LETTER

Impacts of dams and reservoirs on local climate change: a global perspective

Yiyang Zhao 1,2, Suning Liu 1,2 and Haiyun Shi 1,2,*

1 State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, People’s Republic of China
2 Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, People’s Republic of China
* Author to whom any correspondence should be addressed.
E-mail: shihy@sustech.edu.cn

Keywords: dams and reservoirs, reservoir characteristic factors, meteorological changing behaviors, Pearson correlation analysis, multiple linear regression

Abstract

As a primary artificial project in a river basin, the role of large dams and related reservoirs in mitigating the hydrological and meteorological extremes has been widely recognized. Moreover, construction of additional large dams is considered as one of the best available options to meet future increases in water, food, and energy demands, although this may drastically affect the surrounding natural environment. However, the issue about the degree of such effects is still unclear, although the interactive impacts of dams and related reservoirs on the changes in meteorological variables are of great scientific importance. Therefore, this study aims to explore the impacts of reservoir characteristic factors (RCFs) on local climate change near the reservoir through examining the correlations between the RCFs of selected dams and related reservoirs and meteorological variables all over the world. The major findings are: (a) the correlations of the RCFs with evaporation are the opposite of those with precipitation; (b) dams and related reservoirs can have completely opposite (weakening or enhancing) effects on different meteorological variables; and (c) the RCFs outperform the geographical factors in terms of the impact on precipitation, whereas the variation of evaporation is more sensitive to geographical factors. Overall, the outcomes of this study will be useful for policymakers to have a more informed understanding of local climate change affected by large dams and related reservoirs in the future.

1. Introduction

The construction of dams and related reservoirs has essentially been the result of population growth and the associated consequence of increased consumption of water, food, and energy (Chen et al. 2016, Shi et al. 2019). Meanwhile, climate change will cause higher flood risk at the end of the 21st century, but a severe challenge of hydrological extremes can be reduced and controlled by reservoir regulation (Yun et al. 2021). Then, should the construction of dams and related reservoirs be responsible for local or even global climate change? If so, which meteorological variables are more sensible? Such questions have attracted the attention of many researchers. Due to construction of dams and reservoirs, the change in land surface and consequent alteration of water cycle will directly lead to change in radiation and storage of various greenhouse gases, and eventually disturb the interaction between water and atmosphere (Chi et al. 2021, Li et al. 2021), along with the local climate behaviors. As the most important type of artificial land surface change, the construction of dams and related reservoirs can convert the original surface land-cover into water body, store and conserve water, discharge runoff, and alter local water cycle (Keigo et al. 2018, Yun et al. 2021, Zaidel et al. 2021). Dams and related reservoirs have been proven to be effective in watershed tolerance and resistance of meteorological, hydrological, and socioeconomic drought propagation (Guo et al. 2021). Besides, several studies have indicated some unspecific connections between local
climate change and the construction of dams and related reservoirs through analyzing the characteristics of rainfall and temperature changes. For example, Chen et al. (2007) found no significant change in precipitation but an increasing trend in temperature since the water impoundment of Three Gorges Reservoir. Zheng et al. (2017) concluded that the climate effect of Miyun Reservoir was mainly reflected in the summer-half year, with the temperature and precipitation being the most obvious targets. Liang et al. (2017) reported that the annual average temperature, average highest temperature and average lowest temperature in the main stream of Wujiang River all showed warming trends due to the cascade reservoirs. Given the significance of dams and reservoirs in the global water cycle, a better comprehension of the impacts of reservoirs on meteorological changing behaviors (MCBs) would not only be the key to the above questions, but also be of great importance to the mitigation of the climatic risks and hydrological extremes in future (Keigo et al. 2018, Guo et al. 2021). Nonetheless, most current studies on the relationships between the MCBs and dams and related reservoirs have only focused on the responses of reservoirs to global climate change, not the other way around. The responsive behaviors of chemical elements, microbe quantity and water quality in reservoir areas to climate change have been commonly addressed (Tong et al. 2019, Mello et al. 2021). Only a few studies have attempted to find how the construction of dams could reversely affect the local MCBs (Yu et al. 2020, Wu et al. 2021), while no studies have ever tried to build the bridge between the MCBs and reservoir characteristics numerically, especially from a global perspective.

It can be inferred that the physical mechanism behind the process of dams and reservoirs affecting the MCBs remains fundamental to related questions and a better comprehension. To our best knowledge, a land–surface–atmosphere model is commonly used in representing the actual global water–atmosphere circulation and abstracting the physical mechanism (Bram et al. 2021, Dos Santos et al. 2021). Yet such a model is still rarely applied to reservoir areas because there is no direct connection between the required parameters (e.g. soil moisture, porosity, and plant coverage) and reservoir characteristics. Given the deficiency, existing studies have been attempting to bypass the land–surface–atmosphere model and detect the hidden connection between the dams and related reservoirs and the local MCBs, most of which, however, restrictedly focused on only one representative dam and related reservoir as a case (e.g. Three Gorges Reservoir in China). Such research status inevitably brings limitations: (a) conclusions like ‘Annual average temperature in reservoir region has an increasing trend after the construction of reservoir’ are rather general. (b) Even true, no valid evidence can authenticate that the ‘increasing trend of annual average temperature’ is necessarily and exclusively caused by dams and reservoirs. For instance, Rodrigues et al. (2021) found that regional air pollution would render different behaviors on evaporation of several tropical reservoirs in Brazil. (c) Current conclusions from different studies on the MCBs near the reservoirs are sometimes inconsistent (e.g. an increasing trend for one reservoir but a decreasing trend in another one).

