Sayama *et al.*
Progress in Earth and Planetary Science (2025) 12:17
https://doi.org/10.1186/s40645-025-00691-w

Progress in Earth and Planetary Science

RESEARCH ARTICLE Open Access

Parameter regionalization of large-scale distributed rainfall–runoff models using a conditional probability method

Takahiro Sayama[1*], Masafumi Yamada[1], Ayato Yamakita[2] and Yoshito Sugawara[1]

Abstract
Given the evident impact of climate change, the frequency of severe flood events has increased worldwide. For various risk-reduction measures, covering all rivers in a country or regions including small-to-medium-sized rivers, flood risk assessment and real-time forecasting based on large-domain and high-resolution distributed rainfall–runoff models are fundamental. Due to limited observed records in such small-to-medium-sized rivers, the used distributed model must be robust and physically sound with the regionalized model parameters. Specifically, rather than optimizing parameters in many independent river basins, leading to a patched parameter distribution, regionalization should reflect the spatial distribution of hydrological signatures, such as soil and geology types. However, optimizing the parameters with existing methods incurs computational costs, posing difficulties in the parameter regionalization of large-domain and high-resolution distributed runoff models. To address this challenge, we propose a parameter regionalization method based on conditional probability. The key feature of this method is that the calibration phase calculation assumes spatially uniform parameter sets within the calibrating basins, significantly reducing computational costs. However, the resulting parameter sets are spatially distributed corresponding to the region's pre-prepared soil or geological maps. It was achieved by introducing the Bayes' theorem to estimate the conditional probability of the parameter set. The proposed method was applied to the distributed rainfall–runoff–inundation (RRI) model developed for Japan with a resolution of 150 m. The model performance in the validation phase, in which the performance was evaluated with 2723 flood events at 711 gauging stations, the median Nash–Sutcliffe efficiency (NSE) being 0.87, comparable or even improved to the performance in the calibration phase (NSE = 0.83) with 525 flood events at 75 dam reservoirs. Overall, the obtained nationwide high-resolution model is robust with good performance, even in ungauged basins. Furthermore, the proposed regionalization is a simple and useful way reflecting spatially distributed hydrologic signatures in the model parameters, and it can be utilized for any distributed rainfall–runoff model.

Keywords Distributed model, Parameter regionalization, Conditional probability, Soil type, Flood, Hillslope runoff

1 Introduction

The intensification of various hydro-meteorological hazards including river floods is a growing concern amid climate change (IPCC 2022). For real-time flood predictions and climate change impact assessment, distributed hydrological models have been widely used. These models can predict streamflow discharges in both gauged and ungauged basins, allowing for spatially consistent hydrological predictions over large areas (Archfield et al.

*Correspondence:
Takahiro Sayama
sayama.takahiro.3u@kyoto-u.ac.jp
[1] Disaster Prevention Research Institute, Kyoto University, Gokasho Uji, Kyoto 611-0011, Japan
[2] Graduate School of Engineering, Kyoto University, Nishikyo-Ku, Kyoto 615-8530, Japan

2015; Merz et al. 2022). Utilizing a distributed model that explicitly considers basin characteristics, such as topography, soil type, and land use, can better represent hydrological processes than lumped models (Clark et al. 2015; Fan et al. 2019). By comprehensively treating a wide area, it becomes possible to utilize observation data so that the comparison between models and observations can also be carried out at multiple sites (Merz et al. 2022). Such large-sample hydrology (Addor et al. 2020; Gupta et al. 2014) may advance the understanding of hydrological processes and identify hydrological model structures from comparative hydrological perspectives.

Large-domain distributed models typically have grid scales ranging from 100 to 1000 m, making it impractical to express all processes using physical governing equations and determine parameters solely from field observations. Instead, these models approximate the physical behavior and introduce parameters equivalent to properties, such as hydraulic conductivity and soil layer thickness. The parameter estimation process should reflect the spatial signatures present in the basin, such as soil and geological types, land use, and vegetation, through parameter regionalization (Abdulla and Lettenmaier 1997; Carrillo et al. 2011; Schweppe et al. 2022; Young 2006). This makes constructing a more spatially consistent distributed model possible and avoid a patch-like parameter distribution, which results in incontiguous runoff responses. It also helps to clarify how soil and geology affect rainfall–runoff processes by examining the correspondence between estimated parameters and spatial signatures (Mizukami et al. 2017; Samaniego et al. 2010). Furthermore, the collective use of basin information allows hydrological projections in ungauged basins and more robust and stable parameters to be estimated without overfitting in a particular basin.

Several approaches have been proposed to achieve parameter regionalization in large-domain hydrological models (Beven and Chappell 2021). These parameter regionalization methods can be classified based on whether spatially uniform or distributed parameters are used for the first training stages in gauged basins and whether a transfer function relates the spatial signatures and model parameters or estimates the model parameters directly.

An example of a parameter regionalization approach is the regional calibration approach developed by Hundecha and Bardossy, (2004), which calibrates the functional dependency of model parameters on catchment descriptors for all catchments. Another approach is multiscale parameter regionalization (MPR) proposed by Samaniego et al., (2010), which uses transfer functions to convert geophysical properties at their original spatial scale into aggregate model parameters at the desired model resolution. The MPR is a type of simultaneous regionalization that operates at a finer resolution (data input level) to account for the sub-grid variability of basin predictors. It utilizes transfer functions such as pedotransfer functions to link model parameters and basin predictors through transfer functions or global parameters. The effective parameters at a coarser grid scale were obtained by upscaling the calibrated parameters using the appropriate operations. The MPR allows for spatially consistent parameter estimation and provides application-ready estimates of spatial parameter fields over a large geographical domain (Mizukami et al. 2017; Samaniego et al. 2017). Other large-scale modeling approaches include the Parameter Set Shuffling method proposed by (Merz et al. 2020), which combines local calibration and regionalization using machine learning tools without requiring a priori selection of the dominant catchment descriptors for each parameter.

The choice of the parameter regionalization method depends on factors such as the scale and amount of spatial signature information, the computational cost of the model simulation, and the desired model resolution. The selection should also consider the balance between simultaneously incorporating diverse maps and maintaining a high model resolution. For the specific case of a high-resolution model (~ 150 m) focusing on flash flood simulations, it is important to avoid overly complex parameter distributions and consider reductions in spatial dimensionality. Calibration should involve selecting uniform parameter sets within the basin and choosing parameter sets corresponding to soil and geological maps based on likelihood information. Although some parameter regionalization methods have been proposed and applied to large-domain distributed rainfall–runoff models, there is no computationally efficient method that allows to estimate spatially distributed parameters reflecting a signature map based on spatially uniform parameters in the calibration phase.

The objective of this study is to propose a regional integration method based on conditional probability for parameters corresponding to soil and geological maps. In this approach, multiple small basins were targeted, and simulations were conducted using spatially uniform parameters within the basins. The results of these simulations were then used to estimate the parameters corresponding to the soil and geology using conditional probability.

2 Conditional probability parameter regionalization (CPPR)

2.1 Outline of the proposed method

In this study, we propose a method for estimating the spatial distribution of a parameter according to the

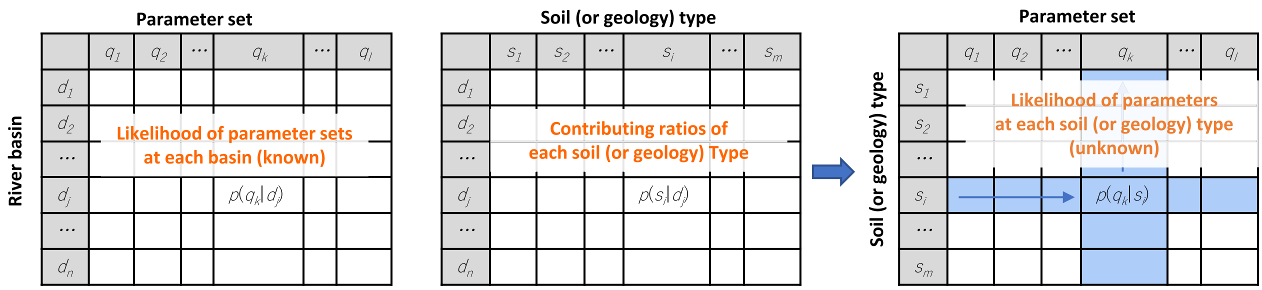

Fig. 1 The steps of the CPPR method for computing the likelihood of parameter sets for each soil or geology type

classification of soil or geological maps based on conditional probability. We call this the conditional probability parameter regionalization (CPPR) method. CPPR comprises the following two steps. The first-step involves conducting multiple simulations in the calibration catchments using spatially uniform candidate parameter sets. The second step is to compute the likelihood of each parameter set for different soil or geological types based on a conditional probability concept derived from the first-step simulation results. The main advantage of the CPPR is that, for the first-step simulation, we can conduct simulations with spatially uniform candidate parameter sets; hence, we can limit the number of first-step simulations. In the second step, the CPPR method identifies the best parameter set for each soil or geologic type among the candidate sets, so that this method enables to estimate spatially distributed parameter sets corresponding to the soil or geological map.

