

Increase of Future Summer Rainfall in the Middle and Lower Reach of the Yangtze River Basin Projected With a Nonhomogeneous Hidden Markov Model

Lianyi Guo, Zhihong Jiang, Laurent Li, Huijun Wang

▶ **To cite this version:**

Lianyi Guo, Zhihong Jiang, Laurent Li, Huijun Wang. Increase of Future Summer Rainfall in the Middle and Lower Reach of the Yangtze River Basin Projected With a Nonhomogeneous Hidden Markov Model. Geophysical Research Letters, 2022, 49, 10.1029/2021GL097325 . insu-03726908

HAL Id: insu-03726908

https://insu.hal.science/insu-03726908

Submitted on 19 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Check for updates

AGU ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters®

RESEARCH LETTER
10.1029/2021GL097325

Key Points:

- Nonhomogeneous hidden Markov model can well capture physical relations between regional rainfall and large-scale atmospheric circulation
- The middle and lower reach of the Yangtze River basin exhibits a general wetness for 1.5°C and 2°C global warming targets, with smaller (larger) increase for the west (east)
- Increases in the occurrence frequency of relevant synoptic weather patterns lead to more precipitation in the region

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Z. Jiang,
zhjiang@nuist.edu.cn

Citation:

Guo, L., Jiang, Z., Li, L., & Wang, H. (2022). Increase of future summer rainfall in the middle and lower reach of the Yangtze River basin projected with a nonhomogeneous hidden Markov model. *Geophysical Research Letters*, *49*, e2021GL097325. https://doi.org/10.1029/2021GL097325

Received 6 DEC 2021
Accepted 12 MAR 2022

© 2022. American Geophysical Union. All Rights Reserved.

Increase of Future Summer Rainfall in the Middle and Lower Reach of the Yangtze River Basin Projected With a Nonhomogeneous Hidden Markov Model

Lianyi Guo[1,2], Zhihong Jiang[2], Laurent Li[3], and Huijun Wang[1,2,4]

[1]Climate Change Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [2]Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters/Key Laboratory of Meteorological Disasters, Ministry of Education/Joint International Research Laboratory of Climate and Environment Change, Nanjing University of Information Science and Technology, Nanjing, China, [3]Laboratoire de Météorologie Dynamique, CNRS, Sorbonne Université, Ecole Normale Supérieure, Ecole Polytechnique, Paris, France, [4]Nansen-Zhu International Research Centre, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Abstract Among the statistical downscaling tools available for regional climate simulation, the nonhomogeneous hidden Markov model (NHMM) is a powerful and efficient algorithm. It consists of establishing statistical relations between predictand and predictors through a hidden layer of Markov process and a covariate nonhomogeneous term as external forcing. Here, it is used to simulate summer daily precipitation in the middle and lower reach of the Yangtze River basin (MLYRB), including future projection. Results show that NHMM well identifies large-scale circulation features that are physically consistent with regional rainfall patterns. MLYRB exhibits a general wetness for 1.5°C and 2°C global warming targets, with smaller (larger) increase for the west (east). Such changes correspond to increases in the occurrence frequency of synoptic weather patterns with stronger and more westward Western Pacific Subtropical High and stronger westerly jet. This contributes to moistening MLYRB with increasing occurrence frequency for those wetter rainfall patterns.

Plain Language Summary Statistical downscaling has been widely used in regional climate simulation and projection. It is generally based on statistical relations between regional climate (predictand) and large-scale atmospheric circulations (predictors). We have to, furthermore, assume that the relations found and established in present-day climate would remain valid for future climate. There are needs and requirements for the statistical tools to find good and robust relations between the predictand and predictors. In this study, we demonstrate that the nonhomogeneous Markov model (NHMM) is an efficient algorithm and can well capture physically consistent relations between regional rainfall patterns in the middle and lower reach of the Yangtze River basin and large-scale atmospheric circulation features around the region. We believe that NHMM is a confidence-inspiring tool for the regional climate simulation and projection. Finally, our study reveals a general increase in precipitation in the region for a future warm climate for both 1.5°C and 2°C global warming targets, and this projection scenario can be well explained by the increase in the occurrence frequency of the synoptic weather patterns with a stronger and westward-shifted Western Pacific Subtropical High as well as a stronger westerly jet in the future.

1. Introduction

The middle and lower reach of the Yangtze River basin (MLYRB) in Eastern China is located in the East Asian monsoon region. It suffers frequent meteorological and climatic disasters causing serious losses of lives and properties. It is of high significance to investigate future changes of its regional climate under the background of global warming.

Global climate models (GCMs) are primary tools and source of information to make future climate change projection. They particularly excel in reproducing large-scale atmospheric circulations. Although models from the Coupled Model Intercomparison Project Phase 6 (CMIP6) are shown to be more skillful than their counterpart in the precedent phase CMIP5 (Dong & Dong, 2021; Xin et al., 2020; Zhu et al., 2020; Zhu, Jiang, & Li, 2021; Zhu, Lee, et al., 2021), they still have important biases in simulating regional climate that can be well described if we have a good observation network covering the region and for a time period long enough to deduce

a reliable climatology. A fruitful approach is thus to explore both advantages offered by GCMs for large-scale circulations and by local weather stations for regional climate. In practice, we need to establish statistical relations between the large-scale circulations and regional climate with present-day conditions. Such relations can then be, subsequently, applied to future scenarios produced by GCMs. We can easily see that this methodology is conceptually very close to any statistical downscaling techniques designed to transform simulations of GCMs to reliable regional climate. It is clear that we have to assume that the relations found in present-day climate would remain valid for future climate. This assumption can be a fragile point of the whole methodology, since there are no, a priori, physical laws used in the approach to explore and explain future climate change (Ehret et al., 2012; Fan et al., 2021). In this study, we attempt to remedy this issue. Our basic idea is to check the physical consistency of the statistical relations, which can inspire more confidence.

Let us take a simple statistical downscaling model as an example, such as the multivariate linear regression that allows us to establish a relationship between present-day observed large-scale circulations (predictor) and meteorological variables (predictand) to simulate future regional climate. This statistical relationship may not clearly describe any valid physical relations in the complex climate system, which consequently does not inspire confidence. In recent literatures, however, more and more studies were reported with more advanced statistical downscaling schemes to perform climate downscaling (Cioffi et al., 2017; Li et al., 2020; Yin et al., 2011).