It is worth noting that, large-scale atmosphere circulation (LSAC) could account for the MCBs (near a reservoir or not) to a larger extent than reservoirs (Dong et al. 2020, Wang et al. 2021). Besides, LSAC with specific temporal/spatial-frequency distribution will disturb climate change with uncertainty (Zhou et al. 2020). Obviously, LSAC will affect the results of our study by introducing ‘noise’ when an exclusive relationship between reservoirs and the MCBs around the reservoir is expected, therefore, the elimination of this ‘noise’ is a necessary validation, which is not easy due to the complexity of LSAC.

In light of the difficulty in elaborating the physical mechanism behind the process of reservoir altering the MCBs, along with all related questions and limitations aforementioned, a major objective of this study would be establishing some relationships that can be numerically quantified between several reservoir characteristic factors (noted as RCFs hereinafter, such as water surface area, storage capacity, and upstream catchment area) and the local MCBs. Besides, given the existence of unexpected ‘noise’, some methods should be adopted to eliminate this noise to the greatest extent, as a supplementary objective. Some intuitive hypotheses like ‘The influencing-scale and scope of a reservoir on the MCBs could be linearly or non-linearly correlated to the RCFs’ and ‘A linear or non-linear function numerically connecting the magnitude of dams and related reservoirs and the MCBs could exist’ is based on the objectives. To the best of our knowledge, no study ever tried to numerically quantify the correlations between the RCFs and the local MCBs, and those original questions and intuitive hypotheses remain unsolved.

Summarily, this study aims to detect the most important RCFs which may somehow affect the local climate near the reservoir through examining the correlations between the RCFs of selected dams and related reservoirs and the MCBs globally. Moreover, this study aims to numerically quantify to what extent the MCBs can be explained by these detected RCFs. In the following, section 2 briefly introduces the research data and the methods used in this study, section 3 presents the results and discussion, and section 4 summarizes the major conclusions of this study. Overall, the outcomes of this study can be of great value for better understanding the impacts of reservoirs on regional climate behaviors,
and quantitatively structuring the potential relationship between the characteristics of dams and related reservoirs and the degree of local climate change. Furthermore, this study will attempt to attribute the main results to the physical mechanism of process as a reverse analysis.

2. Data and methods

2.1. Framework

The logic framework of this study is shown in figure 1. We suggest three approaches to disassemble the complexity of physical mechanism and explore the anticipated objectives:

Approach (1): some RCFs quantitatively describing the reservoirs are selected, and then, statistic method is adopted to explore the correlations between the RCFs and the MCBs, based on the meteorological data right above the reservoir center, following the previous intuitive hypotheses. Multiple indices are set, and the RCFs and the MCBs with significant correlations will be sorted out.

Approach (2): study ranges are expanded from the point right above the reservoir center to an area with a limited scope, namely the buffer zone of reservoir, to investigate the correlations between the RCFs and the MCBs within the buffer zone. This buffer zone is mainly introduced for abating the ‘noise’ caused by LSAC and will be further explained and discussed in the following sections.

Approach (3): since the ‘noise’ caused by LSAC is difficult to eliminate, instead of attempting to separate them, this study adopts some statistical methods to create a general picture of how much the RCFs and LSAC are responsible for the local MCBs, respectively, and then make a comparison between them. This approach is adopted in case of the buffer zone introduced in Approach (2) cannot effectively abate the ‘noise’ caused by LSAC, and therefore, we can turn to obtain a general view of which MCBs are more sensitive to LSAC and which are more sensitive to the RCFs.

2.2. Research data

Two datasets are used in this study, i.e. the datasets of dams and related reservoirs and meteorological variables for each dam and related reservoir. For the datasets of dams and related reservoirs, there are 7320 dams constructed from the 17th century AD to 2017 globally in the Global Reservoir and Dam (GRanD) database (Lehner et al 2011). For each dam and related reservoir, the basic information includes name, year of completion, longitude, latitude, height, area, storage capacity, average depth, installed capacity, catchment area, main use, long-term average discharge, water residence time, and so on. The meteorological data used in this study are derived from a grid-based data set (ERA5-Land monthly averaged data) provided by European Centre for Medium-Range Weather Forecasts (https://cds.climate.copernicus.eu/), with the horizontal resolution of 0.1° × 0.1° and the temporal resolution of 1 month (Muñoz 2019), available from 1981 to 2020 (40 years in total). Meteorological variables such as precipitation, evaporation, and temperature (surface temperature and temperature 2 m above surface) are chosen for this study. After data pretreatment, only 200 reservoirs with the storage capacity larger than 0.1 km² and relatively broad area are selected globally for analysis (figure 2). It is worth noting that, since it is rational to ensure that the contrast analysis be conducted on the same timescale, for each given reservoir in this study, the 40-year time sequence of a given meteorological variable should be divided into two roughly equal sequences (both around 20 years) for the contrast analysis before and after the construction of a reservoir (i.e. pre-reservoir and post-reservoir sequences). Therefore, only dams and related reservoirs completed during the period of 1995–2005 are selected, while some reservoirs, despite their significant magnitudes, are ruled out given that their construction times are outside of this period. The average water surface area of these 200 samples is 67.67 km² (with the largest sample being 1205 km²), and the average storage capacity of these 200 samples is 1.87 km³ (with the largest sample being 54.40 km³). The overall situation of the water surface areas of 200 selected reservoirs is shown in figure 3.