2.2 Steps of the CPPR method

The following are the steps of the proposed CPPR method.

1) Prepare sets of candidate parameters.
2) Perform model simulations using all the candidate parameter sets for all the calibration basins. In this step, simulations were performed using spatially uniform parameter sets within the calibration basins.
3) Compare the calculated and observed discharges and estimate the likelihood of each candidate parameter set for all calibration basins. The method used to estimate the likelihood can be determined according to the objectives of the model application. In this study, as we focused on the robustness of the model for flash flood predictions, we used the number of events satisfying the criteria, composed of the relative peak difference (RPD) and correlation coefficients (CC) between the simulated and observed hydrographs. The details are explained in Sect. 3.4 and the Supplement. The estimated likelihoods $p\left(q_k|d_j\right)$ for each calibration basin are summarized in Fig. 1, where q_k is the potential parameter set k and d_j is the examined river basin j.
4) Compute areal occupation ratios $p\left(s_i|d_j\right)$ of each soil type or geological type from a soil or geological map for all calibration basins, where s_i is the soil or geologic type i.
5) Estimate the probability indicating which candidate parameter sets should be assigned for each soil or geology type $p(q_k|s_i)$ based on the following conditional probability theorem:

$$p(q_k|s_i) = \sum_{j=1}^{n} \frac{p\left(s_i|d_j\right)p\left(d_j\right)}{p(s_i)} p\left(q_k|d_j\right) \quad (1)$$

6) Select for each soil or geology type, a parameter set with the highest $p(q_k|s_i)$.

3 Model and data used in this study

3.1 Nationwide rainfall–runoff–inundation (RRI) model

The distributed model used in this study was the rainfall–runoff–inundation (RRI) model. The RRI model, which was applied to all of Japan and had a spatial resolution of 5 s (approximately 150 m), was used in this study (hereafter referred to as the JRRI model). For the development of the JRRI model, Japan was divided into 14 regions (Figure B.1 in the supplement) so as not to straddle any river basin but to limit the size of each regional model. We used the Japan Flow Direction Map (J-FlwDir) as the basis for the distributed model (Yamazaki et al. 2018). As the original J-FlwDir was prepared with a spatial resolution of 1 s (approximately 30 m) based on digital elevation models provided by the Geographical Survey Institute of Japan, the dem and flow directions were upscaled to 5 s by maintaining the main flow directions using the algorithm proposed by Masutani et al., (2006).

The RRI model simultaneously simulates both rainfall runoff and flood inundation. It applies a two-dimensional diffusive wave model to the land and a one-dimensional diffusive wave model to the river channel. The model also computes water interaction between lands and rivers by considering river cross sections, including embankment heights, based on overtopping formulas. For the land calculations, the simulated rainfall–runoff processes were categorized into three types: vertical infiltration dominant, lateral subsurface dominant, or only lateral surface flow. The RRI model can include three runoff processes in a single model based on the parameter settings. Typically, we apply the vertical infiltration process to the gentle topography of paddy or crop fields, the lateral subsurface process for mountainous forests, and overland flow for urban areas. We used a land use map to classify it into three categories. Among the three classes, we particularly focus on the second one, rainfall–runoff processes from forest mountains, since about 70 percent of Japan is covered by forest, which is the dominant source of flood runoff and makes it difficult to regionalize the model parameters. For flat areas, including paddy, cropping, and urban areas, we estimated a parameter set for each land use and excluded it from our calibration. Note that most of the basins are occupied by mountainous forests, owing to their topographic features. For mountainous forests, the current RRI model uses the following stage-discharge relationship to represent unsaturated and saturated subsurface and overland flows (Sayama and McDonnell 2009; Tachikawa et al. 2004). The first line of Eq. (2) is to represent unsaturated subsurface flow in soil matrix. The second line is to simulate both saturated subsurface flow in soil matrix and macro pores. The third line represents saturated subsurface flow in the soil and overland flow:

$$q_x = \begin{cases} -k_m d_m \left(\frac{h}{d_m}\right)^{\beta} \frac{\partial H}{\partial x} (h \le d_m) \\ -k_a (h - d_m) \frac{\partial H}{\partial x} - k_m d_m \frac{\partial H}{\partial x} (d_m < h \le d_a) \\ -\frac{1}{n} (h - d_a)^{\frac{5}{3}} \sqrt{\left|\frac{\partial H}{\partial x}\right|} sgn\left(\frac{\partial H}{\partial x}\right) - k_a (h - d_m) \frac{\partial H}{\partial x} - k_m d_m \frac{\partial H}{\partial x} (d_a < h) \end{cases} \tag{2}$$

where q_x is the flow rate in the x-direction, h is the water depth from the interface between the bedrock and soil layer, and H is the water stage from the datum. The k_a is lateral saturated hydraulic conductivity, d_a is the soil depth multiplied by the effective porosity, d_m is the water depth equivalent to the maximum water content in the capillary pores, and n is Manning's roughness on the land. To ensure the continuity of the discharge change when $h = d_m$, the lateral hydraulic conductivity in the unsaturated zone (k_m) was computed as $k_m = k_a/\beta$. Hence, β is the parameter to be identified instead of k_m. Five parameters were used to characterize the stage-discharge relationship.

When d_m is set to zero, the model does not consider unsaturated flow and becomes a saturated subsurface and surface flow model. In this case, the model had only three parameters and did not tend to underestimate the flood discharge. For our previous model application study (Sayama et al. 2020), we used a spatially uniform parameter set without considering the unsaturated flow, whose values were $d_a = 0.471$ m, $d_m = 0$ m, $k_a = 0.1$ m/s and $n = 0.4$ m$^{-1/3}$ s called this model parameter setting as the default parameter set of the JRRI model. According to our previous study of JRRI model applications (Sayama et al. 2020) to a torrential rainfall event in the western part of Japan in July 2018 and Typhoon Hagibis in October 2019, the former case focusing on the western part of Japan showed relatively high performance even with the default parameter setting, while the latter case in the eastern part of Japan showed overestimations of peak discharges with the default parameter setting. Sayama et al. (2020) discussed a possible reason for the difference in geologic settings. The eastern part of Japan, particularly in the northern Kanto and part of Tohoku regions, has volcanic soil and geology that store more precipitation in the soil and bedrock, even during such severe storm events.

Yamada et al. (2022) incorporated 26,032 cross section data for the length of 7734.7 km, accounting for 72.9% of the total river length managed by the MLIT. Because the cross-sectional data include the height of the embankment, the JRRI model used in this study can consider their impact. For other small-to-medium-sized rivers, whose detailed cross-sectional information is not available, we represent them in a rectangular shape. Widths and depths were estimated as functions of the upstream contributing area. This empirical equation was obtained from our previous study conducted in the Chikusa River Basin, and its applicability was validated in other river basins (Yamada et al. 2020).

3.2 Selection of flood events and calibration basins

This study selected 121 dam reservoir basins managed nationwide by the MLIT and the Japan Water Agency as candidates for the calibration basins. The reasons for focusing on the major dam reservoirs as calibration basins were the comparatively high quality of the inflow data and their positions, typically in mountainous forest regions. Among the potential 121 dam basins, we selected 100 based on the availability of observed records.

