Major Moisture Pathways and Their Importance to Rainy Season Precipitation over the Sanjiangyuan Region of the Tibetan Plateau

YU ZHANG

State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China

WENYU HUANG

Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China

DEYU ZHONG

Joint-Sponsored State Key Laboratory of Plateau Ecology and Agriculture, School of Water Resources and Electric Power, Qinghai University, Xining, and State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China

(Manuscript received 10 March 2019, in final form 11 June 2019)

ABSTRACT

Knowledge of the quantitative importance of moisture transport pathways of the Sanjiangyuan region (known as the “water tower” of China) can provide insights into the regional atmospheric branch of the hydrological cycle over the Sanjiangyuan region. A combined method with a clustering algorithm [Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN)] and a Lagrangian moisture source diagnostic is developed to identify the major moisture transport pathways and quantify their importance to three types of consecutive precipitation events—extreme precipitation (EP) events, moderate precipitation (MP) events, and extreme aridity (EA) events—for the Sanjiangyuan region during the rainy season (June–September 1960–2017). The results indicate that moisture paths from the northwest covering northwest China and central Asia (the N.W. pathway) and moisture paths from southern and southeastern China (the S.S. pathway) are stable moisture transport pathways during EP and MP events [importance (precipitation contribution in percentage): N.W. pathway, 18.4% (EP), 32.2% (MP); S.S. pathway, 25.9% (EP), 28.5% (MP)]. Affected by the western edge of a significant anticyclone anomaly centered around 35°N, 115°E, the moisture paths via the Bay of Bengal (the B.B. pathway) can reach the target region and become a supplementary moisture contributor (14.9%) to EP events. Moisture paths via the Arabian Sea and Indian peninsula (the A.I. pathway) are also active but the contributions are limited [4.9% (EP) and 5.6% (MP)]. For EA events, the fast-moving trajectories from farther western Asia (the F.W. pathway) play a dominant role and all major moisture pathways (F.W., N.W., and S.S. pathways) carry limited moisture to the target region.

1. Introduction

As early as the 1980s, the Tibetan Plateau (TP) was found to be a key region closely related to severe weather events over eastern, southern, and northern China through the low-level shear lines and vortices over and

around the TP (Xue 1980; Tao and Ding 1981). Located along the eastern edge of the TP shear lines [see Fig. 4 in Xu et al. (2014) and Fig. 3 in Zhang et al. (2016)], the Sanjiangyuan region (also known as the “Three-River Headwaters” region), which covers more than 350 000 km², is characterized by complex atmospheric circulation conditions due to interactions among circulation systems in the broader East Asian region, such as the South Asian monsoon, midlatitude westerlies, the East Asian monsoon, and high-latitude northerlies (see Fig. 2.1 in McGregor 2016). As the site of the world’s largest alpine wetland ecosystem (Tong et al. 2014), the

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-19-0196.s1>.

Corresponding author: Deyu Zhong, zhongdy@tsinghua.edu.cn

Sanjiangyuan region plays an irreplaceable role in water source conservation, run-off adjustment, and biodiversity maintenance and thus acts as an important ecological barrier of China. Moreover, the Sanjiangyuan region contains the headwaters of three major Asian rivers (Yangtze, Yellow, and Lancang), which provide approximately 49%, 20%, and 15% of the runoff of the Yellow, Yangtze, and Lancang Rivers, respectively (Cao and Pan 2014). Because of the large amount of water yield and the high altitude (with an average elevation of 4500 m), the Sanjiangyuan region is known as the “water tower” of China. The complex atmospheric circulation conditions and the roles as an ecological barrier and as the water tower of China make the Sanjiangyuan region a key region on the TP. A detailed investigation of the hydrological processes of the region contributes to the formulation of effective strategies for regional ecosystem and water resource conservation. As the atmospheric moisture transport process is an important part of the hydrological cycle, knowledge of where moisture comes from (moisture sources) and how it is transported (moisture transport pathways) is of great value to provide a full picture of the hydrological cycle in the Sanjiangyuan region, and this issue deserves a systematic and in-depth study. The primary task in this study is to identify the major moisture transport pathways and quantify their importance to the precipitation in the Sanjiangyuan region.