As shown in figures 3(a) and (b), 200 dams and related reservoirs can be roughly divided into three categories. Most reservoirs have the water surface areas less than 100 km² and these reservoirs can be fully covered by a 1 × 1 grid (i.e. category I). There are less than 20 reservoirs with the areas that can be fully covered by a 3 × 3 grid-group (i.e. category II), and four mega reservoirs can only be fully covered by a 5 × 5 grid-group (i.e. category III). The 1 × 1 grid with meteorological data right above the reservoir center for each dam will be calculated and arranged in a 200-point sequence, and correlation analysis between that and the RCFs will be conducted first.

However, given the previous studies regarding the influencing radius, also known as the buffer zone of a reservoir (Chen et al 2007, Zheng et al 2017, 2021, Dong et al 2020), it is rational to assume that there is spatial heterogeneity regarding the interactions between reservoirs and local MCBs, that is, the underlying surface change might cause a disparity in spatial distribution of climate change, leading to the MCBs above the periphery area of a reservoir having an even closer correlation with the RCFs than that right above the reservoir center (see figure 3(c)). This may render the analysis of 1 × 1 grid in reservoir center imprecise. Normally, 10 km is a believed effective influencing radius of a reservoir, and a study on the Three Gorges Reservoir (Tian et al 2021) claimed that the effective influencing radius of mega reservoirs
as huge as the Three Gorges Reservoir could reach to 100 km, which is also supported by Guo et al. (2021), yet none of these radii have been technically proven by any physical and mathematical theory. The 200 selected dams in this study vary dramatically in magnitude, and it is unable to establish a clear relationship between the magnitude of a reservoir and its potential buffer zone. Therefore, a rational assumption claiming the effective influencing radius (i.e. the buffer zone) of a reservoir as 30–40 km (within the range from 10 km to 100 km) can be made since most of these reservoirs do not reach the potential effective influencing radius of 100 km. As mentioned before, the 200 reservoirs can be roughly divided into three categories according to their water surface area, and thus, a group of $7 \times 7$ grids (i.e. 49 grids in total), $9 \times 9$ grids (i.e. 81 grids in total), and $11 \times 11$ grids (i.e. 121 grids in total) of the meteorological data with the center grid right above the reservoir center are collected for each category, respectively. Therefore, a rational buffer zone (with the effective influencing radius of 30–40 km) is attached to each reservoir with different water surface areas.

The main reason why this buffer zone theory coordinating with Approach (2) is introduced here is to reach our secondary objective, that is, since the 200 dams are dispersive globally, the sequence consisting of $1 \times 1$ grid meteorological data above the reservoir center could possibly be a noisy sequence based on the former explanation regarding LSAC. Whereas within

---

**Figure 1.** The framework of this study.

**Figure 2.** Global distribution map of 200 selected reservoirs. Note: the watershed boundary is obtained from FAO-UN (2011).
a buffer zone, LSAC could be assumed to exert almost equal influence on each grid (Wang et al. 2021). In such case, if any discrepancy within buffer zone is found (e.g., the disparity between the grid with a maximum meteorological value and a minimum one), we can believe that such discrepancy is caused mainly by the reservoir with the larger confidence. Therefore the ‘noise’ caused by LSAC in meteorological sequence can be attenuated to a large extent, and the correlations between the RCFs and the MCBs will be guaranteed with a higher reliability.

2.3. Indices for analysis
After analysis of basic information of dams and related reservoirs, six quantitative RCFs are selected, including water surface area (AREA), storage capacity (CAP), average depth (DEP), catchment area (CATCH), long-term average discharge (DIS), and degree of residence time of water in reservoir (DOR). The first three RCFs (i.e., AREA, CAP, and DEP) are basic factors describing the physical magnitude of reservoirs from a narrow spatio-temporal dimension and they are often taken into consideration in research of large-scale temporal/spatial water cycle or water–atmosphere interaction all over the world (Biemans et al. 2011, Cheng et al. 2019, Matta et al. 2019, Tong et al. 2019, Yang et al. 2020). Adeloye et al. (2019) claimed that, water surface area (AREA), storage capacity (CAP) and average depth (DEP) of a reservoir could functionally related with evaporation-loss in reservoir area. Matta et al. (2019) claimed that, catchment area (CATCH) of a reservoir, including tributaries, branches, wetland landscape and water diversion channels, could all contribute to the alteration of hydrodynamics and local climate extremes. Long-term average discharge (DIS) along with water residence time (DOR) could describe the impact of a reservoir on local water cycle from a temporal perspective (Cheng et al. 2019). CATCH, DIS and DOR are more like the reservoir ‘characteristics’ whereas AREA, CAP and DEPTH are like reservoir ‘appearance’.