The flood events used for calibration and validation were selected between June 2002 and December 2018. We set the start year as 2002 because of the availability of finer spatial resolution of the rainfall data used in this study. We used radar and gauged composite data provided by the Japan Meteorological Agency and the spatial resolution became finer (2.5 km and 1 km after 2006) compared to 5 km before 2002. The accuracy has also improved since then. As described above, the JRRI model comprises 14 regional models covering Japan. Therefore, we extracted 10 storm events for each region. To extract the 10 storm events, we checked the maximum daily inflow at 100 dams and listed 10 storm events for each dam in each region. Based on this information, we selected 10 storm events for each region by prioritizing the larger inflow events for each dam. It should be noted that because several dam basins exist in each region, not all of the selected 10 storm events resulted in major flood events in each dam basin. Therefore, although we ran the model for each region with 10 storm events, we focused only on the largest top seven flood events in terms of peak discharges at each dam basin for the following calibrations and validations.

3.3 Soil and geological data

3.3.1 Soil map

This study utilized a 1:200,000 scale National Digital Soil Map provided by the National Agriculture and Food Research Organization. This soil map is available as polygon in the Shapefile format and is distributed as open data (CC BY 4.0). Based on the latest classification criteria established in 2011 (Comprehensive soil classification system of Japan-first approximation), the map is available for cultivated areas and forests. The map contains different categories in a hierarchical manner: the soil great group, soil group, soil subgroup, and soil series group. Assuming that we could represent runoff processes with a single parameter set in each soil great group, we used 10 different soil types in the soil great group. These include man-made soils, organic soils, Podzols, Andosols, Dark Red soils, lowland soils, red-yellow soils, stagnant soils, Brown Forest soils, and Regiosols. For the following analysis, we added another type, Bare Rock, for areas without any soil type.

3.3.2 Geological map

As for the geological map, the study adopts the 1:200,000 Seamless Geological Map V2 provided by the Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology. This map includes geological information about the era, rock type, and lithology, with over 2400 legends. Here, we primarily used classifications based on rock types, which include igneous rocks, sedimentary rocks, metamorphic rocks, and accretionary complexes. Following the classification used by (Mushiake et al. 1981) in the geology runoff analysis, we divided the igneous rocks into tertiary (and before tertiary) volcanic rocks, quaternary volcanic rocks, and plutonic rocks. Consequently, we used six types in total.

3.4 Estimation of likelihood $p\left(q_k|d_j\right)$

The CPPR method requires the determination of the likelihood of each parameter set in all calibration basins, described as $p\left(q_k|d_j\right)$. We introduced different metrics to evaluate model performance, including the CC and RPD:

$$\mathrm{RPD} = \frac{Q_{p,s} - Q_{p,o}}{Q_{p,o}} \tag{3}$$

$$\mathrm{CC} = \frac{\sum_{t=1}^{T}\left(Q_o^t - \overline{Q_o}\right)\left(Q_s^t - \overline{Q_s}\right)}{\sqrt{\left(\sum_{t=1}^{T}\left(Q_o^t - \overline{Q_o}\right)^2\right)\left(\sum_{t=1}^{T}\left(Q_s^t - \overline{Q_s}\right)^2\right)}} \tag{4}$$

where Q_o^t and Q_s^t are the observed and simulated discharges at time step t, $\overline{Q_o}$ and $\overline{Q_s}$ are the temporal means, and $Q_{p,o}$ and $Q_{p,s}$ are the peak discharges of Q_o^t and Q_s^t.

Because we emphasize the robustness of the model, particularly the model's capability to reproduce different magnitudes of flood discharge with a parameter set, we evaluate $p\left(q_k|d_j\right)$ according to the number of events (N_{pass}) satisfying the following criteria:

$$|\mathrm{RPD}| < 0.15 \text{ and } \mathrm{CC} > 0.85 \tag{5}$$

For our particular objective of model application, it was essential to include RPD in the criteria, while the shape of the hydrographs could be examined using the CC. Nevertheless, we used the NSE and Kling-Gupta efficiency (KGE) for model validation, as described below. Further detailed procedure of computing $p\left(q_k|d_j\right)$ is provided in the Supplement.

3.5 Experimental settings

As described in Sect. 3.1, there are five parameters to be calibrated. Let us suppose the five parameters are discretized into six, as described later, within realistic ranges. In that case, the candidate parameter set becomes 6^5 (=15,625), which is practically very difficult or impossible to test all the candidate parameter sets in many calibration basins. However, according to our preliminary investigation, among the 15,625 candidate parameter sets, several sets resulted in nearly identical outputs. Therefore, we decided to conducted preliminary simulations in two river basins (Hiyoshi and Shimouke dam reservoir basins) focusing on two different storm events at each basin with all the 15,625 candidate parameter sets. Based on the simulated hydrographs from the two basins, we clustered these parameters using the k-means method in to 40 parameter sets. In the following part, the number of candidate parameter are limited to 40, whose parameter values were shown in Table B.1 in the Supplement.

To evaluate the proposed method, we compared the model performance for the following five cases: The first case, referred to as the "Soil Case," utilizes parameters reflecting the soil map with the CPPR method. The second case, "Geology Case," employs parameters reflecting the geology map with the CPPR method. The third case, labeled "Default Case," uses default parameters from the JRRI model. The fourth and fifth cases, "Optimized Case 1" and "Optimized Case 2," involve optimization without considering soil or geological maps.

Optimized Case 1" uses parameter sets that yield the highest evaluation metrics among the 40 sets of parameters. Although this case was constrained by the spatial uniformity inside the forest areas of each dam basin, the independent selection of parameter sets for each dam allowed flexibility. It is important to recall that the main objective of the parameter regionalization cannot be achieved by this "Optimized 1" because parameter sets cannot be decided outside of the calibrating dam basins, unlike the "Soil Case" or "Geology Case" of CPPR.

Optimized 2" selected a single parameter set that yielded the best performance for all dam basins. Performance was evaluated for 75 dams of classes A and B (see the Supplement for the details on classes A and B), comprising 525 events (75 dams and seven events). Note that the parameter set No. 12 in Table B.1 satisfied the most events (189 among the 525) satisfying the above criteria and was adopted as the "Optimized 2" parameter set. Because this case assigns a single parameter set for the entire mountainous region, applying the same parameter set (i.e., No. 12) is possible outside the calibrating dam basins. Nevertheless, this case does not consider the spatial distribution of the model parameters and therefore, cannot achieve the main objective of this study. The "Optimized Cases 1 and 2" will be used to compare the performance of the "Soil Case" and "Geology Case" of the CPPR. We used NSE and KGE, RPD, and CC as evaluation indices to evaluate model performance.

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{T} \left(Q_s^t - Q_o^t\right)^2}{\sum_{t=1}^{T} \left(Q_o^t - \overline{Q_o}\right)^2} \tag{6}$$

$$\mathrm{KGE} = 1 - \sqrt{(CC-1)^2 + (\gamma-1)^2 + (\alpha-1)^2} \tag{7}$$

$$\alpha = \frac{\overline{Q_s} / \sqrt{\frac{1}{T}\sum_{t=1}^{T} \left(Q_s^t - \overline{Q_s}\right)^2}}{\overline{Q_o} / \sqrt{\frac{1}{T}\sum_{t=1}^{T} \left(Q_o^t - \overline{Q_o}\right)^2}} \tag{8}$$

$$\gamma = \frac{\overline{Q_s}}{\overline{Q_o}} \tag{9}$$

3.6 Method for verification at many observation points in Japan

In addition to assessing the model performance at the 121 dam reservoir basins whose observed inflow was used to calibrate the model, we validated the model at other gauging stations (a total of 711 stations), which are not used for the calibration. Among the storm events measured at all gauging points, we selected 2723 events. We selected flood events for validation based on the following criteria: The first criterion ensures the station's reliability; therefore, the station must have a record of more than 10 years. The second criterion ensures the magnitude of flood events; therefore, the peak discharge during a flood event should be greater than the national average annual maximum discharge. To estimate the national average annual maximum discharge for the whole country considering basin areas, the following equation approximated from the Creager curve was used:

$$Q = CA^{A^{-0.05}} \tag{10}$$

where Q is the annual mean maximum discharge (m^3/s), A is the basin area (km^2), and C is the coefficient estimated from the observed records. The least-squares method was used for observation points where the annual maximum discharge was recorded for over 10 years. Consequently, we obtained the parameter C = 6.41 for all of Japan. This study uses C^*, defined as the ratio of peak discharge of interest against the standard annual peak discharge Q estimated by (10) to quantify the magnitude of the flood event for different river basin sizes.

