* **556**: 100–118. DOI: 10.1016/j.jhydrol.2017.11.010.
- Tanaka T, Tachikawa Y, Ichikawa Y, Yoroze K. 2018. Flood risk curve development with probabilistic rainfall modelling and large ensemble climate simulation data: a case study for the Yodo River basin. *Hydrological Research Letters* **12**: 28–33. DOI: 10.3178/hrl.12.28.
- Tanaka T, Kawai Y, Tachikawa Y. 2019. Evaluating reproducibility of maximum basin-averaged rainfall of d4PDF-RCM in all class-A rivers in Japan. *Journal of Japan Society of Civil Engineering, Ser. B1 (Hydraulic Engineering)* **75**: I_1135–I_1140 (in Japanese with English summary).
- 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: Atmospheres* **117**: D23114. DOI: 10.1029/2012JD018192.
- Watanabe S, Hirabayashi Y, Kotsuki S, Hanasaki N, Tanaka K, Mateo CMR, Kiguchi M, Ikoma E, Kanae S, Oki T. 2014. Application of performance metrics to climate models to projecting future river discharge in the Chao Phraya River basin. *Hydrological Research Letters* **8**: 33–38. DOI: 10.3178/hrl.8.33.