Similarly, four meteorological variables commonly adopted by previous studies are selected to describe the local climate change: evaporation (E), precipitation (P), surface temperature (ST), and temperature 2 m above surface (T2M). Here, two temperature indices ST and T2M are introduced in case the existence of a vertical disparity in
temperature, which could be potentially caused by reservoirs. To better quantify the relationships between the six selected RCFs and four selected meteorological variables, the following six indices with different physical meanings are designed for each meteorological variable, taking evaporation (E1–E6) as an example:

- E1 represents the Variance of post-reservoir evaporation sequence, a given 1 × 1 grid.
- E2 represents the difference between the Variance of pre-reservoir and post-reservoir evaporation sequences, a given 1 × 1 grid.
- E3 represents the Long-term Annual Average Value (LAAV) of post-reservoir evaporation sequence, a given 1 × 1 grid.
- E4 represents the difference between the LAAV of pre-reservoir and post-reservoir evaporation sequences, a given 1 × 1 grid.
- E5 represents the difference in the LAAV of the post-reservoir sequence between the maximum grid and the minimum grid within the grid-group (considering the buffer zone).
- E6 represents the difference in the Variance of the post-reservoir sequence between the maximum grid and the minimum grid within the grid-group (considering the buffer zone).

It is worth noting that, P1–P6, ST1–ST6, and T2M1–T2M6 are designed for P, ST, and T2M in the same way, respectively. The above indices are set based on the following reasons: first, this study focuses on the changes in the LAAV and variance of each meteorological variable, which are two important properties for time series. Second, it is worth investigating whether construction of dams and related reservoirs could trigger static (i.e. the behavior of post-reservoir meteorological series) or dynamic (i.e. the disparity between the behaviors of pre-reservoir and post-reservoir meteorological series) impacts on meteorological variables. Third, the buffer zone theory is introduced to create E5–E6 according to Approach (2), to address the issue of heterogeneity assumption and noisy sequence due to LSAC.

2.4. Analysis methods

The analysis methods adopted in this study are Pearson correlation analysis and multiple linear regression (Wang et al 2021). Pearson correlation analysis describes the degree of compactness by calculating the linear relationship magnitude between two variables. Correlation coefficient value greater than zero indicates a positive correlation between two variables, and a negative correlation between two variables is indicated by the correlation coefficient value less than zero. The equation for calculating Pearson correlation coefficient is as follows:

\[
\text{COR}(X, Y) = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N} (X_i - \bar{X})^2 (Y_i - \bar{Y})^2}}
\]

where \(N\) is the number of samples, \(X_i\) and \(Y_i\) are the observed values at point \(i\) corresponding to the variables \(X\) and \(Y\), \(\bar{X}\) and \(\bar{Y}\) are the sample mean for \(X\) and \(Y\).

The absolute value of correlation coefficient varies between 0 and 1, indicating a weak correlation when correlation coefficient is close to 0 and a strong correlation when correlation coefficient is close to 1. To evaluate whether there is a significant correlation between two variables, a statistical significance test of the correlation coefficient is required (T-test is applied in this study), and the obtained p-value reflects the probability. Normally, \(p < 0.05\) and \(p < 0.01\) are two common significance levels, representing that the chance of the correlation happens to be non-statistical is less than 5% and 1%, respectively. In this study, the correlation coefficients passing the significance test at the significance levels of \(p < 0.05\) and \(p < 0.01\) are labeled as “\(a\)” and “\(b\)”, respectively.

Multiple linear regression is mainly adopted in Approach (3) as a supplemental method to analyze the relationship between all the RCFs as a whole and the MCBs, quantifying the holistic responsibility of the RCFs to local climate change. According to Approach (3), multiple linear regression is also adopted for the contrast between the impacts on the MCBs from RCFs and those from LSAC. Since most of these LSAC have close relationships with geographic factors like longitude, latitude, and elevation, some of which are even linear (Wang et al 2021, Wei et al 2021, Zhang et al 2021), it is rational to hypothetically represent LSAC with geographic factors.

In line with Approach (2), in this study, rather than only focusing on the 1 × 1 grid of MCBs in each reservoir center, the groups of 7 × 7, 9 × 9, and 11 × 11 grids representing reservoir buffer zone are also analyzed for each dam and related reservoir. Based on heterogeneity consumption and buffer zone theory introduced in Approach (2), a new research idea, also a sparking point, is generated in this study: the maximum and minimum values of the MCBs are collected among buffer zone for all the 200 reservoirs, and similar to the center-grid sequence, the max-grid sequence and the min-grid sequence can be created by changing the position of the given 1 × 1 grid for indices E1–E4. It is worth noting that, given the fact that the max/min-grid can be anywhere within the buffer zone, for a given reservoir, the position of the max/min-grid for a given reservoir can be totally different with one another. Then, three correlation patterns will be created (i.e. center-grid to RCF, max-grid to RCF and min-grid to RCF) and
further analysis can be conducted based on these three correlation patterns.

3. Results and discussion

3.1. Correlations between the RCFs and meteorological variables

This subsection is mainly in line with Approach (1). Correlations between the six selected RCFs and the selected four meteorological variables are analyzed in the center-grid and the max/min-grid, respectively. Then, the comparisons between the correlation patterns (i.e. the distribution of figure-cells with \(a\) and \(b\)) in the center-grid and max/min-grid are conducted.