In our recent study (Guo et al., 2019), we used the nonhomogeneous hidden Markov model (NHMM) to construct a relationship between large-scale circulation patterns (predictor) and the probability distribution of regional precipitation (predictand) (Hughes & Guttorp, 1994). It continuously learns the evolution characteristics of synoptic weather situations over time, so that the statistical property of meteorological variables can be reproduced more comprehensively and accurately (Cioffi et al., 2017; Fu et al., 2013; Greene et al., 2011; Robertson et al., 2004; Tan et al., 2013). NHMM comprises of an input layer (predictor), a hidden layer, and an output layer (predictand). Our NHMM work focused on rainfall simulation over MLYRB. Here, we pursue this investigation to physical mechanisms. We try to focus on the hidden layer, explore the relations between the hidden states representing relevant rainfall patterns, and atmospheric circulation regimes in the statistical model, and finally provide an explanation for future changes in predictand.

2. Data and Methodology

2.1. Data Sets

We selected daily rainfall records in summer (June–August) during 1961–1990 at 56 stations in MLYRB (27.5–32°N, 110–122°E) from the China Meteorological Administration network. Daily atmospheric circulation fields during 1961–1990 are from ERA-40 reanalysis with 2.5° × 2.5° resolution from the European Centre for Medium-Range Weather Forecasts. They are used to establish and calibrate the NHMM.

Four GCMs from CMIP6 used in our NHMM are EC-Earth3 (0.7° × 0.7°) from the EC-Earth consortium, GFDL-CM4 (2.5° × 2.0°) from National Oceanic and Atmospheric Administration, IPSL-CM6A-LR (2.5° × 1.3°) from French Pierre Simon Laplace Institute, and MPI-ESM1-2-HR (0.9° × 0.9°) from Germany Max Planck Institute. They all show good performance in simulating precipitation in China and large-scale atmospheric circulation around the region (Dong & Dong, 2021; Xin et al., 2020; Zhu, Jiang, & Li, 2021; Zhu, Lee, et al., 2021). The emission scenario is SSP5-8.5. Large-scale atmospheric variables under the global warming of 1.5°C and 2°C are applied into NHMM to obtain future climate projection. For details of data processing, please refer to Ding et al. (2016) and Guo et al. (2019).

2.2. Nonhomogeneous Hidden Markov Model (NHMM)

NHMM is a double stochastic process consisting of hidden states that cannot be directly found and an observed state sequence (Hughes et al., 1999; Hughes & Guttorp, 1994). It makes two conditional assumptions. The first states that the multivariate precipitation R_t on day t is conditionally independent of all other variables, given the hidden state S_t on day t. The second assumption says that the hidden state S_t on day t depends only on the predictor vector X_t on day t and the hidden state S_{t-1} on day $t-1$.

$$P(R_t|S_{1:T}, R_{1:t-1}, X_{1:T}) = P(R_t|S_t) \quad (1)$$

$$P(S_t|\, S_{1:t-1}, X_{1:T}) = P(S_t|\, S_{t-1}, X_t) \tag{2}$$

where R_t and S_t represent observed precipitation and the hidden states of precipitation (also called rainfall patterns, hereafter) on day t ($t = 1, 2\ldots T$), respectively. Both are defined at W stations over the study area, that is, 56 for MLYRB. T is the time sequence length in days, that is, 92 for the boreal summer. X_t represents atmospheric circulation field on day t and thus $X_{1:T} = (X_1, X_2\ldots X_T)$ denotes the time series of circulation field from $t = 1$ to $t = T$.

The number of rainfall patterns is noted as N and determined by optimizing the Bayesian Information Criterion (BIC).

$$BIC_N = 2L\left(\Theta_N^*\right) - p\log T \tag{3}$$

where Θ_N^* is the estimated maximum likelihood parameter vector as found through the expectation maximization algorithm on the training data for a model with N states. p is the number of parameters in the N-state model. Normally, the least BIC score corresponds to an optimal model that greatly interprets the training data. In this paper, the optimal number of rainfall patterns is $N = 8$.

Our previous work (Ding et al., 2016; Guo et al., 2019) shows that with four large-scale predictors, including sea level pressure (12.5–35°N, 105–120°E), geopotential height at 500 hPa (10–25°N, 95–170°E), relative humidity at 500 hPa (27.5–32.5°N, 105–125°E), and zonal wind at 500 hPa (25–32.5°N, 95–140°E), we are able to establish a good NHMM simulating summer daily precipitation over MLYRB. In practice, we apply the principal component analysis on joint field of the four predictors to filter out noises. The first 26 principal components corresponding to the cumulative explained variance of 90% are selected as actual predictors to establish the model.

2.3. Time Windows for 1.5°C and 2°C Global Warming Targets

The 1.5°C and 2°C thresholds are relative to the pre-industrial period (1861–1900). Time series of the global averaged temperature anomaly for each GCM is first smoothed by a 21-year moving average. We then select the year for which the 1.5°C or 2°C threshold is first reached. Finally, a period of 21 years in total is defined with 10 years before and after the nominal year, as shown in Table S1 of Supporting Information S1. It is noted that future changes in precipitation at the two warming targets are relative to the reference period (1986–2005).

3. Results

3.1. Relations Between Rainfall Patterns and Circulation Regimes

In our case, there are eight rainfall patterns that were optimized by BIC through its calibration with station-originated precipitation and reanalysis. We further investigate their corresponding circulation regimes. In practice, these circulation regimes are just a composite of synoptic conditions comprised in each rainfall pattern (i.e., the hidden states of the NHMM).

The eight rainfall patterns obtained by NHMM are shown in Figure 1 (left panels) in the form of anomalies or deviations from their average. The rainfall patterns 2, 3, and 4 (P2, P3, and P4) reveal spatial structures with deficit precipitation almost over whole MLYRB, with only 7, 1, and 6 stations of weak opposite signal, while P7 shows surplus precipitation all over, except 6 stations with weak deficit. The remaining rainfall patterns are transitional ones, including P1 with positive deviation in the south and negative deviation in the north, P5 with above-normal in the east and below-normal in the west, and P8 with surplus in the west and deficit in the east. P6, however, shows only insignificant deviation from the normal all over.