Enormous efforts have been made to identify the moisture transport conditions and moisture sources of the precipitation over the Sanjiangyuan region and the TP. Earlier studies based on the Eulerian-based moisture transport flux fields have found that the major moisture pathways for the Sanjiangyuan region and the TP are those from the Arabian Sea, the Bay of Bengal, the South China Sea, and the midlatitude westerlies (e.g., Simmonds et al. 1999; Li et al. 2009; Feng and Zhou 2012). The rapid development of atmospheric moisture tracking models, such as the Eulerian atmospheric water tracer (AWT; e.g., Sodemann et al. 2009; Sodemann and Stohl 2013; Dominguez et al. 2016), the dynamic recycling model (DRM; Dominguez et al. 2006; Martinez and Dominguez 2014), the water accounting model (WAM; van der Ent et al. 2010, 2013), and the Lagrangian-based surface freshwater flux method ($E - P$ flux method, where E is evaporation flux and P precipitation flux) (Stohl and James 2004, 2005) and its variant method, referred to as the S08 method hereafter from its explication in Sodemann et al. (2008), which can quantify the precipitation contributions of different moisture sources, has led to quantitative studies designed to reveal the importance of different moisture sources to precipitation over the TP. Chen et al. (2012)

identified the moisture sources of TP precipitation by developing a climatology of moisture sources during boreal summer via the $E - P$ flux method and found a dominant source band extending from the Arabian Sea to the Indian subcontinent and two additional sources extending from the Bay of Bengal to the northwestern TP. Sun and Wang (2014) proposed a modified version of the S08 method (referred to as the SW14 method hereafter) to quantify the total contribution of a predefined source region to the total precipitation in the target region and determine the moisture sources of semiarid grassland in the eastern TP (including the Sanjiangyuan region). The authors concluded that continental sources (e.g., moisture sources from the Eurasian continent) contribute more to the precipitation in the target region than oceanic sources (e.g., moisture sources located in the Arabian Sea and the Bay of Bengal). Quan et al. (2016) applied the $E - P$ flux method to diagnose moisture sources for the Sanjiangyuan region. The dominant moisture sources of the Sanjiangyuan region exhibit intraseasonal variations, changing from the western TP in June to the Arabian Sea in July, and finally to the Bay of Bengal in August. The northern TP is a stable moisture source during summer. Zhang et al. (2017) investigated the changes in moisture sources of the western-central TP precipitation during 1979–2013 using the WAM and suggested that the Indian summer monsoon, the East Asian summer monsoon, and the westerlies all contribute significantly to the summer precipitation over the targeted region. Hua et al. (2017) applied the DRM to provide a full picture of the precipitation recycling ratio over mainland China and indicated a recycling ratio of 0.23 in the eastern TP. Huang et al. (2018a) adopted the S08 method to help explore the formation mechanism for wintertime extreme precipitation events over the southeastern TP and revealed that the terrestrial sources south of the TP account for 78.7% of the total moisture for these extreme events. Pan et al. (2019) embedded the AWT in the Community Atmosphere Model version 5.1 (CAM5.1) and identified the moisture sources of the TP and found that the dominant moisture source region of the northern TP during summer is the TP itself (accounting for approximately 25% of the total moisture) and the tropical Indian Ocean is the most significant moisture supplier (around 28%) for the southern TP during summer.

The two Lagrangian methods, $E - P$ flux and the S08 method, are preferred in the study of moisture sources of the TP and the Sanjiangyuan region. Compared with the Eulerian-based moisture tracking models, such as AWT and WAM, the Lagrangian methods provide a straightforward approach to analyzing moisture