3.1.1. Correlations between meteorological variables in the center-grid and the RCFs

Figures 4(a) and (b) represent the correlation patterns of the RCFs with evaporation and precipitation, respectively, which are two completely opposite physical processes in hydrological cycle. For example, Ma (2020) reported that the impoundment of the Xiaolangdi Reservoir could render a downward trend on evaporation but an upward trend on precipitation. From an overall perspective towards figure 4, despite the generally low absolute values of correlation coefficients (ranging from \(-0.35\) to \(+0.35\)), it can still be trusted that these correlations are real and significant by virtue of significance test. Significance test excludes those insignificant correlations, and we can only focus on those correlations passing the significance test (i.e. figure-cells with ‘a’ and ‘b’).

Focusing on evaporation as shown in figure 4(a), among these six indices, E3 has the most significant correlations with the RCFs, with four of the six RCFs (i.e. AREA, CAP, DIS and DOR) being significantly correlated with E3. Among them, the correlation between E3 and CAP passes the significance test at the level of \(p < 0.05\), and the correlations of E3 with AREA, DIS and DOR pass the significance test at the level of \(p < 0.01\). Moreover, each of E1, E2 and E5 has significant correlations with two RCFs, passing the significance test at the level of \(p < 0.05\). In contrast, E4 and E6 do not have significant correlations with any of the six RCFs. On the other hand, focusing on precipitation as shown in figure 4(b), among these six indices, P4 has the most significant correlations with the RCFs, with five of the six RCFs (except AREA) passing the significance test. Then, P2 has significant correlations with three RCFs (i.e. AREA, CAP and CATCH), while each of P1, P3, P5 and P6 has a significant correlation with only one RCF (i.e. CATCH).

Based on the comparison of figures 4(a) and (b), it is interesting that the distribution of figure-cells with ‘a’ and ‘b’ for evaporation (figure 4(a)) is completely different from that for precipitation (figure 4(b)). For instance, the RCFs can affect more significantly on the LAAV of post-reservoir evaporation sequence than the corresponding precipitation sequence (see E3 and P3) but affect less significantly on the difference between the LAAV of pre-reservoir and post-reservoir evaporation sequences than the corresponding precipitation sequences (see E4 and P4). Especially, CATCH, compared to other RCFs, can affect precipitation more significantly but can barely affect any evaporation indices. Moreover, the RCFs which can exert significant impacts on evaporation have little influence on precipitation, and vice versa. The study of Chen et al. (2021) regarding evapotranspiration and precipitation could provide a potential explanation for such findings: there was a poor correlation between the two variables in alpine wetland (in contrast to meadow) where surplus rainfall could infiltrate into deep soil more easily, and the behavior was largely due to soil moisture of underlying surface in a complicated way. In our case, the reservoir can be regarded as an extreme case of ‘mega wetland’ and the RCFs can be regarded as the parameters describing ‘moisture volume’ of a reservoir.

With reference to the two variables of temperature (i.e. ST and T2M), they have similar correlation patterns, indicating that there is no significant vertical disparity in temperature caused by reservoirs. Both seem to have more significant correlations with the six RCFs, since most of the correlations can pass the significance test (see figures 4(c) and (d)). Among them, the fourth index, which indicates the difference between the LAAV of pre-reservoir and post-reservoir sequences, is regarded as an exception. Both ST4 and T2M4 do not have a significant correlation with any RCF. From the perspective of the RCFs, it is worth noting that, the first three RCFs (i.e. AREA, CAP and DEP) have less significant correlations with temperature indices as a whole, whereas the last three RCFs (i.e. CATCH, DIS and DOR) have great impacts on temperature indices, especially for DIS and DOR. It is possibly due to that the long-term average discharge (i.e. DIS) and the degree of residence time of water in reservoir (i.e. DOR) can have direct impacts on the dynamic processes of nitrogen, oxygen, and carbon in reservoirs (Tong et al. 2019), leading to the concentration and diffusion of related greenhouse gases (e.g. carbon dioxide and nitrous oxide which can affect the local temperature) above the reservoirs.

3.1.2. Correlations between meteorological variables in the max/min-grid and the RCFs

As the results shown in section 3.1.1, the highest value of correlation coefficient is only up to 0.35. Validated by the significance test though, it still poorly reaches the expectation that any of these RCFs numerically and linearly correlated with the MCBs in the reservoir center. In this case, it is rational to resort to heterogeneity in Approach (2). Therefore, as mentioned before, we find the grids...
Correlations between the RCFs and meteorological variables in the center-grid. with the maximum and minimum meteorological data among the range considering the buffer zone of each reservoir (7 × 7, 9 × 9, or 11 × 11 grid-group) and do the same with them as that in section 3.1.1. This subsection focuses on the former four indices (e.g. E1–E4) for each meteorological variable, since the last two indices (e.g. E5 and E6) are identical for the three correlation patterns (i.e. center-grid to the RCFs, max-grid to the RCFs, and min-grid to the RCFs).

In view of figures 5 and 6, no distinct correlation coefficient value can be found, which means that the RCFs do not have a close correlation directly with the max/min-grid within the range considering the buffer zone. However, when comparing figure 4(a) with figures 5(a) and 6(a) respectively, interesting correlation patterns of evaporation (i.e. the distribution of figure-cells with ‘a’ and ‘b’) are found. E1 and E2 in the center-grid share the similar correlation patterns with those in the min-grid, which indicates that E1 and E2 in the center-grid happen to be the minimum sequences among the range considering the buffer zone; in contrast, E3 and E4 in the center-grid share the similar correlation patterns with those in the max-grid, which indicates that E3 and E4 in the center-grid happen to be the maximum sequences among the range considering the buffer zone. The results suggest that a reservoir can have a weakening effect on the variance (E1 and E2) but an enhancing effect on the LAAV (E3 and E4).