We can remark that, for all P2, P3, P4, and P6 with deficit rainfall or relatively insignificant deviation from normal (middle panels in Figure 1), the Western Pacific Subtropical High (WPSH) is all eastward retreated (142°, 144°, 136°, and 143°E, respectively). For P2 and P6, a wave train at 850 hPa is emanated from the equatorial Pacific (Figures 1b2 and 1f2). For P4, there is an anomalous cyclone at 850 hPa over Taiwan (Figure 1d2). These configurations lead to weaker summer monsoon. For P3, our study area is controlled by an anomalous anticyclone at 850 hPa (Figure 1c2). Right panels in Figure 1 display the zonal wind and meridional overturning circulation averaged along 110°–122°E over MLYRB (27.5–32°N). We can observe air descending on the whole layer over

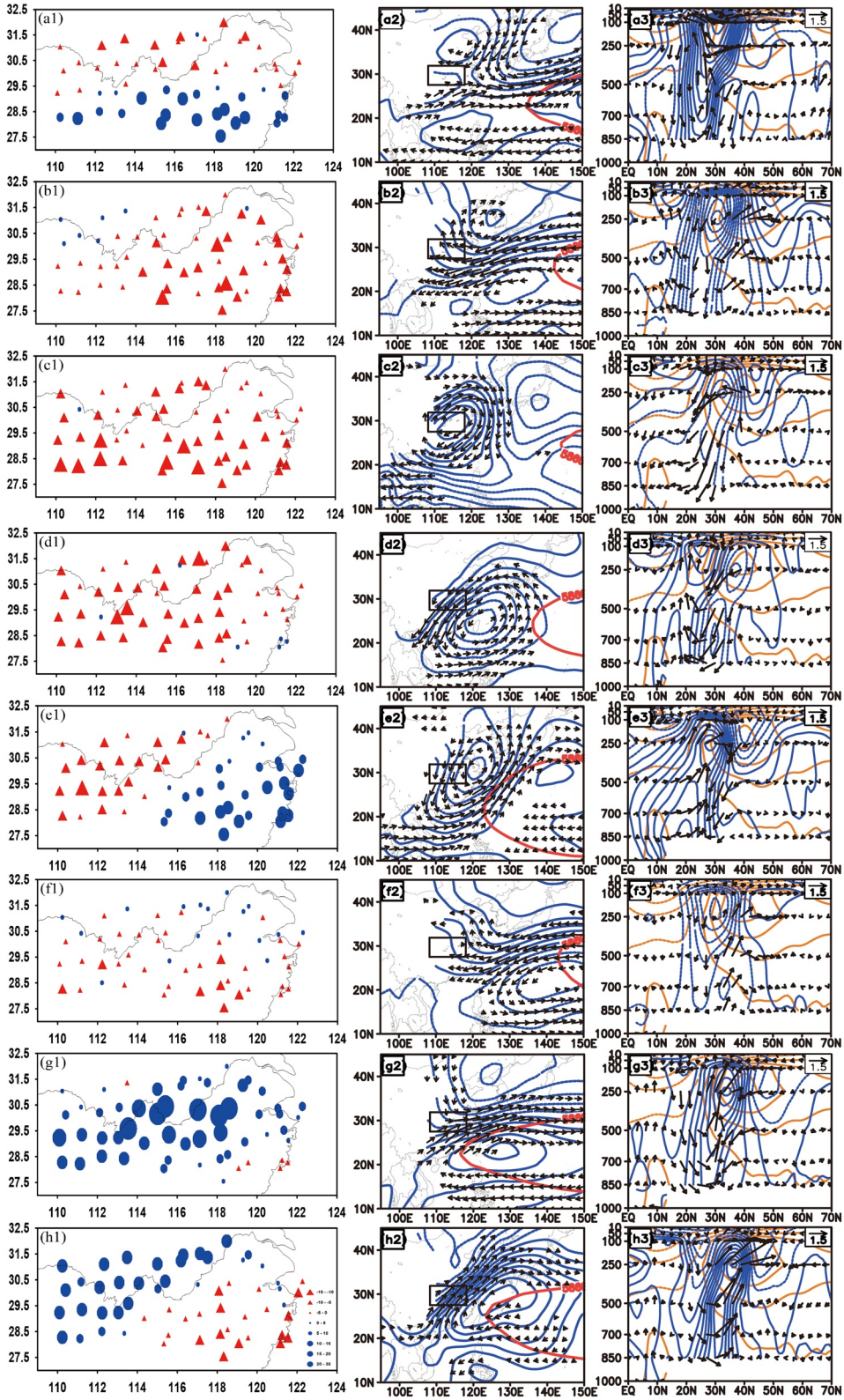


Figure 1.

MLYRB for the four patterns. Except for P4, the westerly jet is northward shifted (50°, 44°, and 49°N for P2, P3, and P6, respectively) compared to climatology. This circulation configuration is not favorable for precipitation.

For P7, WPSH extends westward (116°E), leading to an anomalous anticyclone at 850 hPa (Figure 1g2). There is an anomalous convergence at 850 hPa with water vapor transport from the South China Sea. The vertical profile shows air ascending on the whole layer over MLYRB (Figure 1g3). The westerly jet is southward shifted (33°N) and stronger compared to climatology (Figure S2 in Supporting Information S1). This circulation configuration is favorable to receive more rain.

For P1, WPSH is located around 135°E which is close to the climatology (Figure 1a2). There is an anomalous anticyclone at 850 hPa over the Philippines. A wave train at 850 hPa occurs in the middle and high latitudes, with an anomalous anticyclone in North China. This configuration is able to generate stronger moisture convergence. The zonal-vertical profile shows air ascending in the south of 30°N and air descending in the north of 30°N (Figure 1a3). The westerly jet is southward shifted (30°N) compared to the climatology. This circulation configuration is favorable to wetness in the south and dryness in the north.

For P5, WPSH extends westward to the southeast side of MLYRB (123°E), leading to an anomalous cyclone at 850 hPa in eastern MLYRB (Figure 1e2). There are regions with anomalous convergence in the east and anomalous divergence in the west at 850 hPa (Figure 1e3). This circulation regime, visibly, brings wetness to the east and dryness to the west.

Finally, for P8, WPSH extends westward (123°E) with an anomalous anticyclone at 850 hPa in the east of MLYRB and leads to stronger summer monsoon across the western MLYRB (Figure 1h2). The westerly jet is southward shifted (35°N) (Figure 1h3). This circulation configuration is favorable for wetness in the west and dryness in the east.