transport pathways in regions of interest by generating Lagrangian air parcel trajectories. Another important feature of these methods is the ability to provide information of the en route changes in moisture content, which enables us to quantify the en route loss of moisture originating from certain regions. Specifically, in terms of the $E - P$ flux method, a high negative $E - P$ flux in certain location indicates a great loss of moisture when air parcels pass over that location. However, a high positive $E - P$ flux in a certain location does not mean that the location is also a moisture source of the target region since moisture obtained at that location might be released before reaching the target region (Sodemann et al. 2008). The S08 method was developed by Sodemann et al. (2008) to address this issue and quantify the moisture loss amounts by taking into account the effects of subsequent evaporation and precipitation on previous evaporation sources. The performance of the S08 method has been demonstrated with the Eulerian AWT models (Winschall et al. 2014). In addition to the TP, the S08 method has also been widely applied to explore the moisture sources of different regions in China, such as Sichuan (Huang and Cui 2015), the Loess Plateau (Hu et al. 2018), East China (Sun and Wang 2015), and South China (Huang et al. 2018b). When diagnosing the moisture sources, some studies also investigated the moisture pathways of the region of interest by presenting a qualitative schematic diagram (e.g., Sun and Wang 2015; Wang et al. 2018) or conducting clustering algorithms, such as k -means clustering (e.g., Li and Dolman 2016), self-organizing maps (SOMs; e.g., Tan et al. 2018), total spatial variance (TSV)-based clustering (e.g., Tang et al. 2015), and expertise-based techniques (e.g., Li et al. 2016; Huang et al. 2018b). However, certain restrictions, such as predefined parameters (k -means and SOM), random initialization (k -means and SOM), computational efficiency (TSV-based algorithm), and lack of universality (expertise-based clustering), might limit the performances of these algorithms to identify moisture pathways. Furthermore, existing studies mainly focused on the contributions of predefined regions to precipitation in the region of interest (e.g., Sun and Wang 2014, 2015; Huang and Cui 2015; Wang et al. 2018; Hu et al. 2018; Huang et al. 2018b; Pan et al. 2019). The quantitative importance of moisture pathways was rarely addressed; thus, the moisture pathways of the Sanjiangyuan region require further quantitative understandings. Considering these two points, the present study introduces a stable and computationally efficient clustering algorithm that does not require a predetermined cluster number, the hierarchical density-based spatial clustering of applications with noise (HDBSCAN), to identify the

major moisture pathways of the rainy season precipitation over the Sanjiangyuan region. Then, the principles of the S08 and SW14 methods are adopted to quantify the importance of each major moisture pathway. As noted by Wang et al. (2006), atmospheric circulation conditions differ during different types of precipitation in the Sanjiangyuan region. Therefore, major moisture pathways are investigated separately for different precipitation types. With the aid of long-term precipitation data (1960–2017) from 14 meteorological stations in the target region, three precipitation types (extreme precipitation events, moderate precipitation events, and extremely arid events) are specified. The air parcel trajectories are generated using the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT) (Draxler and Hess 1997, 1998), which can directly use long-term NOAA–NCEP–NCAR reanalysis 1 (NNR1) data (from 1948 to the present; Kalnay et al. 1996) as meteorological field inputs. Details of the dataset and methodology are described in section 2. The results of major moisture pathway identification are shown in section 3. Analyses of the importance of each major moisture pathway are described in section 4. Discussions regarding the methodology and the findings are presented in section 5, and the major conclusions are provided in section 6.

2. Data and methodology

a. Definition of three types of precipitation events

Rainy season (June–September) daily precipitation data (1960–2017) from 14 meteorological stations located inside the Sanjiangyuan region (Fig. 1) are used to define three types of precipitation events. The daily precipitation data are provided by the China National Meteorological Information Center (Ren et al. 2012). Additional information about the availability of daily precipitation data of 14 meteorological stations can be found in Table S1 in the online supplemental material. Consecutive precipitation events are considered in the present study since consecutive precipitation events over the target region might have relatively persistent meteorological circulation conditions with similar patterns, compared with 1-day extreme precipitation events (Chen and Zhai 2013). The consecutive precipitation events are defined in the following steps. First, the region-averaged daily precipitation rate (mm day^{-1}) is obtained by averaging the valid precipitation rate of the 14 stations. Next, the 5th, 45th, 55th, and 95th percentiles of precipitation rates are calculated. Three types of consecutive precipitation events are classified as follows: 1) extreme precipitation (EP) events, which are

FIG. 1. Topography (colors; m) and meteorological stations in the Sanjiangyuan region. In the top panel, the thin black line indicates the boundary of the TP, the bold black line indicates the boundary of the Sanjiangyuan region, and the deep blue lines denote the Yangtze, Yellow, and Lancang Rivers. In the bottom panel, the yellow dots represent the locations of meteorological stations.