4 Results and discussion

4.1 Calibration at dam reservoir basins

The box plot (Fig. 2) and the cumulative distributions (Fig. 3) show the distributions of the NSE, RPD, KGE, and CC with different parameter settings. The cases include CPPR based on soil and geologic maps ("Soil" and "Geology" cases), the default parameter settings ("Default"), and two different optimization cases ("Optimized 1" and "Optimized 2"). Both "Soil" and "Geology" cases by the CPPR method outperform the "Default" and "Optimized 2" cases, both of which assume uniform parameter set in mountainous regions. The median NSE values for the "Soil" and "Geology" cases were 0.83 and 0.80, respectively, surpassing the values of 0.70 for the "Default" case and equivalent to 0.82 for the "Optimized 2" case. Furthermore, the lower quartile values in the box plot

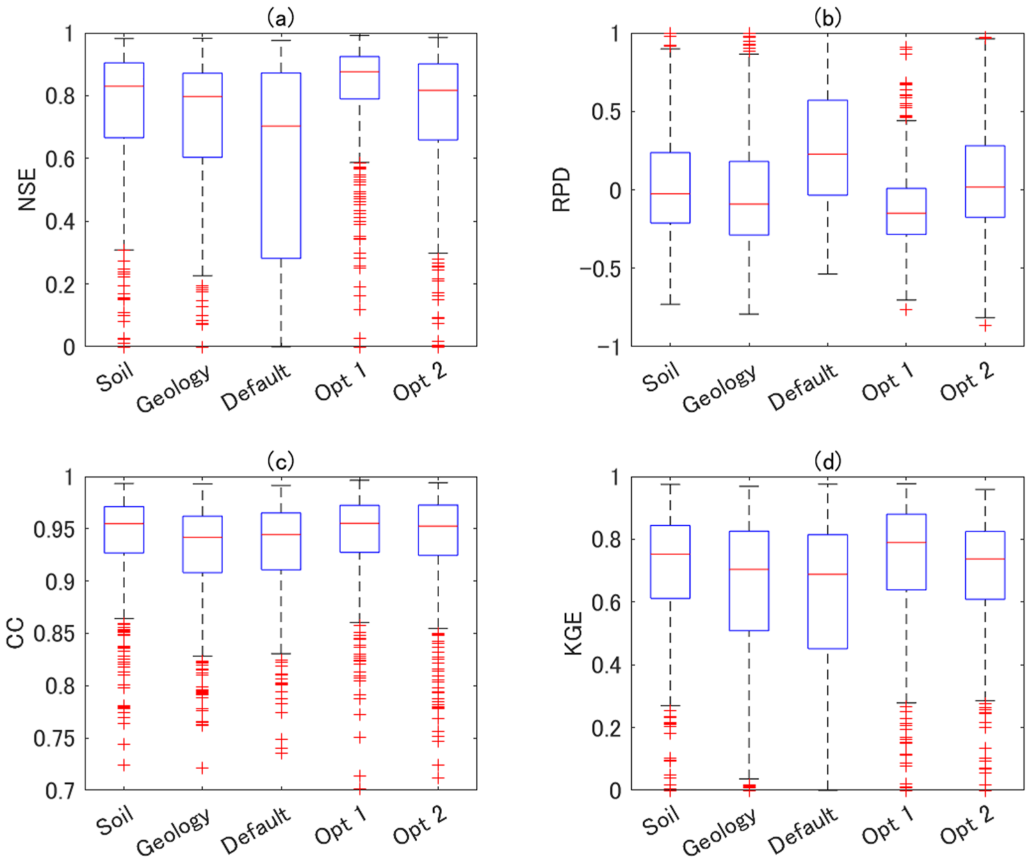

Fig. 2 Box plots comparing the different parameter settings with four evaluation indices (**a** NSE, **b** RPD, **c** CC, **d** KGE) at 75 dam reservoir basins (total 525 events). Soil: parameter regionalization by CPPR with a soil map, Geology: parameter regionalization by CPPR with a geology map, Default: spatially uniform default parameter setting, Opt 1: the best parameter set for each dam reservoir basin, Opt 2: the best parameter set applied uniformly for the entire Japan

highlight the superior accuracy of the soil case compared to the accuracy of the geological case.

In the "Soil" case context, the median NSE value of 0.83 closely aligns with the NSE value of 0.88 achieved by the "Optimized 1" case. The RPD metric evaluation reveals an underestimation bias (RPD = −0.15) in the "Optimized 1" case. Conversely, employing soil distribution by the CPPR results in a smaller bias, with a median RPD of − 0.05. The implementation of regionalization based on the soil map indicates that 50% of flood events fall within the range of ± 20% concerning the relative peak error. Among the two cases by CPPR, the "Soil" case exhibits superior results to the "Geology" case, emphasizing the efficacy of considering soil distribution. Consequently, parameters reflecting soil maps are considered superior alternatives to "Default" or "Optimized 2" parameter sets for the nationwide RRI model.

4.2 Validation at streamflow gaging stations

This section presents the validation results, focusing specifically on the soil case and the comparison with the "Optimized 2" case. The total number of events tested was 2723. The median NSE was 0.87, which was also good compared to the dam basin validation results shown in Fig. 2. Figures 4 and 5 show the box plots and cumulative distributions of NSE, RPD, KGE, and CC based on the parameter set obtained by the CPPR with the soil map (soil) and by the "Optimized 2" case. The results suggest that all indices are better or equivalent by the "soil" case compared to the "Optimized 2" case in the validation. Overall, the model performance was robust, with the NSE, RPD, KGE, and CC medians at 2723 sites being 0.87, −0.071, 0.74, and 0.96, respectively by the "soil" case.

4.3 Effects of river basin sizes

Figure 6 illustrates the variation in accuracy owing to differences in basin area. The smaller basins tended to have lower RPD, whereas the larger basins tended to have higher RPD. Specifically, in the three categories with basin areas less than 1000 km^2, the median RPD was below zero, whereas in the two categories with areas