To sum up, all the rainfall patterns produced by NHMM are well consistent with their corresponding large-scale circulation regime. It is clear that our NHMM is an efficient statistical downscaling algorithm with excellent capacity in climate diagnosis (Cioffi et al., 2016, 2020; Fu et al., 2013; Khalil et al., 2010; Pineda & Willems, 2016; Robertson et al., 2004). It is worthy of note that NHMM, although purely statistical, has the excellent capacity of "learning" from station-based rainfall observation and reanalysis data of atmospheric circulation, and digesting them to produce physically consistent behaviors and relations.

3.2. Future Projection of Precipitation and Its Possible Cause

Our previous paper (Guo et al., 2019) showed that NHMM effectively enhances the ability of CMIP5 GCMs in simulating the spectral distribution and spatial pattern of summer precipitation in MLYRB in eastern China. We also drew the same conclusion by conducting an independent validation of NHMM with CMIP6 GCMs during 1991–2010 (Figure S1 in Supporting Information S1). In Section 3.1 of the present work, we furthermore demonstrated that NHMM has an excellent capacity in capturing the relation between regional rainfall patterns and atmospheric circulation regimes. It explains why NHMM has an efficient capacity in climate simulations. We are thus quite confident now to apply the established NHMM to projection of future precipitation under the global warming of 1.5°C and 2°C.

It is to be recalled that the precipitation R (output layer) is jointly determined by the hidden states S (hidden layer) and the covariate large-scale predictors X (input layer). For covariates as predictors, we use the first 26 principal components of the joint field from four geophysical variables. NHMM assumes that these rainfall patterns and the spatial structures of the first 26 principal components of predictors remain unchanged in the future, but their occurrence frequencies and intensities can vary over time. This allows us to explore the underlying physical mechanisms of precipitation changes from the perspective of changes in rainfall patterns and predictors.

Figure 1. Hidden states or rainfall patterns produced by nonhomogeneous hidden Markov model (a1–h1; unit: mm/day) in the form of deviation from their average and their corresponding large-scale anomalous circulations (a2–h2; a3–h3). The second column (a2–h2), the black rectangular depicts the middle and lower reach of the Yangtze River basin. The red bold line (586gpm) indicates the position of the Western Pacific Subtropical High; the blue contours show geopotential height anomaly at 850 hPa (unit: hPa). Wind anomaly at 850 hPa (vector; unit: m/s) with a speed larger than 0.5 m/s is plotted. The third column (a3–h3), zonal-vertical profile averaged along 110°–122°E. Zonal wind climatology (yellow contour), zonal wind anomaly (blue contour), and wind anomaly (vector) are plotted.

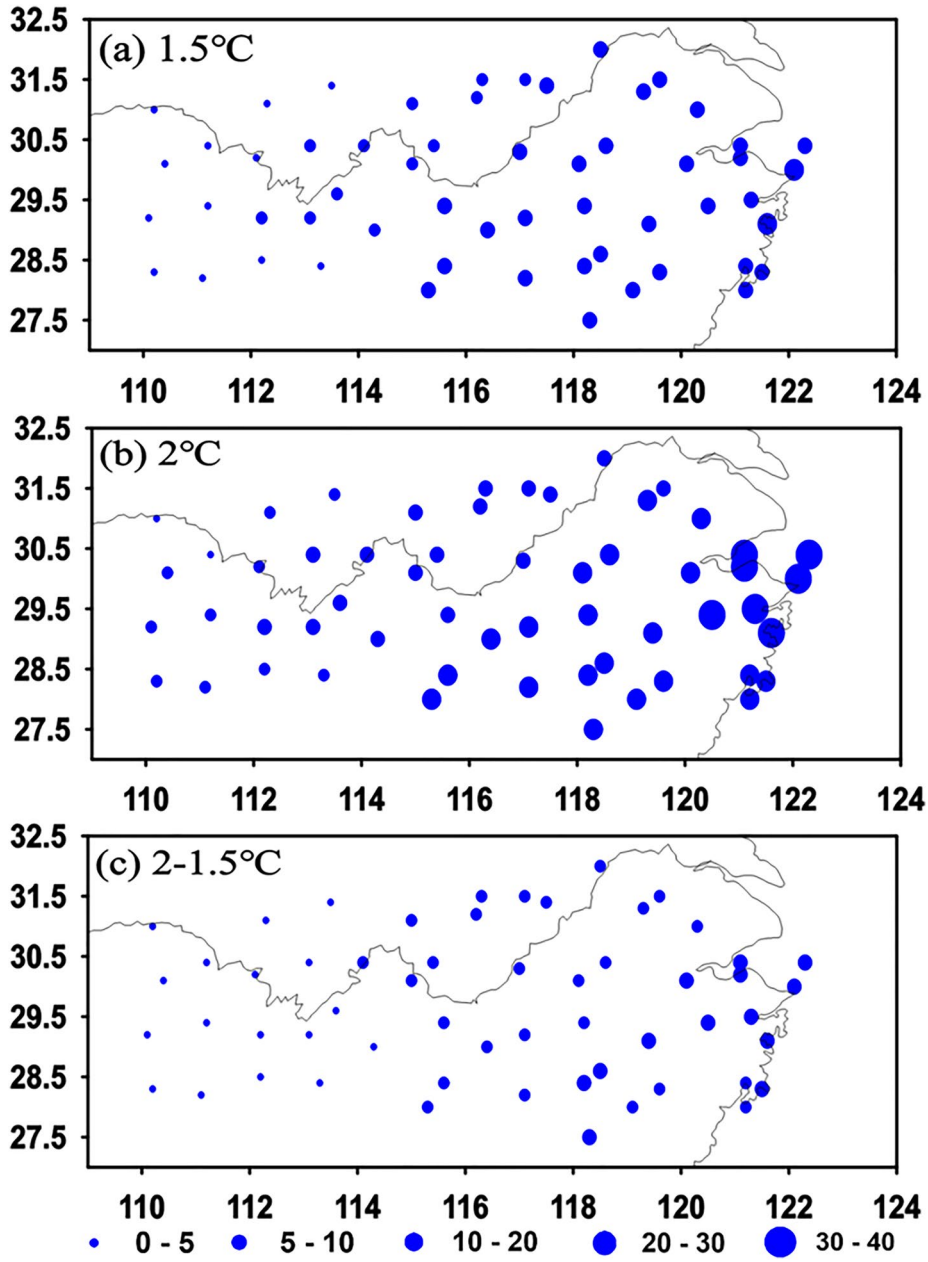


Figure 2. Relative changes of total precipitation under the global warming of 1.5°C (a), 2°C (b), and half-degree warming from 1.5 to 2°C (c), relative to 1986–2005 (unit: %). Yellow cross represents the station that passes the 95% significance test.