represented by consecutive days during which daily precipitation rates exceed the 95th percentile; 2) moderate precipitation (MP) events, which are represented by consecutive days during which daily precipitation rates fall between the 45th and 55th percentiles; and 3) extremely arid (EA) events, which are represented by consecutive days during which daily precipitation rates stay below the 5th percentile. All three types of consecutive precipitation events allow 1-day interruptions between any two eligible days during one precipitation event. The percentiles and corresponding event numbers are shown in Table 1. To verify this definition method, gridded daily precipitation data ($0.25^\circ \times 0.25^\circ$) from the Asian Precipitation—Highly Resolved Observational Data Integration Toward Evaluation of Water Resources [APHRODITE; Yatagai et al. 2012; versions APHRO_MA_V1101 (1960–2007) and APHRO_MA_V1101_EX_R1 (2008–15)] project are used to depict the precipitation conditions of the corresponding

precipitation events. The results (see Table S2 and Fig. S1) indicate satisfactory consistency between the station data and the APHRODITE data in both the spatial and temporal variability of the precipitation over the Sanjiangyuan region.

b. Trajectory generation description

Backward air parcel trajectories are generated by the HYSPLIT model (January 2017 release) developed by Draxler and Hess (1997, 1998). The reprocessed NNR1 data (HYSPLIT compatible format) are used to drive the model. A release grid ($1^\circ \times 1^\circ$) clipped by the regional boundary with 200 vertical levels is defined as the release positions of the trajectories [note that a broad vertical range, 0–10 000 m above ground level (AGL) with an interval of 50 m, is defined here in case of the existence of mid- to high-level moisture pathways in certain types of precipitation events, especially the EA events]. The trajectory simulation runs backward in time for 240 h, which is the average residence time of moisture in the atmosphere (Trenberth 1998, 1999; Numaguti 1999). To cover the period of each precipitation event, the simulation starting time of each trajectory is chosen to be 0000 UTC on the day after the last day of each precipitation event. The latitude, longitude, height, pressure, specific humidity, and mixed-layer depth are selected as the output parameters, and the output time interval is 1 h. The total numbers of trajectories generated for three precipitation types are shown in Table 1. The integration errors (Draxler et al. 2018) of the trajectories are calculated to exclude trajectories with significant calculation errors. The number of valid trajectories is shown in Table 1. In addition, not all air trajectories bring moisture to the target region, and only those that steadily release moisture in the target area should be treated as the stable moisture paths. Therefore, moisture paths in the present study are defined as those that have statistically significant ($p < 0.05$) negative trends of moisture content inside the target region. Table 1 shows the number of moisture paths of each precipitation type.

TABLE 1. Percentiles used to identify three types of precipitation events, number of events, and numbers of generated trajectories, valid trajectories, and moisture paths.

	Percentiles	Critical precipitation rate (mm day^{-1})	No. of events	No. of trajectories	No. of valid trajectories	No. of moisture paths
Extreme precipitation (EP) events	95th	8.08	16	115 200	61 761	11 435
Moderate precipitation (MP) events	45th, 55th	1.96, 2.72	22	158 400	92 233	9890
Extreme arid (EA) events	5th	0.01	35	252 000	167 393	15 666

c. Methodology

1) MOISTURE PATH CLUSTERING METHOD

There are always noise points in real-world data that should be discarded during the clustering process, which is also the case for the model-generated moisture path data. The density-based spatial clustering of applications with noise (DBSCAN) (Ester et al. 1996) is a density-based clustering algorithm that is widely used for discovering clusters in large spatial datasets with noise, and it has been used in the clustering of vehicle and aircraft trajectories (e.g., Gariel et al. 2011; Kim and Mahmassani 2015). However, the performance of DBSCAN in some situations is unsatisfactory because a single set of the key parameters (“minPts”, “Eps”), which represent fixed density, does not suit all clusters. The hierarchical DBSCAN (HDBSCAN) algorithm proposed by Campello et al. (2013) and Campello et al. (2015) is a hierarchical version of DBSCAN that can be regarded as an improvement over DBSCAN in the aspects of concept and algorithm to detect clusters with varying densities. A new distance metric called “mutual reachability distance” (MRD) is proposed to represent the relationship between a pair of objects. The MRD also makes HDBSCAN robust to noise data. Then, an extension of the minimum spanning tree (MST) of the MRD graph is built to construct the cluster hierarchy, and finally, the clustering patterns are extracted by cutting the tree at different heights to select varying density clusters based on cluster stability. HDBSCAN has also been used for aircraft trajectory clustering (e.g., Wilson et al. 2016).