Regarding precipitation in figures 5(b) and 6(b), the indices P1–P4 in the center-grid (figure 4(b)) notably share the similar correlation patterns with those in the max-grid, which means that the precipitation changes in the center of reservoir happen to be the maximum sequences among the range considering the buffer zone. Therefore, a reservoir can basically have an enhancing effect on the precipitation changes, including both the variance and the LAAV.

In addition, the two variables of temperature (i.e. ST and T2M) in the center-grid both have similar correlation patterns with those in the min-grid (see figures 4(c), (d), 6(c) and (d)), which means that the temperature changes in the center of reservoir happen to be the maximum sequences among the range considering the buffer zone. Therefore, a reservoir can have a weakening effect on the temperature changes, including both the variance and the LAAV.

3.2 Multiple linear regression of meteorological variables and the RCFs
Multiple linear regression is adopted to create the macroscopic relationships between the MCBs and the whole RCFs considering that the correlation between
Correlations between the RCFs and meteorological variables in the max-grid. The MCBs and individual RCF could be imprecise (Liao et al 2020). Focusing on those blue bars labeled as 'R^2-RCFs' as shown in figure 7(a), most evaporation indices do not have significant multiple linear correlations with the RCFs except for E3 (with the R^2 of 0.34), whereas in figure 7(b), all precipitation indices correlate with the RCFs significantly, especially for P3 (with the R^2 of 0.57). As to ST and T2M (see figures 7(c) and (d)), most of these indices can pass the significance test at the level of p < 0.05 or even p < 0.01, expect for ST4 and T2M4. In general, the results of multiple linear regression prove that the RCFs can have a less significant impact on E than on P; moreover, the LAAVs of both post-reservoir E and P sequences (i.e. E3 and P3) have strong correlations with the RCFs. ST and T2M have the similar moderate correlations with the RCFs.

The complexity of LSAC has been proven difficult to analyze (Shi et al 2016, 2017, Wang et al 2021), and the fact that the selected reservoirs in this study are dispersive globally and affected by different kinds of LSAC only makes the work even harder. In line with Approach (3), this study provides a side-contrast towards the correlations between the RCFs and the MCBs. The multiple linear regression between MCBs and geographic factors (i.e. latitude, longitude, and elevation) is analyzed (orange bars in figure 7). Then a comparison between blue bars (R^2-RCFs) and orange bars (R^2-geographic factors) is made. As shown in figure 7(a), for most evaporation indices, the R^2-RCFs are smaller than R^2-geographic factors, indicating that the impact of the RCFs on evaporation is less significant than that of geographic factors. For precipitation, opposite conclusion is found that the impact of the RCFs is more significant than that of geographic factors (see figure 7(b)). Regarding the two variables of temperature (i.e. ST and T2M) in figures 7(c) and (d), geographic factors are found to have more significant impact than the RCFs, which is more obvious for T2M. As inferred from the above results, compared to evaporation and temperature, precipitation is more sensitive to the RCFs rather than geographical factors.

3.3. Discussion

In this subsection, the reverse extrapolation to physical mechanism will be discussed from the approaches and corresponding results. Initially, this study attempts to find a general correlation pattern between the RCFs and the MCBs. Four meteorological variables (i.e. E, P, ST and T2M) are introduced and six indices are created for each of them. Ideally, significant linear correlations between several MCBs and RCFs are expected from the
Pearson correlation analysis, but the obtained results in sections 3.1 and 3.2 show that the highest absolute value of correlation coefficient is up to only 0.35, meaning that only a maximum of 35% of local climate change can be explained by the construction of dams and related reservoirs. It is not a satisfying correlation coefficient statistically. However, even though the correlation coefficients are relatively low, the correlation significance test has proved that more than half of such correlations are basically valid, and it is credible that some RCFs have up to 35% linear relationship with the MCBs.

Whereas not all correlations of the RCFs and the MCBs pass the significance test, only those correlations passing the significance test are selected for further analysis. Based on the distribution of figure-cells with ‘a’ and ‘b’ for different meteorological variables, it is found that: (a) the correlation pattern (with the RCFs) of evaporation and that of precipitation are overall opposite, that is, the RCFs which are significantly correlated with evaporation have barely any significant correlations with precipitation, and vice versa. This finding potentially provides a penetrating point to physical mechanism since it coincides to a great extent with several previous studies regarding evaporation and precipitation on wetland (Chen et al 2021), where soil moisture is deemed as the possible reason. Soil moisture content is often considered as the reserve of evaporation, and it determines the volume of total evaporation (Shi et al 2017, Zhou et al 2021). This correlation pattern is tenable since reservoir can be abstracted as a wetland with 100% soil moisture. As for precipitation, it largely depends on air temperature and humidity (Shi et al 2016, Dong et al 2020, Wang et al 2021). Since such correlation pattern is particularly obvious for CATCH which is correlated with all the precipitation indices and has barely any significant correlations with the evaporation indices, CATCH could be more significantly correlated with soil moisture, air temperature and humidity. (b) the correlation patterns of both temperature variables show that, the RCFs from a narrow spatio-temporal dimension seem to have less significant correlations with temperature variables than the RCFs from a relatively broader spatio-temporal dimension (especially for DIS and DOR). The influences of greenhouse gases (e.g. carbon dioxide and nitrous oxide) on the changes in temperature variables near the reservoirs could be the potential explanation. One assuming explanation is that the transformation and transmission of nitrogen, carbon, and oxygen elements (and related greenhouse gases) from water into atmosphere are large-scale water-atmosphere interaction (Tong et al 2019), and they significantly correlate with the RCFs from a
relatively broader spatio-temporal dimension such as DIS and DOR.