3.2.1. Changes in Precipitation Geographic Distribution

Figure 2 presents the spatial distribution of total precipitation changes in the multi-model ensemble mean at the two warming targets, relative to 1986–2005. Results show that total precipitation increases at all stations, with wetter conditions in the east than in the west. The additional half-degree warming from 1.5°C to 2°C induces more pronounced wetness without passing the 95% significance test at all stations, and the difference between the west and the east tends to amplify. Specifically, total precipitation increases with relative changes keeping below 20% (40%) in the east and 8% (10%) in the west at the 1.5°C (2°C) warming target. It is noted that compared with raw GCMs, NHMM makes a minor adjustment on the amplitude and spatial patterns of total precipitation changes (Figure S3 in Supporting Information S1). Uncertainties of projection among the four GCMs at all 56

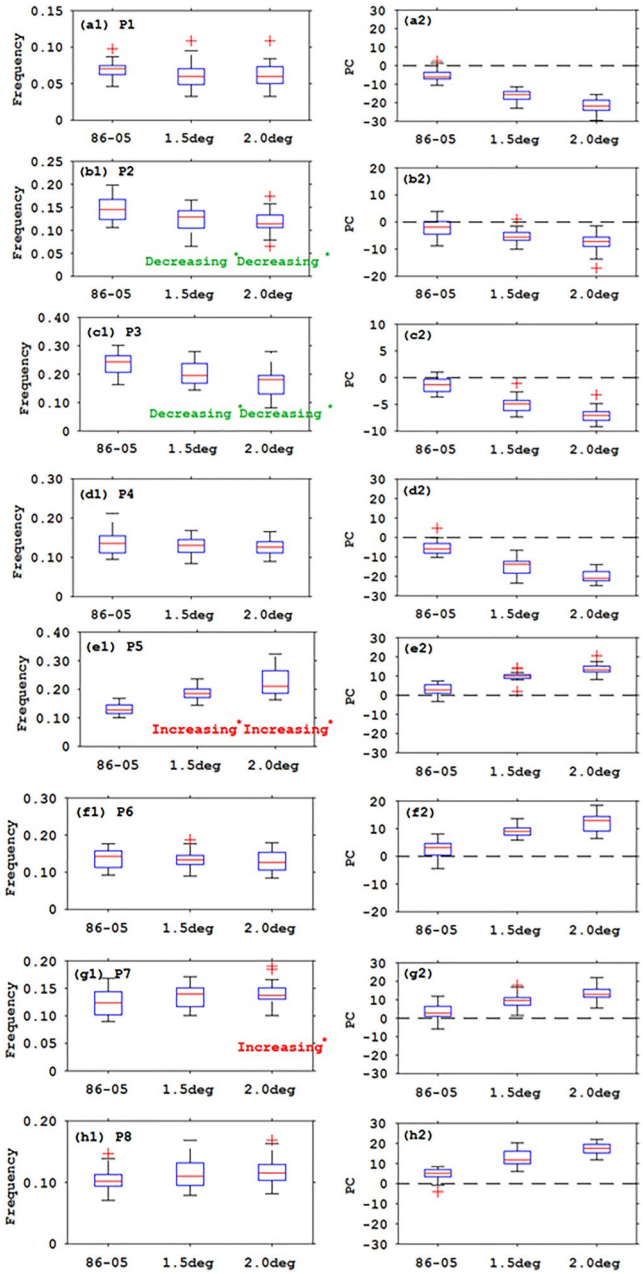


Figure 3. Boxplots showing occurrence frequency for eight rainfall patterns (a1–h1) and projection coefficients of corresponding circulation fields for eight rainfall patterns (a2–h2) during the three time periods (historical and 2 global warming targets). "Increasing*" and "Decreasing*" indicate the increasing and decreasing trend that passes the 95% significance test using Student's *t*-test, respectively.

stations decrease after the application of NHMM (Figure S4 in Supporting Information S1). It demonstrates that NHMM can enhance the credibility of climate projection.

Actually, our previous work also produced a future projection, but with CMIP5 models (Guo et al., 2019). The result there is close to the present work using CMIP6 models, that is, MLYRB exhibits a general wetter trend for its future precipitation. The only difference is in western MLYRB, showing a weak decrease with CMIP5 and a slight increase with CMIP6.

It is to be noted that many results reported in the literature using GCMs or statistical methods had similar conclusions. Guo et al. (2011) used a multivariate linear regression model to simulate summer precipitation in MLYRB, with driving forcing from multiple GCMs. The downscaled result provided a clearer increase in rainfall in MLYRB, compared to the original GCMs. Pan et al. (2016) showed that all categories of precipitation in MLYRB are projected to increase in frequency using 31 GCMs from CMIP5. Han et al. (2014) showed that the intensity and the occurrence frequency of heavy rainfall events are projected to increase more rapidly in the east than in the west using eight GCMs from CMIP5. However, Gu et al. (2014) indicated that most of MLYRB stands out as "hotspots" of climate change in China at the end of the 21st century, with an increase in annual precipitation in northern MLYRB and a decrease in southern MLYRB using the regional climate model, RegCM4.0. In a word, although models and methods used could have an impact on the regional precipitation projection in MLYRB, our results are basically consistent with most other works reported in the literature, projecting wetter conditions.

3.2.2. Response of Rainfall Patterns and Their Corresponding Circulation Regimes

The implementation of our NHMM can record the weather situation each day and label it with one of the rainfall patterns. We can thus calculate the occurrence frequency of each rainfall pattern and explore their significance in simulating rainfall during the three time periods (historical and 2 global warming targets) (left panels in Figure 3). We further investigate occurrence frequencies of their corresponding circulation regimes for the eight rainfall patterns. To quantitatively measure their occurrence, we can project (formally, a dot product of two vectors or matrices) the composite structure onto each actual synoptic situation. Hereafter, we will use the term "projection coefficient (PC)" to designate this dot product. Results are now shown in Figure 3 (right panels) to assess the variations of their PCs among the three periods.