The primary parameters of the HDBSCAN algorithm to be adjusted are the “minimum samples” and “minimum cluster size.” The minimum samples parameter determines the number of samples in a neighborhood for a point to calculate its “core distance” for MRD determination. A higher value of minimum samples results in a more conservative result since many data points may be pushed farther and treated as noise, whereas clusters are restricted to much denser areas. The parameter minimum cluster size determines the smallest size to be considered a cluster. A higher value of minimum cluster size results in fewer clusters. For more algorithm details, please refer to the original work of Campello et al. (2013) and Campello et al. (2015). The HDBSCAN Python module developed by McInnes et al. (2017) is used to implement the identification of major moisture path groups.

2) MOISTURE SOURCE ATTRIBUTION METHOD

Here, following the ideas of the S08 and SW14 methods, a synthetic moisture path is proposed to

illustrate the calculation steps of the moisture source attribution method used in the present study. For simplicity, it is assumed that the moisture path is below the boundary layer [only moisture uptake positions below the boundary layer are considered as the potential moisture sources of the target region as noted by Sodemann et al. (2008)]. The geographic locations and specific humidity changes along the path are given in Fig. 2. The method used to diagnose moisture sources of the moisture path consists of the following steps.

- 1) Identify the moisture uptake and release locations along the moisture path by calculating the moisture changes between neighboring time points. A threshold $e_c = 0.2 \text{ g kg}^{-1} (6 \text{ h})^{-1}$ is set to select the moisture uptake locations with moisture uptake amount Δq^0 larger than e_c , as done by Sodemann et al. (2008), and negative moisture changes represent moisture release. Furthermore, the moisture release outside the target region is treated as moisture loss Δl , while the moisture release inside the target region is treated as moisture release Δp based on the SW14 method. Hereafter, for simplicity, we use the subscript of the time point to denote the corresponding geographic location. In terms of the synthetic moisture path, three moisture uptake locations ($\Delta q_{-30}^0 = q_{-30} - q_{-36} = 7.9 \text{ g kg}^{-1}$, $\Delta q_{-24}^0 = q_{-24} - q_{-30} = 4.1 \text{ g kg}^{-1}$, and $\Delta q_{-6}^0 = q_{-6} - q_{-12} = 3.6 \text{ g kg}^{-1}$), one moisture loss location ($\Delta l_{-18} = 7 \text{ g kg}^{-1}$), and one moisture release location ($\Delta p_0 = 8.2 \text{ g kg}^{-1}$) are identified.
- 2) Traverse the locations of moisture uptake, loss, and release forward in time. The initial values of discounted moisture uptake amount Δq are set to Δq^0 . Assume we are at location n .

- (i) If location n is a moisture uptake location, then calculate the fractional contribution f of each previous uptake location m :

$$f_m = \frac{\Delta q_m}{q_n}, \quad \text{for all } m \leq n \quad \text{and} \\ m \in \{\text{uptake locations}\}. \quad (1)$$

- (ii) If location n is a moisture loss location, then discount all the previous moisture uptake Δq_m values as follows:

$$\Delta q_m = \Delta q_m + \Delta l_n \times f_m, \quad \text{for all } m < n \quad \text{and} \\ m \in \{\text{uptake locations}\}, \quad (2)$$

and record the portion of moisture loss $\Delta l_{m,n}$ that should be attributed to the moisture uptake location m :

FIG. 2. Synthetic example of a moisture path. Time points $t = -36, -30, -24$, and -18 are outside and $t = -12, -6$, and 0 are inside the target region. (top) The geographic location of the path. Orange segments indicate moisture uptake processes between the neighboring time points, blue segments indicate moisture release processes, and black segments indicate that neither uptake processes nor release processes happen. (bottom) The specific humidity changes along the path. Orange blocks indicate moisture uptake amounts, and blue blocks indicate moisture release amounts that can be attributed to certain moisture uptake locations.