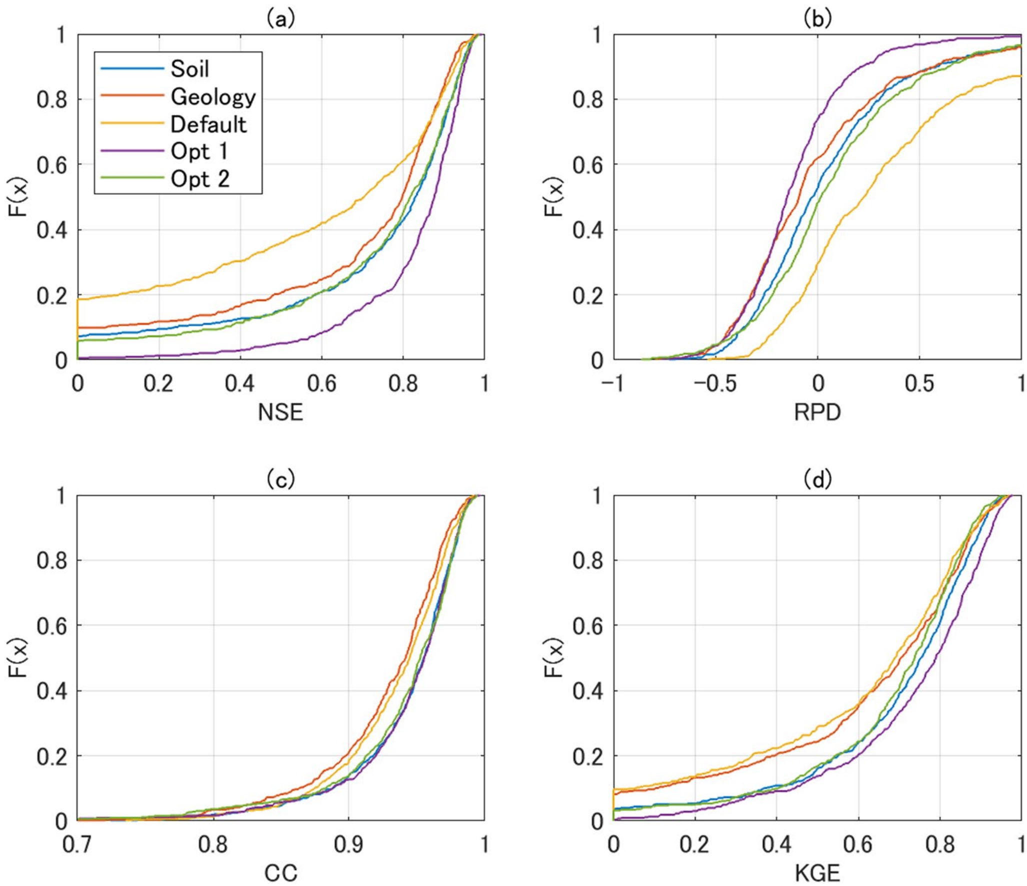

Fig. 3 CDFs compare the different parameter settings with four evaluation indices (**a** NSE, **b** RPD, **c** CC, **d** KGE) at 75 dam reservoir basins (total 525 events). The explanations of the parameter settings for Soil, Geology, Default, Opt 1, and 2 are the same as those in Fig. 1

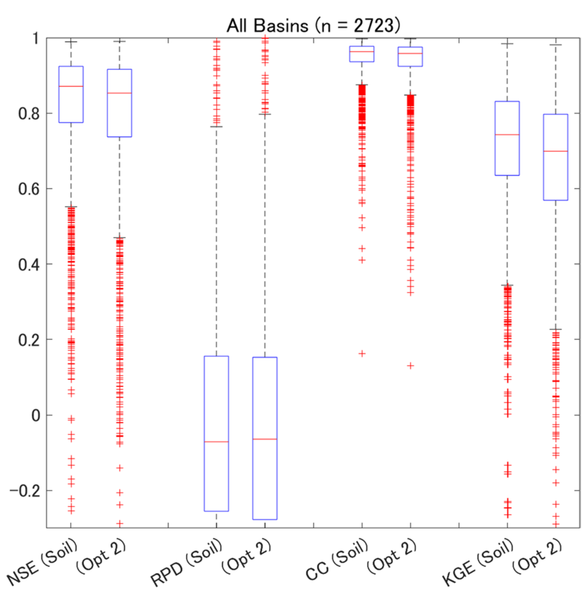

Fig. 4 Box plots of the evaluation indices at 711 stations (total 2723 events) by CPPR with a soil map and by Opt 2

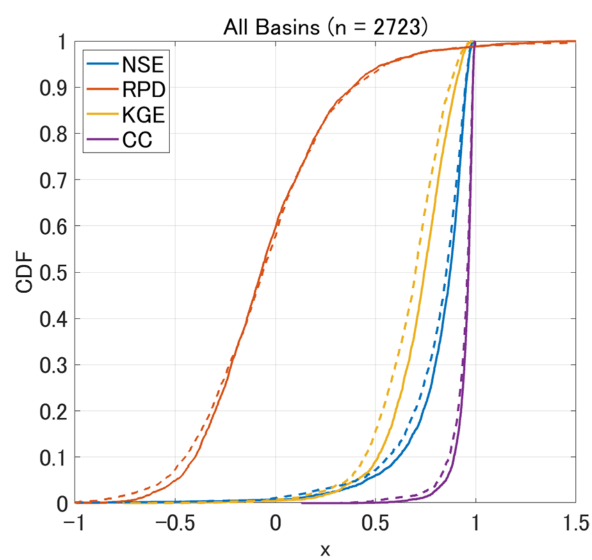

Fig. 5 CDFs of the evaluation indices at 711 stations (total 2723 events) by CPPR with a soil map (solid line) and by Opt 2 (dashed line)

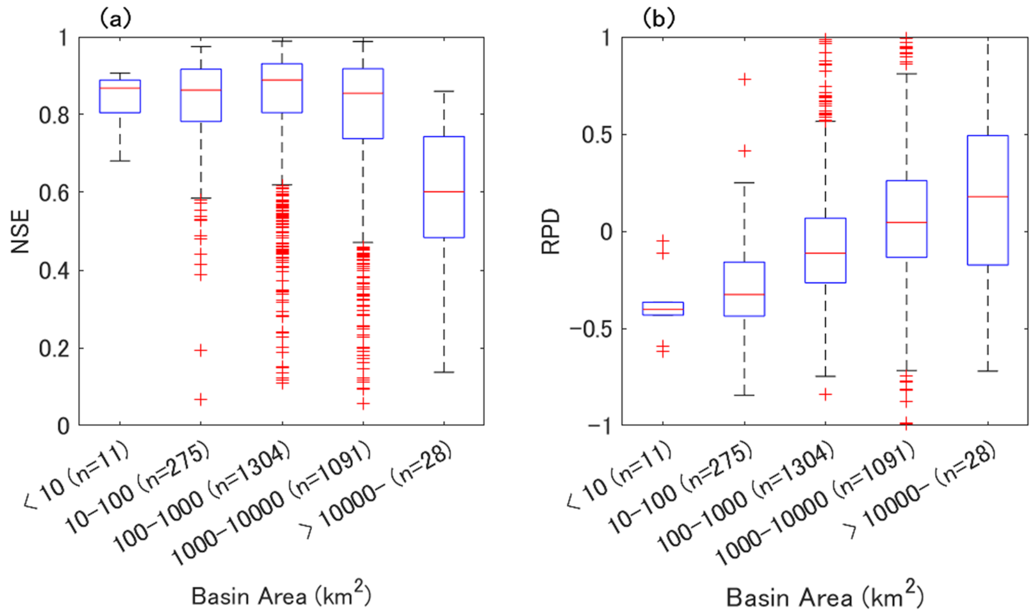

Fig. 6 Box plots of the evaluation indices with **a** NSE and **b** RPD at 711 stations (total 2723 events) comparing different basin areas

greater than 1000 km^2, the median RPD became positive. Although the sample size was limited to 11 in basins with areas below 10 km^2, there were no events where the peak discharge was overestimated, and all events resulted in underestimation. One of the possible reasons of the underestimations of peak discharges in small river basins is associated to the model structure and the settings. The RRI model has a setting parameter called river threshold, which defines how many upstream contributing grid-cells are needed to start the river channel in the model. In this study, we set 50 grid-cells, equivalent to approximately 1 km^2 of the upstream contributing area. For small river basins (i.e., 10 ~ 100 km^2), the drainage density becomes lower compared to that for large river basins. Our investigations of small river basins revealed the importance of adjusting the starting position of the channel to avoid the underestimation bias in small river basins. Meanwhile, we have to compromise with the reasonable threshold to avoid the increase of computational costs for more river calculations.

In contrast, it was observed that the median NSE drops to 0.60 in basins larger than 10,000 km^2. These points are limited to the lower part of the Tone and Shinano River basins, primarily affected by timing discrepancies resulting from the inadequate reflection of flood retention effects inside the main river channels. A dynamic wave model, rather than the diffusive wave used in the RRI model, may be essential for flood tracking in major river channels. Overall, the reproducibility of this model for the entire country was particularly favorable in medium-sized basins. Specifically, for events ranging from 100 km^2 to 1000 km^2, the median NSE was 0.89, and the first quartile showed a value of 0.81.