Left panels in Figure 3 provide the annual occurrence frequency for the eight rainfall patterns during the three time periods (historical and 2 global warming targets). Results show that occurrence frequencies of P5, P7, and P8 increase, while those of P1, P2, P3, P4, and P6 decrease. We further apply Student's t-test to the occurrence frequency for each rainfall pattern. It shows that only P2 and P3 with strong below-normal precipitation for the whole domain significantly decreases for the two warming targets relative to the historical period. The decrease passes the 95% significance test. On the contrary, P5 with wetness in the east and dryness in the west at the two warming targets, and P7 with above-normal deviation of precipitation over the whole region at the 2°C warming target significantly increases at 95% of the significance test.

Right panels in Figure 3 provide boxplots showing the distribution of PC for the eight rainfall patterns during the three time periods. We find that future changes of PC are consistent with those of the corresponding occurrence

frequency, except for P6. Specifically, PCs of the corresponding circulation regimes for P1, P2, P3, and P4 (P5, P6, P7, and P8) are almost negative (positive) and decreases (increases) with time. It means that occurrence frequencies of their corresponding circulation regimes will decrease (increase) in the future.

To sum up, considering the good consistence between rainfall patterns and large-scale circulation regimes, we can draw the conclusion that the occurrence frequency of synoptic weather patterns with stronger and more westward WPSH as well as stronger westerly jet will increase in the future.

3.2.3. Role of the Nonhomogeneous Term in NHMM

The nonhomogeneous term in our NHMM, also called covariable term or input layer X, plays an important role in controlling the behaviors of the Markov process. It can be considered as the large-scale predictor coming from GCMs. We now assess their changes to further understand the underlying physical mechanisms leading to rainfall changes in MLYRB. As we mentioned previously, the first 26 principal components of the joint fields of four predictors are used to establish the NHMM. It is thus useful to examine the leading EOFs and the associated principal components.

Figure 4 exhibits the first two principal components and their associated regression fields (spatial pattern). The principal components cover the historical period 1986–2005 and two periods under the global warming of 1.5°C and 2°C. The explained variance of the first principal component is 33.0% (Figure 4a). The high-value center of geopotential height at 500 hPa over the South China Sea easily brings water vapor to MLYRB. Relative humidity anomaly at 500 hPa also increases. Therefore, the first spatial mode of the predictors shows beneficial situations to increasing precipitation over MLYRB. The first principal component's time coefficient is almost positive and increases with time (Figure 4b). It contributes to rainfall abundance in the future.

For the second principal component that explains 10.1% of the total variance (Figure 4c), the relative humidity at 500 hPa obviously shows negative values in the east and positive values in the west, but the corresponding principal component is almost negative and decreases with time (Figure 4d). By consequence, the second mode brings significant wetness in the east and slight wetness in the west in future.

In a word, the above two synoptic weather patterns that occur more frequently at 1.5°C target contribute to a general wetter trend over MLYRB, remarkable wetness in the east, and modest wetness in the west. These trends would be more pronounced at the 2°C global warming target.

4. Conclusions and Discussion

A statistical scheme based on NHMM was constructed and used to project responses of summer rainfall in MLYRB to the global warming of 1.5°C and 2°C. The driving forcing is provided by four GCMs from CMIP6, that is, EC-Earth3, GFDL-CM4, IPSL-CM6A-LR, and MPI-ESM1-2-HR. We further explore the causes of precipitation changes from the perspective of changes in rainfall patterns and predictors over time. Main conclusions are as follows.

1. NHMM has an excellent capacity in capturing the relationship between the rainfall patterns and their corresponding regional atmospheric circulation regime.
2. MLYRB exhibits a general wetter climate in the future, with modest increase in precipitation for the west and larger increase for the east at the two warming targets. The additional half-degree warming from 1.5°C to 2°C induces more pronounced wetness, and the difference between the west and the east tends to amplify.
3. Such rainfall changes correspond to increases in the occurrence frequency of these synoptic weather patterns with stronger and more westward WPSH as well as stronger westerly jet. This constitutes favorable conditions to bring more rain over MLYRB, and thus leads to an increase in the occurrence frequency of the rainfall patterns with more precipitation.

In our work, NHMM makes the assumption that the rainfall patterns and the spatial structures of predictors remain unchanged in future, but their occurrence frequencies and intensities can vary over time. We demonstrate that NHMM well captures the relationship between local rainfall patterns and large-scale atmospheric circulation