Then an innovative point of this study was brought out according to buffer zone theory. For each dam and related reservoir, the grids with the maximum and minimum meteorological data among buffer zone were selected to generate two new sequences. Then, the Pearson correlation analysis was conducted between the RCFs and the sequences of center-grid, max-grid, and min-grid, respectively. Known from the comparison results, the correlation patterns of some meteorological variables (e.g. precipitation) with the RCFs in the center-grid happen to be coincident with those in the max-grid, whereas the correlation patterns of other meteorological variables (e.g. temperature) with the RCFs in the center-grid happen to be coincident with those in the min-grid. This sparkling finding could potentially imply that, the RCFs can basically have an enhancing effect on precipitation and a weakening effect on temperature.

To find holistic correlations between the RCFs and the MCBs, this study conducts the multiple linear regression through joining the six RCFs as a whole. The results show that the RCFs seem to have a more significant correlation with precipitation whereas a less significant correlation with evaporation. The correlations between the RCFs and temperature variables are moderate. Furthermore, this study also conducts the multiple linear regression between the MCBs and geographical factors (i.e. latitude, longitude, and elevation) which represent LSAC, and compare it with the correlation between the MCBs and the RCFs. The results evidently show that the RCFs have relatively more significant correlations with precipitation than geographical factors, whereas for evaporation and temperature, the results are on the contrary. This finding indicates that, among all meteorological variables, precipitation may be the least sensible to LSAC, but the most susceptible to the construction of dams and related reservoirs. Such finding is slightly counterintuitive, since most prevailing studies on LSAC tend to believe that, compared to evaporation and temperature, precipitation is the most direct indicator of LSAC (Wei et al. 2021, Zhang et al. 2021, Zhou et al. 2021). The reason is that, when it narrows down to a given area with a reservoir being the center, the impacts of LSAC on precipitation could be overwhelmed by that from the reservoir. As for evaporation and temperature, a reasonable assumption could be that they are affected by some other meteorological factors such as the distribution of solar radiation and monsoon. Undoubtedly, further study is required for a more satisfying explanation, but such finding can be novel and insightful for future studies on the attributions of changing meteorological variables due to land surface change or LSAC, respectively.

Figure 7. Results of multiple linear regression.
4. Conclusions

Dams and related reservoirs are functioning indispensably both in human society and nature mainly by altering the characteristics of land surface and increasing the proportion of water surface. Based on the basic characteristics of 200 selected dams and related reservoirs globally and 40-year (i.e. 1981–2020) meteorological data for each dam according to the geographic coordinates, this study aims to explore the potential connection between the RCFs and meteorological variables (e.g. evaporation, precipitation and temperature, which are important indices for evaluation of hydrology extremes and climate change) at the global scale, and it is an important step to better comprehension of the physical mechanism behind the process of a reservoir affecting the MCBs. It is not the simple case that a reservoir with larger size would have more significant impacts on the MCBs because their relationships are rather complicated, and the existence of LSAC only makes it harder to isolate such relationships from noisy samples. Based on the objectives and approaches specially designed for this study, the major contributions of this study can be summarized as follows:

First, based on the results of Pearson correlation analysis, the correlations of the RCFs with evaporation indices are the opposite of those with precipitation indices. Both surface temperature and temperature 2 m above surface have more significant correlations with the RCFs from a relatively broader spatio-temporal dimension (i.e. CATCH, DIS and DOR). Second, comparisons between center-grid and max/min-grid indicate that dams and related reservoirs can have completely opposite (weakening or enhancing) effects on different meteorological variables. Third, the RCFs outperform the geographical factors in terms of the impact on precipitation, whereas the variation of evaporation is clearly more sensitive to geographical factors.

Overall, although no equation has been established to accurately describe the relationships between the RCFs and the MCBs, the results of correlation analysis is still effective in manifesting the potential patterns of how dams and related reservoirs can affect local climate change. Moreover, the sensitivity of the MCBs to the RCFs and geographical factors has been distinguished.

Limitations still exist in this study. For example, ‘noise’ generated by LSAC is the most complex problem during the whole period. The heterogeneity consumption and buffer zone theory introduced in this study may have addressed the stochastic distribution of spatial-frequency and attenuated the corresponding disparity, but the stochastic distribution of temporal-frequency (several LSAC are seasonal) is still difficult to address. Besides, the physical mechanism behind water–atmosphere process is hard to thoroughly analyze. Nevertheless, the approaches designed in this study can partly address the above questions, and the outcomes of this study still successfully explore the roles of large dams and related reservoirs in affecting local climate change, which will be useful and referable in the face of future hydrological extremes and climate change.