4.4 Effects of flood event magnitudes

The model's predictive accuracy varies depending on the scale of flood events. In this context, the nationwide average annual maximum flood discharge corresponding to the basin area was estimated using Eq. (10), and the peak flow of each discharge was normalized relative to this result. In other words, events with higher C^* values in Fig. 7 indicate larger-scale events, and an approximate C^* equal to 1 suggests an event of approximately the annual maximum discharge size. According to the evaluation using NSE, as the event size increased, the relative accuracy improved, particularly in the smallest class of C^*: 1–1.5, where the average NSE was the lowest. The RRI model is physically based on a simple structure and exhibits a higher reproducibility for larger floods. For events with C^* greater than 1.5, the median NSE was 0.85, and there was no significant decrease in accuracy for floods larger than this scale. Regarding RPD, although there was a slight tendency toward overall underestimation, there was no noticeable difference in the results based on the scale of discharge events.

4.5 Differences among regions

The results of evaluating the accuracy of the model for the 14 regions across Japan, with the country divided into these regions, are shown in Fig. 8. Note that for

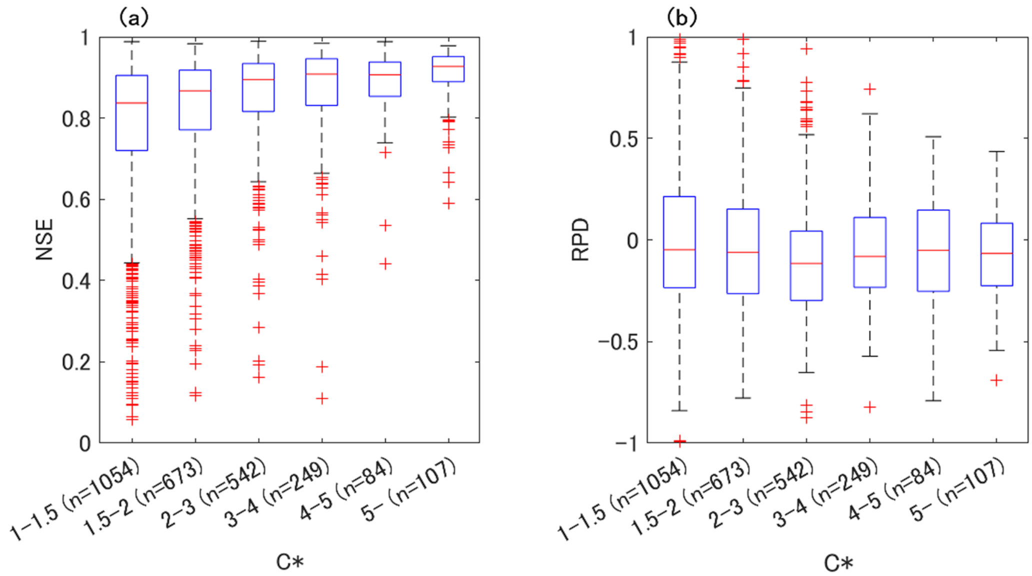

Fig. 7 Box plots of the evaluation indices with **a** NSE and **b** RPD at 711 stations (total 2723 events) comparing different event sizes: larger *C** represents the scale of flood peak discharges, where *C** = 1 is considered annual maximum flood discharges considering the basin area

the southern island regions of Japan, the results are not shown as no observational data were meeting the criteria for our model validation. The spatial distribution maps of the reproducibility indices are shown in Fig. 9. These maps illustrate the average values of each index for events at each observation point, represented by a continuous color map.

Significant regional variations in accuracy were observed. NSE generally yielded favorable results in western Japan, whereas regions such as Hokkaido, Hokuriku, and Okinawa showed room for improvement. Particularly in Okinawa, there were many events with small CC values or significantly underestimated events, with instances in which the shape of the hydrographs did not correspond to the observations. Severe floods can occur in the Okinawa region, where the basin area is relatively small, particularly when directly affected by typhoons. The mismatch in the shapes of the hydrographs may be associated with the spatial and temporal distributions and the accuracy of the radar

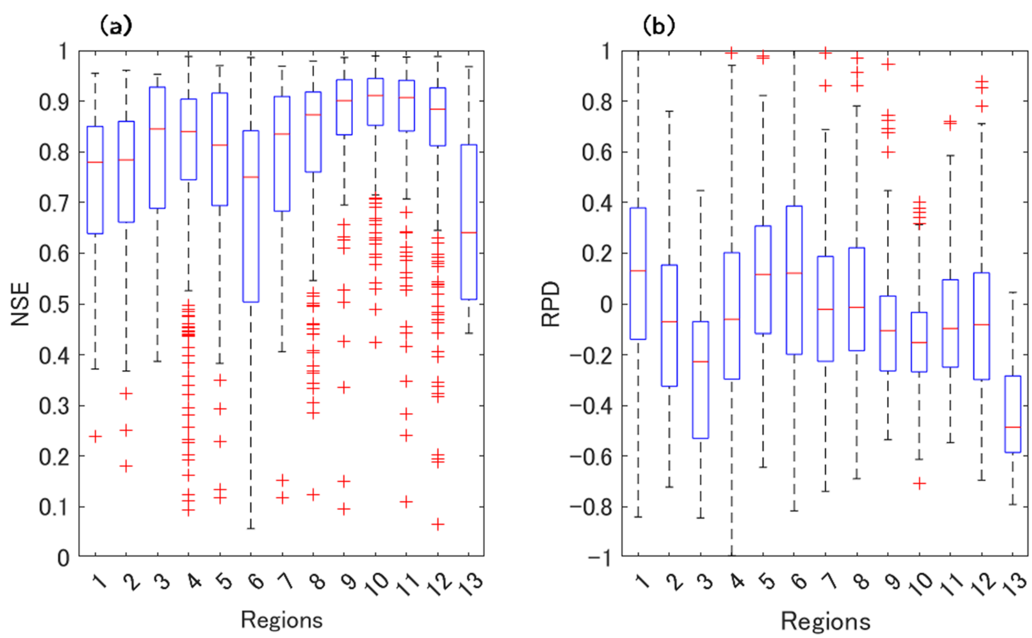

Fig. 8 Box plots of the evaluation indices with **a** NSE and **b** RPD at 711 stations (total 2723 events) comparing different regions

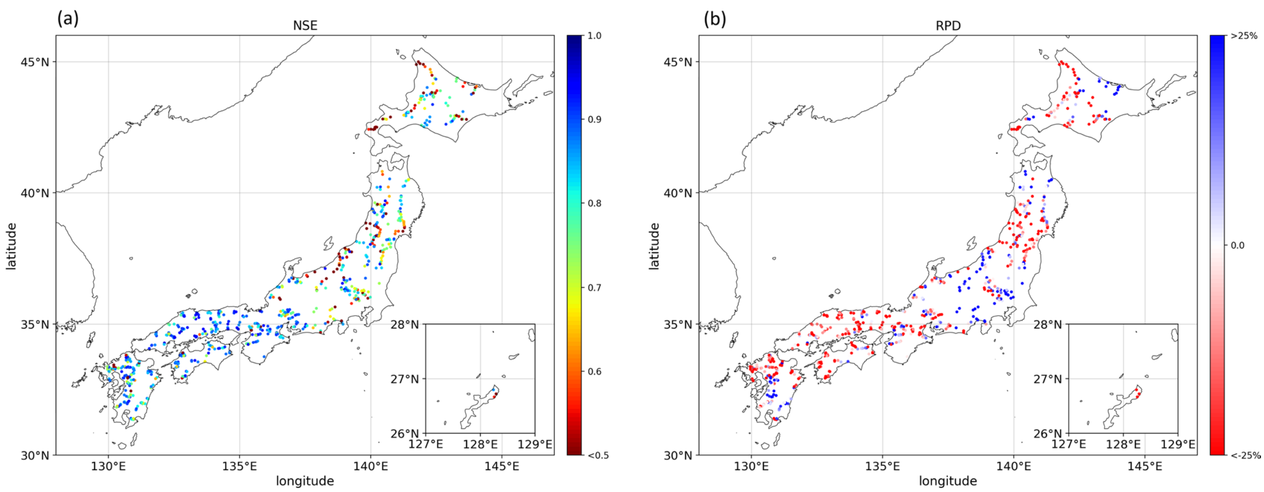

Fig. 9 Spatial distributions of average **a** NSE and **b** RPD values by the CPPR with a soil map at 711 stations (total 2723 events)

and gauged composite data used in this study. Among the regions where the NSE showed good values, the RPD was slightly underestimated in the Chugoku, Shikoku and Kyushu regions.