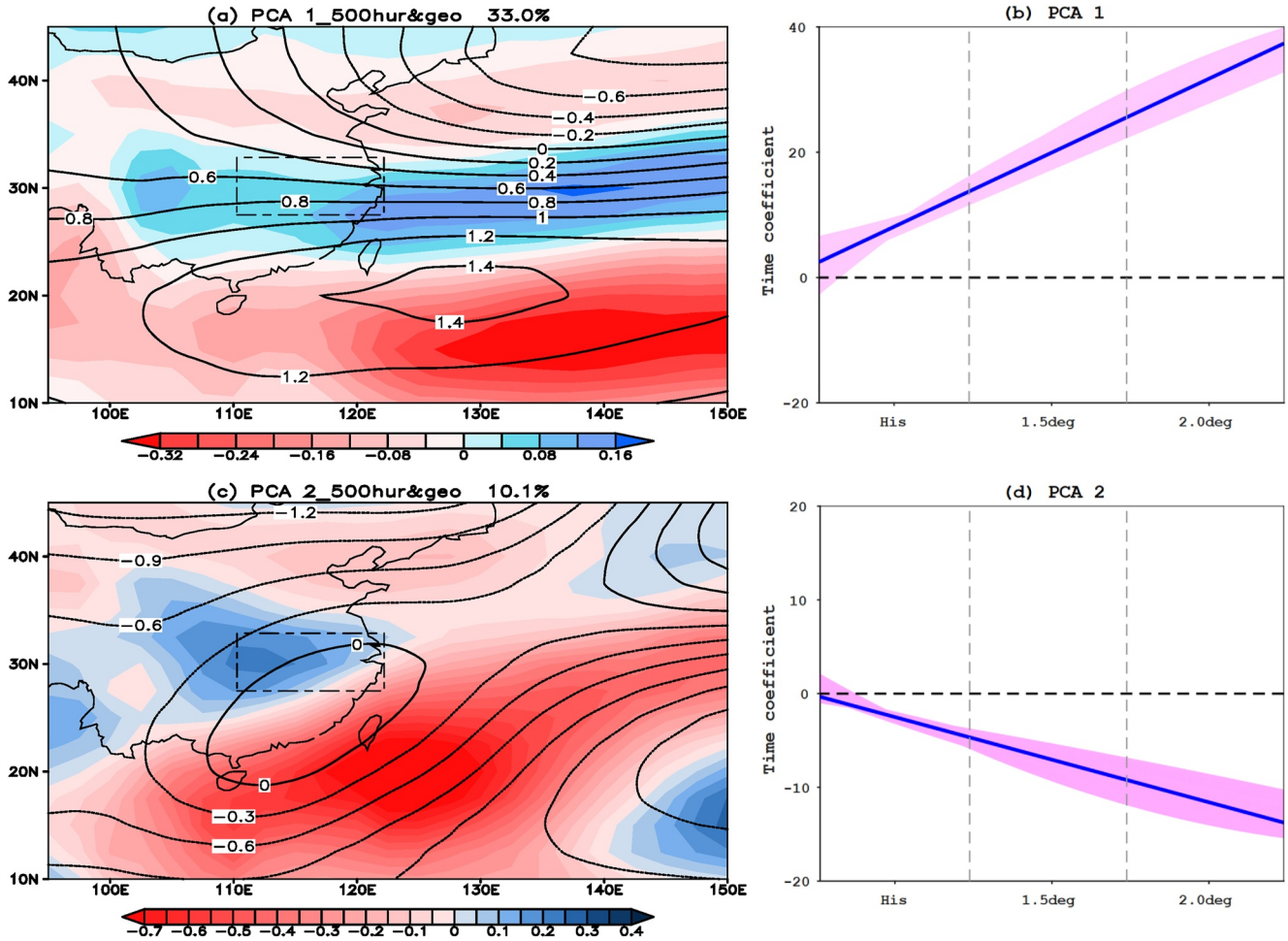


Figure 4. Regression fields of the first two spatial modes of the 26 principal components (a and c) and their time coefficients during the three time periods (historical and 2 global warming targets) (b and d). The predictors of the left picture include relative humidity at 500 hPa (shaded) and geopotential height at 500 hPa (black contour). The gray dashed rectangle represents the middle and lower reach of the Yangtze River basin. For the right picture, the blue solid line corresponds to the trend of multimodel ensemble mean. Red shadow indicates the difference among models.

features, reflecting the excellent capacity of NHMM in climate diagnosis. This is the reason why NHMM generally shows good performance in regional climate simulations. We believe that what we showed in this study can inspire confidence for future climate projection performed with NHMM. It also provides an opportunity to explain causes of future changes.

Finally, some alternative NHMM approaches and more advanced machine-learning algorithms, as abundantly reported in the recent scientific literature, may lead to further improvement in regional climate simulations (Bulla et al., 2010; Holsclaw et al., 2016; Irrgang et al., 2021; Long et al., 2019; Wang et al., 2021). Further investigations are worthy of exploration in the future.

Data Availability Statement

The practical realization of nonhomogeneous hidden Markov model is through the toolbox "HMMTool" developed and maintained in the International Research Institute (https://iri.columbia.edu/our-expertise/climate/tools/hidden-markov-model-tool/hmmtool/). Daily rainfall records are available at http://data.cma.cn/. ERA40 reanalysis data are available at https://apps.ecmwf.int/datasets/data/era40-daily/levtype=sfc/. Coupled Model Intercomparison Project Phase 6 model data are available at https://esgf-node.llnl.gov/search/cmip6/.

Acknowledgments
This study was supported by the National Key Research and Development Program of China (2017YFA0603804) and the National Natural Science Foundation of China (41991283). Laurent Li acknowledges French GENCI and IDRIS for the allocation of computing resources. Lianyi Guo is supported by the Special Research Assistant Funding Project of Chinese Academy of Sciences.

References

Bulla, J., Bulla, I., & Nenadić, O. (2010). Hsmm-An R package for analyzing hidden semi-Markov models. *Computational Statistics & Data Analysis*, *54*(3), 611–619. https://doi.org/10.1016/j.csda.2008.08.025

Cioffi, F., Conticello, F., & Lall, U. (2016). Projecting changes in Tanzania rainfall for the 21st century. *International Journal of Climatology*, *36*(13), 4297–4314. https://doi.org/10.1002/joc.4632

Cioffi, F., Conticello, F., & Lall, U. (2020). Stochastic scenarios for 21st century rainfall seasonality, daily frequency, and intensity in South Florida. *Journal of Water Resources Planning and Management*, *146*(8), 04020058. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001250

Cioffi, F., Conticello, F., Lall, U., Marotta, L., & Telesca, V. (2017). Large scale climate and rainfall seasonality in a Mediterranean Area: Insights from a non-homogeneous Markov model applied to the Agro-Pontino plain. *Hydrological Processes*, *31*(3), 668–686. https://doi.org/10.1002/hyp.11061

Ding, M., Jiang, Z. H., & Chen, W. L. (2016). Simulation and evaluation of summer daily precipitation based on nonhomogeneous hidden Markov model over the Yangtze-Huaihe River Basin. *Journal of Meteorological Research*, *74*(5), 757–771. https://doi.org/10.1002/joc.5882

Dong, T. Y., & Dong, W. J. (2021). Evaluation of extreme precipitation over Asia in CMIP6 models. *Climate Dynamics*, *57*, 1751–1769. https://doi.org/10.1007/s00382-021-05773-1

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., & Liebert, J. (2012). HESS Opinions "Should we apply bias correction to global and regional climate model data?" *Hydrology and Earth System Sciences*, *16*, 3391–3404. https://doi.org/10.5194/hess-16-3391-2012

Fan, X. W., Jiang, L., & Gou, J. J. (2021). Statistical downscaling and projection of future temperatures across the Loess Plateau, China. *Weather and Climate Extremes*, *32*, 100328. https://doi.org/10.1016/j.wace.2021.100328

Fu, G. B., Charles, S. P., & Kirshner, S. (2013). Daily rainfall projections from general circulation models with a downscaling nonhomogeneous hidden Markov model (NHMM) for south-eastern Australia. *Hydrological Processes*, *27*(25), 3663–3673. https://doi.org/10.1002/hyp.9483