Acknowledgments

This study was supported by the National Natural Science Foundation of China funded project (Grant No. 51909117), Basic and Applied Basic Research Foundation of Guangdong Province funded project (Grant No. 2020A1515110642), Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control, and State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://cds.climateresearchdata.grid.copernicus.eu/cdsapp#!/search.

ORCID iD

Haiyun Shi https://orcid.org/0000-0001-5793-1138

References

Adeloye A J, Wumi J Y, Dau Q V, Soundharajan B S and Kasiviswanathan K S 2019 Height–area–storage functional models for evaporation-loss inclusion in reservoir-planning analysis Water 11 1413
Biemans H, Haddeland I, Kabat P, Ludwig F, Høgset D W A, Heinke J, Von Bloh W and Gerten D 2011 Impact of reservoirs on river discharge and irrigation water supply during the 20th century Water Resour. Res. 47 W03509
Bram H, Pierre-Erik I, Nadeau D F and Steeve P 2021 Can a physically-based land surface model accurately represent evapotranspiration partitioning? A case study in a humid boreal forest Agric. For. Meteorol. 304–305 108410
Chen J L, Wen J, Kang S C, Meng X H and Yang X Y 2021 The evapotranspiration and environmental controls of typical underlying surfaces on the Qinghai–Tibetan Plateau Sci. Cold Arid Reg. 13 53–61
Chen J, Shi H Y, Sivakumar B and Peart M R 2016 Population, water, food, energy and dams Renew. Sustain. Energy Rev. 56 18–28
Chen X Y, Zhang Q, Ye D X, Liao Y M, Zhu C H and Zou X K 2007 The characteristics of rainfall and temperature changes since the water storage of three Gorges reservoirs Agric. Meteorol. 304–305 108410
Cheng Y, Mu Z, Wang H Y, Zhao F X, Li Y and Lin L 2019 Water residence time in a typical Tributary Bay of the three Gorges reservoir Water 11 1585
Chi Q, Zhou S H, Wang L J, Zhu M Y, Liu D D, Tang W C, Cui Y P and Lee J 2021 Exploring on the eco-climatic effects of land use changes in the influence area of the Yellow River Basin from 2000 to 2015 Land 10 601
Dong Z Y, Peng T, Dong X H, Liu J, Chang W J and Lin Q X 2020 Spatiotemporal variation characteristics of extreme precipitation events in the three Gorges reservoir area during 1960–2016 J. Water Resour. Water Eng. 31 93–101
Dos Santos T, KeppelAleks G, De Roo R and Steiner A L 2021 Can land surface models capture the observed soil moisture control of water and carbon fluxes in temperate-to-boreal forests? J. Geophys. Res. Biogeosci. 126 e2020G005999

Food and Agriculture Organization of the UN (FAO–UN) 2011 Major hydrological basins of the world (available at: https://data.apps.fao.org/map/catalog/srv/api/records/7707086-dafsc±41cc-a8a5-323db069b2d1 (accessed 25 August 2021))

Guo Y, Huang Q, Huang S Z, Leng G Y, Zheng X D, Fang W, Deng M J and Song S B 2021 Elucidating the effects of mega reservoir on watershed drought tolerance based on a drought propagation analytical method J. Hydrol. 598 125738

Kaveh K, Hosseinianzadeh H and Hosseini K 2013 A new equation for calculation of reservoir’s area-capacity curves KSCE J. Civil Eng. 17 1149–56

Keigo N, Junya H, Masaomi K and Kazuo O 2018 Debates over the cold region in Northeast China still getting warmer under climate change impact? Atmos. Res. 202 104043 Y Zhao et al

Zaidel P A, Roy A H, Houle K M, Lambert B, Letcher B H, Yang H, Wang J H, Xiao W H, Lu F, Wang Y and Jarso J 2020 Relationship between hydroclimatic variables and reservoir wetland landscape pattern indices: a case study of the Sanmenxia Reservoir wetland on the Yellow River, China J. Earth Syst. Sci. 129 1–13

Yu S H, Li S, Zhan L K and Nilufo Z J 2020 Analysis of temperature variation characteristics in the three gorges reservoir area after impoundment of the three gorges dam Acta Ecol. Sin. 41 384–9

Yun X B, Tang Q H, Li J B, Lu H, Zhang L and Chen D L 2021 Can reservoir regulation mitigate future climate change induced hydrological extremes in the Lancang-Mekong River Basin? Sci. Total Environ. 785 147322

Zhang L, Liu Y F, Jin M G and Li X 2021 Spatiotemporal variability in extreme temperature events in an arid-semiarid region of China and their teleconnections with large-scale atmospheric circulation J. Earth Sci. in press (https://kns.cnki.net/kcms/detail/42.1788.P.20210720.0945.004.html)

Zheng Y H, Wang L, Fu Z Y, Peng Z R and Chen C 2021 Using GRACE and hydrological data to estimate changes of evapotranspiration in the three gorges reservoir Earth. Sci. Envi. 660 012093

Zhou Z Q, Shi H Y, Fu Q, Su Y B, Li T X, Yang W and Liu S N 2021 Characteristics of propagation from meteorological drought to hydrological drought in the Pearl River Basin J. Geophys. Res. Atmos. 126 e2020JD033859

Zhou Z Q, Shi H Y, Fu Q, Li T X, Gan T Y, Liu S N and Liu K 2020 Is the cold region in Northeast China still getting warmer under climate change impact? Atmos. Res. 237 104864