In contrast, the median RPD in Chubu and Kinki was close to zero, with RPD values of 0.022 and 0.001, respectively, indicating small biases. The Hokkaido and Tohoku regions generally have fewer floods than western Japan. In addition, there are volcanic regions in northeastern Kanto, including Hokkaido and Tohoku. The current model may not reproduce the rainfall–runoff process well, such as storage effects in deeper parts of the bedrock, which may contribute to an overall decrease in accuracy in these regions.

4.6 On the selected parameter sets for different soil types

Table B.2 in the Supplement shows the selected parameter set by the CPPR method for each soil type. The smaller IDs of the parameter sets indicate quicker runoff responses leading to larger peak flow, while the larger IDs result in slower runoff responses leading to smaller peak flow. Among the 11 soil types, the brown forest soil, Andosols and Podosols are the three dominant soil types, where parameter sets No. 22, No. 27 and No. 8 were assigned, respectively. To understand the sensitivity of the parameter sets, we applied the three identified parameter sets to simulate a same storm event in the Katsura river basin (blue lines in Fig. 10). Compared to the 40 possible parameters (gray lines in Fig. 10), the one by No. 22 shows average behavior, while the ones by No. 27 and No. 8 showed about 20% smaller or larger of the simulated peak discharges. Note that the figure shows the simulated results also by the other possible parameters that were identified as the 2nd, 3rd and 4th order of the parameter sets for each soil type (green lines in Fig. 10). They behave also similar to the best parameter set (blue lines) for each soil type. Such difference is reflected by incorporating different soil types in the model.

4.7 Limitations of this study

The results of this analysis indicated that the accuracy of the runoff calculations was slightly higher when using the soil map than when using the geological map. This can be attributed to the JRRI model employed in this study, which predominantly represents runoff processes in the soil layer. Geological information may become more crucial when we add deep ground water module to represent bedrock groundwater. Adding this component is necessary especially for river basins where no single parameter set can reproduce observed hydrographs. More specifically, to improve the simulation for the Class C river basins (see the detail in the Supplement), the inclusion of the deep groundwater, accordingly the geologic information, will be more important. Therefore, although the CPPR with the soil map exhibited slightly better performance than the geology map, it is challenging to conclusively assert that the former information is more critical for regionalization.

Another limitation of this approach is the model's initial conditions, a crucial aspect of event-based simulations. Despite the sensitivity of initial conditions to flood discharges depending on the complexity of model structures and our observation that our model, representing only the topsoil layer, has comparatively lower sensitivity to initial conditions, the assumption of no storage at the beginning of the simulations may also impact the calibrated parameters. Owing to the high computational demands of the 150 m resolution nationwide model, we

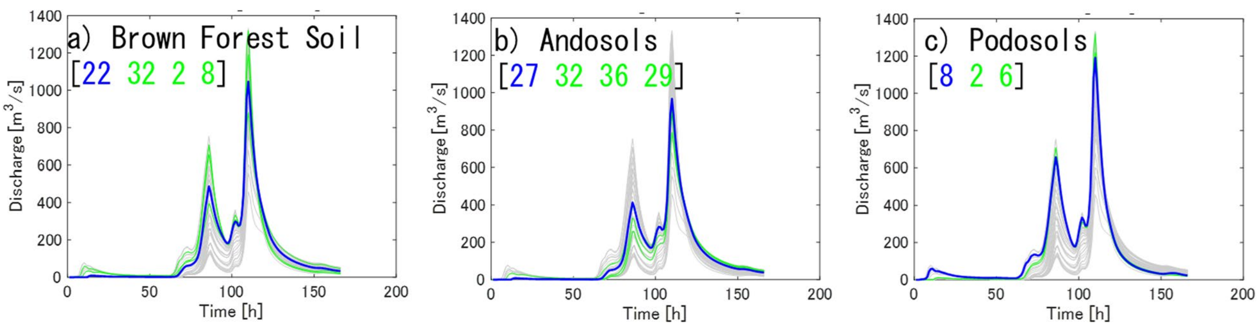

Fig. 10 The differences in simulated hydrographs by the three identified parameter sets applicable to the major soil types. The background gray lines show the hydrographs by all the 40 possible parameters, while the green lines show the results by the other possible parameter sets for each soil type (**a** Brown forest soil, **b** Andosols, **c** Podosols)

could not conduct continuous simulations with numerous parameters. Assessment of the impacts of the initial conditions through a comparison of event-based and continuous simulations is left for future work.

5 Conclusions

Parameter regionalization is crucial in improving the performance of large-domain hydrological modeling. Challenges associated with parameter regionalization include reflecting spatially distributed signatures, such as soil or geologic maps, and maintaining seamless distributions of state variables and fluxes in large domains under the prohibitively high computational cost of the model running. We propose a parameter regionalization method based on conditional probability to address these challenges. The advantage of the proposed CPPR method is that the first phase of model running can be achieved with spatially uniform parameter sets in many calibration basins to estimate the likelihood of each parameter in each calibration basin, that is $p(q_k|d_j)$. Based on the calibration and areal occupation ratios $p(s_i|d_j)$ of each soil or geology type in the calibration basins, we can compute the likelihood of parameter sets that should be assigned for each soil or geology type $p(q_k|s_i)$. The proposed method was applied to a nationwide distributed RRI model covering Japan with a 150 m spatial resolution. The conclusions are summarized as follows:

1) The "Soil" case by the CPPR method outperforms the "Optimized 2" case, which uses a spatially uniform best parameter set across all calibration basins. The median NSE of the "Soil" case by the CPPR method was 0.83, which closely aligns with the NSE value of 0.88 achieved by the "Optimized 1," which selects the best parameter set individually at each calibration basin. While the "Optimized 1" case shows an underestimation bias (RPD = −0.15), the "Soil" case by the CPPR results in a much smaller bias (RPD = − 0.05), whose characteristics are important for practical uses in flood predictions.
2) Among the two cases by CPPR, the "Soil" case exhibits superior results compared to the "Geology" case, emphasizing the efficacy of considering soil distribution. This is likely because the model structure and parameters used in this application focused on the subsurface of the soil layer and surface flow. Geology information should play a more dominant role if we extend our model to bedrock groundwater, which may further improve model performance.
3) To validate the model, we evaluated its performance with 2723 flood events at 711 gauging stations based on the regionalized parameter set by the CPPR method with the "Soil" case. The median NSE value was 0.87, comparable to the above calibration results. The median CC was high at 0.96, and over 75% of the tested events had a CC of 0.94 or higher.
4) Flood events of larger magnitude tended to show higher NSE values, whereas the event size less affected the RPD. The median RPD was −0.1 to 0.0 for all magnitude categories, indicating some underestimation bias. Regarding the model performance and basin size, the scale of 100 km^2–1000 km^2 was the highest, with a median NSE of 0.89 for 1304 events. The first quartile of NSE was 0.81, indicating good performance for this basin scale. Regional differences in model accuracy were significant, with NSE generally yielding good results in western Japan.

While the proposed parameter regionalization method improved the model accuracy, it was not able to perfectly represent all observed flood events. It is necessary to consider the uncertainties associated with the model structure, initial conditions, and input data to enhance model performance. Despite these challenges, the nationwide model holds promise for comprehensive flood risk assessment across Japan. The proposed CPPR method is computationally efficient and effective for obtaining seamless state variables and fluxes that reflect hydrological signature distributions.

Abbreviations

CC	Correlation coefficient
CPPR	Conditional probability parameter regionalization
JRRI	Japan rainfall–runoff–inundation
J-FlwDir	Japan flow direction map
KGE	Kling–Gupta efficiency
MLIT	Ministry of Land, Infrastructure, Transport, and Tourism
MPR	Multiscale parameter regionalization
NSE	Nash–Sutcliffe efficiency
RPD	Relative peak difference
RRI	Rainfall–runoff–inundation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40645-025-00691-w.

Additional file 1.

Acknowledgements
We thank MLIT for providing hydrological and river data. We would also like to thank Associate Professor Dai Yamazaki for providing the J-FlwDir dataset, which is an important RRI model input.