Greene, A. M., Robertson, A. W., Smyth, P., & Triglia, S. (2011). Downscaling projections of Indian monsoon rainfall using a non-homogeneous hidden Markov model. *Quarterly Journal of the Royal Meteorological Society*, *137*, 347–359. https://doi.org/10.1002/qj.788

Gu, H. H., Yu, Z. B., Wang, G. L., Wang, J., Ju, Q., Yang, C., & Fan, C. (2014). Impact of climate change on hydrological extremes in the Yangtze River Basin, China. *Stochastic Environmental Research and Risk Assessment*, *29*, 693–707. https://doi.org/10.1007/s00477-014-0957-5

Guo, L. Y., Jiang, Z. H., Ding, M., Chen, W. L., & Li, L. (2019). Downscaling and projection of summer rainfall in Eastern China using a nonhomogeneous hidden Markov model. *International Journal of Climatology*, *39*(3), 1319–1330. https://doi.org/10.1007/s13351-019-8105-2

Guo, Y., Li, J. P., & Li, Y. (2011). Statistically downscaled summer rainfall over the middle-lower reaches of the Yangtze River. *Atmospheric and Oceanic Science Letters*, *4*(4), 191–198. https://doi.org/10.1080/16742834.2011.11446928

Han, L. Q., Han, Z., & Li, S. L. (2014). Projection of heavy rainfall events in the middle and lower reaches of the Yangtze River valley in the 21st century under different representative concentration pathways. *Transactions of Atmospheric Sciences*, *37*(5), 529–540. https://doi.org/10.1002/joc.4043

Holsclaw, T., Greene, A. M., Robertson, A. W., & Smyth, P. (2016). A Bayesian hidden Markov model of daily precipitation over South and East Asia. *Journal of Hydrometeorology*, *17*(1), 3–25. https://doi.org/10.1175/JHM-D-14-0142.1

Hughes, J. P., & Guttorp, P. (1994). A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water Resources Research*, *30*(5), 1535–1546. https://doi.org/10.1029/93WR02983

Hughes, J. P., Guttorp, P., & Charles, S. P. (1999). A non-homogeneous hidden Markov model for precipitation occurrence. *Journal of the Royal Statistical Society*, *48C*, 15–30. https://doi.org/10.1111/1467-9876.00136

Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., & Saynisch-Wagner, J. (2021). Towards neural Earth system modelling by integrating artificial intelligence in Earth system science. *Nature Machine Intelligence*, *3*, 667–674. https://doi.org/10.1038/s42256-021-00374-3

Khalil, A. F., Kwon, H. H., Lall, U., & Kaheil, Y. H. (2010). Predictive downscaling based on non-homogeneous hidden Markov models. *Hydrological Sciences Journal-Journal des Sciences Hydrologiques*, *55*(3), 333–350. https://doi.org/10.1080/02626661003780342

Li, M., Jiang, Z. H., Zhou, P., Le Treut, H., & Li, L. (2020). Projection and possible causes of summer precipitation in eastern China using self-organizing map. *Climate Dynamics*, *54*, 2815–2830. https://doi.org/10.1007/s00382-020-05150-4

Long, Y. N., Tang, R., Wang, H., & Jiang, C. B. (2019). Monthly precipitation modeling using Bayesian Non-homogeneous Hidden Markov Chain. *Hydrology Research*, *50*(2), 562–576. https://doi.org/10.2166/nh.2018.077

Pan, Z. T., Zhang, Y. J., Liu, X. D., & Gao, Z. (2016). Current and future precipitation extremes over Mississippi and Yangtze River basins as simulated in CMIP5 models. *Journal of Earth Science*, *27*(1), 22–36. https://doi.org/10.1007/s12583-016-0627-2

Pineda, L. E., & Willems, P. (2016). Multisite downscaling of seasonal predictions to daily rainfall characteristics over Pacific–Andean river basins in Ecuador and Peru Using a nonhomogeneous Hidden Markov model. *Journal of Hydrometeorology*, *17*, 481–498. https://doi.org/10.1175/JHM-D-15-0040.1

Robertson, A. W., Kirshner, S., & Smyth, P. (2004). Downscaling of daily rainfall occurrence over northeast Brazil using a hidden Markov model. *Journal of Climate*, *17*(22), 4407–4424. https://doi.org/10.1175/JCLI-3216.1

Tan, W. L., Yusof, F., & Yusop, Z. (2013). Non-homogeneous hidden Markov model for daily rainfall amount in peninsular Malaysia. *Jurnal Teknologi*, *63*(2), 75–80. https://doi.org/10.11113/jt.v63.1916

Wang, H., Asefa, T., & Sarkar, A. (2021). A novel non-homogeneous hidden Markov model for simulating and predicting monthly rainfall. *Theoretical and Applied Climatology*, *143*(1), 627–638. https://doi.org/10.1007/s00704-020-03447-2

Xin, X. G., Wu, T. W., Zhang, J., Yao, J. C., & Fang, Y. J. (2020). Comparison of CMIP6 and CMIP5 simulations of precipitation in China and the East Asian summer monsoon. *International Journal of Climatology*, *40*, 6423–6440. https://doi.org/10.1002/joc.6590

Yin, C., Li, Y., Ye, W., Bornman, J. F., & Yan, X. (2011). Statistical downscaling of regional daily precipitation over southeast Australia based on self-organizing maps. *Theoretical and Applied Climatology*, *105*, 11–26. https://doi.org/10.1007/s00704-010-0371-y

Zhu, H. H., Jiang, Z. H., Li, J., Li, W., Sun, C. X., & Li, L. (2020). Does CMIP6 inspire more confidence in simulating climate extremes over China? *Advances in Atmospheric Sciences*, *37*, 1119–1132. https://doi.org/10.1007/s00376-020-9289-1

Zhu, H. H., Jiang, Z. H., & Li, L. (2021). Projection of climate extremes in China, an incremental exercise from CMIP5 to CMIP6. *Science Bulletin*, *66*(24), 2528–2537. https://doi.org/10.1016/j.scib.2021.07.026

Zhu, X., Lee, S. Y., Wen, X. H., Ji, Z. M., Lin, L., Wei, Z. G., et al. (2021). Extreme climate changes over three major river basins in China as seen in CMIP5 and CMIP6. *Climate Dynamics*, *57*, 1187–1205. https://doi.org/10.1007/s00382-021-05767-z