Author contributions
TS conceptualized the topic, methodology and wrote the original draft. MY contributed the method, data curation and visualization of the results. AY conducted simulations and investigation. YS contributed to the model development and validation of the model.

Funding
The Cross-Ministerial Strategic Innovation Promotion Program (SIP) by the Cabinet Office, Government of Japan, Japan Society for the Promotion of Science (JSPS) KAKENHI [grant JP19H02248,24H00334 for PI: Sayama,JP21H04597 for PI: Yamori, JP22H01600 for PI: Tachikawa,23K26202 for PI: Yokoo, and 20H00434 for PI: Kosugi], Japan Society for the Promotion of Science, KAKENHI JP19H02248,Takahiro Sayama, 24H00334, Takahiro Sayama.

Availability of data and materials
The RRI model is available at https://www.pwri.go.jp/icharm/research/rri/index.html. The flow direction data can be downloaded from https://hydro.iis.u-tokyo.ac.jp/~yamadai/JapanDir/. The other simulation dataset supporting the conclusions of this article is available upon request.

Declarations

Competing interests
The authors declare that they have no competing interest.

Received: 30 September 2024 Accepted: 1 February 2025
Published online: 24 February 2025

References

Abdulla FA, Lettenmaier DP (1997) Development of regional parameter estimation equations for a macroscale hydrologic model. J Hydrol 197:230–257. https://doi.org/10.1016/S0022-1694(96)03262-3
Addor N, Do HX, Alvarez-Garreton C, Coxon G, Fowler K, Mendoza PA (2020) Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. Hydrol Sci J 65:712–725. https://doi.org/10.1080/02626667.2019.1683182
Archfield SA, Clark M, Arheimer B, Hay LE, McMillan H, Kiang JE, Seibert J, Hakala K, Bock A, Wagener T, Farmer WH, Andréassian V, Attinger S, Viglione A, Knight R, Markstrom S, Over T (2015) Accelerating advances in continental domain hydrologic modeling. Water Resour Res 51:10078–10091. https://doi.org/10.1002/2015WR017498
Beven KJ, Chappell NA (2021) Perceptual perplexity and parameter parsimony. Wires Water 8:e1530. https://doi.org/10.1002/wat2.1530
Carrillo G, Troch PA, Sivapalan M, Wagener T, Harman C, Sawicz K (2011) Catchment classification: hydrological analysis of catchment behavior through process-based modeling along a climate gradient. Hydrol Earth Syst Sci 15:3411–3430. https://doi.org/10.5194/hess-15-3411-2011
Clark MP, Fan Y, Lawrence DM, Adam JC, Bolster D, Gochis DJ, Hooper RP, Kumar M, Leung LR, Mackay DS, Maxwell RM, Shen C, Swenson SC, Zeng X (2015) Improving the representation of hydrologic processes in earth system models. Water Resour Res 51:5929–5956. https://doi.org/10.1002/2015WR017096
Fan Y, Clark M, Lawrence DM, Swenson S, Band LE, Brantley SL, Brooks PD, Dietrich WE, Flores A, Grant G, Kirchner JW, Mackay DS, McDonnell JJ, Milly PCD, Sullivan PL, Tague C, Ajami H, Chaney N, Hartmann A, Hazenberg P, McNamara J, Pelletier J, Perket J, Rouholahnejad-Freund E, Wagener T, Zeng X, Beighley E, Buzan J, Huang M, Livneh B et al (2019) Hillslope hydrology in global change research and earth system modeling. Water Resour Res 55:1737–1772. https://doi.org/10.1029/2018WR023903
Gupta HV, Perrin C, Blöschl G, Montanari A, Kumar R, Clark M, Andréassian V (2014) Large-sample hydrology: a need to balance depth with breadth. Hydrol Earth Syst Sci 18:463–477. https://doi.org/10.5194/hess-18-463-2014
Hundecha Y, Bárdossy A (2004) Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. J Hydrol 292:281–295
IPCC (2022) 2022: impacts, adaptation, and vulnerability. Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change climate change. Cambridge University Press, Cambridge, UK and New York, p. 3056 https://doi.org/10.1017/9781009325844
Masutani K, Akai K, Magome J (2006) A new scaling algorithm of gridded river networks. J Jpn Soc Hydrol Water Resour 19:139–150. https://doi.org/10.3178/jjshwr.19.139
Merz R, Tarasova L, Basso S (2020) Parameter's controls of distributed catchment models—how much information is in conventional catchment descriptors? Water Resour Res 56(WR026008):e2019. https://doi.org/10.1029/2019WR026008
Merz R, Miniussi A, Basso S, Petersen KJ, Tarasova L (2022) More complex is not necessarily better in large-scale hydrological modeling: a model complexity experiment across the contiguous United States. Bull Am Meteorol Soc 103:E1947–E1967. https://doi.org/10.1175/BAMS-D-21-0284.1
Mizukami N, Clark MP, Newman AJ, Wood AW, Gutmann ED, Nijssen B, Rakovec O, Samaniego L (2017) Towards seamless large-domain parameter estimation for hydrologic models. Water Resour Res 53:8020–8040
Musiake KT, Takahasi Y, Ando Y (1981) Effects of basin geology on river-flow regime in mountainous areas of Japan. Proc Jpn Soc Civil Eng 1981:51–62. https://doi.org/10.2208/jscej1969.1981.309_51
Samaniego L, Kumar R, Attinger S (2010) Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. Water Resour Res. https://doi.org/10.1029/2008WR007327
Samaniego L, Kumar R, Thober S, Rakovec O, Zink M, Wanders N, Eisner S, Müller Schmied H, Sutanudjaja EH, Warrach-Sagi K, Attinger S (2017) Toward seamless hydrologic predictions across spatial scales. Hydrol Earth Syst Sci 21:4323–4346. https://doi.org/10.5194/hess-21-4323-2017
Sayama T, McDonnell JJ (2009) A new time-space accounting scheme to predict stream water residence time and hydrograph source components at the watershed scale. Water Resour Res 45:W07401. https://doi.org/10.1029/2008WR007549
Sayama T, Yamada M, Sugawara Y, Yamazaki D (2020) Ensemble flash flood predictions using a high-resolution nationwide distributed rainfall-runoff model: case study of the heavy rain event of July 2018 and Typhoon Hagibis in 2019. Prog Earth Planet Sci 7:75. https://doi.org/10.1186/s40645-020-00391-7
Schweppe R, Thober S, Müller S, Kelbling M, Kumar R, Attinger S, Samaniego L (2022) MPR 1.0: a stand-alone multiscale parameter regionalization tool

for improved parameter estimation of land surface models. Geosci Model Dev 15:859–882. https://doi.org/10.5194/gmd-15-859-2022
Tachikawa Y, Nagatani G, Takara K (2004) Development of storage-discharge relationship equation incorporating saturated-unsaturated flow mechanism (in Japanese). Ann Hydraul Eng Jpn Soc Civ Eng 48:7–12
Yamada M, Togashi S, Sayama T (2020) Estimation method of cross-section of small rivers for distributed rainfall-runoff-Iunundation models (in Japanese). Adv River Eng 26:211–216. https://doi.org/10.11532/river.26.0_211
Yamada M, Sayama T, Yamazaki D, Watanabe M (2022) Development of methodology for introducing surveyed river cross-sections into nationwide distributed hydrological model throughout Japan and its effect on water level prediction accuracy (in Japanese). J Japan Soc Civil Eng Ser (Hydraulic Engineering) B1(78):7–22. https://doi.org/10.2208/jscejhe.78.1_7
Yamazaki D, Togashi S, Takeshima A, Sayama T (2018) High-resolution flow direction map of Japan. J Soc Civ Eng B1(74):I_163-I_168. https://doi.org/10.2208/jscejhe.74.5_I_163
Young AR (2006) Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. J Hydrol 320:155–172. https://doi.org/10.1016/j.jhydrol.2005.07.017

Publisher's Note

Ayato Yamakita is a former Master course student at graduate school of engineering, Kyoto University, and currently working at Ministry of Land Infrastructure, Transport and Tourism (MLIT), Japan.