$$\Delta l_{m,n} = \Delta l_n \times f_m, \quad \text{for all } m < n \quad \text{and} \\ m \in \{\text{uptake locations}\}. \quad (3)$$

- (iii) If location n is a moisture release location, then record the portion of the moisture release $\Delta p_{m,n}$ that should be attributed to the moisture uptake location m :

$$\Delta p_{m,n} = \Delta p_n \times f_m, \quad \text{for all } m < n \quad \text{and} \\ m \in \{\text{uptake locations}\}. \quad (4)$$

Then, discount Δq_m in the same way as in condition ii.

In terms of the synthetic moisture path, at $t = -30$, the moisture uptake and fractional contribution are calculated: $\Delta q_{-30} = \Delta q_{-30}^0 = 7.9 \text{ g kg}^{-1}$, $f_{-30} = \Delta q_{-30}/q_{-30} = 94.05\%$. At $t = -24$, $\Delta q_{-24} = \Delta q_{-24}^0 = 4.1 \text{ g kg}^{-1}$, $f_{-24} = \Delta q_{-24}/q_{-24} = 32.80\%$, and f_{-30} is updated by $f_{-30} = \Delta q_{-30}/q_{-24} = 63.20\%$. At $t = -18$, the moisture loss attributed to the uptake location $t = -30$ and $t = -24$ is $\Delta l_{-30,-18} = f_{-30} \times \Delta l_{-18} = 4.42 \text{ g kg}^{-1}$ and $\Delta l_{-24,-18} = f_{-24} \times \Delta l_{-18} = 2.30 \text{ g kg}^{-1}$, respectively; Δq_{-30} and Δq_{-24} are discounted by $\Delta q_{-30} = \Delta q_{-30} + \Delta l_{-18} \times f_{-30} = 3.48 \text{ g kg}^{-1}$ and $\Delta q_{-24} = \Delta q_{-24} + \Delta l_{-18} \times f_{-24} = 1.80 \text{ g kg}^{-1}$, respectively. And at $t = -6$ and $t = 0$, the update of f and discount of Δq are repeated.

- 3) Sum $\Delta l_{m,n}$ and $\Delta p_{m,n}$ over all moisture loss and release locations to obtain the total attributed moisture loss

ΔL_m and total attributed moisture release ΔP_m of each moisture uptake location m :

$$\Delta L_m = \sum_{n=1}^{N_1} \Delta l_{m,n}, \quad m \in \{\text{uptake locations}\}, \\ n \in \{\text{moisture loss locations}\}, \quad (5)$$

$$\Delta P_m = \sum_{n=1}^{N_2} \Delta p_{m,n}, \quad m \in \{\text{uptake locations}\}, \\ n \in \{\text{moisture release locations}\}, \quad (6)$$

where N_1 and N_2 are the total number of moisture loss and release locations, respectively. In terms of the synthetic moisture path, the amounts of moisture loss and release attributed to each uptake location are $\Delta L_{-30} = \Delta l_{-30,-18} = 4.42 \text{ g kg}^{-1}$, $\Delta L_{-24} = \Delta l_{-24,-18} = 2.30 \text{ g kg}^{-1}$, $\Delta P_{-30} = \Delta p_{-30,0} = 3.10 \text{ g kg}^{-1}$, $\Delta P_{-24} = \Delta p_{-24,0} = 1.60 \text{ g kg}^{-1}$, and $\Delta P_{-6} = \Delta p_{-6,0} = 3.21 \text{ g kg}^{-1}$.

- 4) Calculate the total moisture release amount inside the target region R_{total} :

$$R_{\text{total}} = \sum_{i=1}^N \Delta p_i, \quad (7)$$

where i represents the i th moisture release location and N is the total number of moisture release

FIG. 3. Workflow of the combination method applied to the EP events.

locations. The R_{total} of the synthetic moisture path is equal to p_0 (8.2 g kg^{-1}).

3) QUANTIFICATION OF IMPORTANCE OF MOISTURE PATHWAYS: COMBINATION OF MOISTURE PATH CLUSTERING AND MOISTURE SOURCE ATTRIBUTION

As noted in the introduction, the majority of previous Lagrangian model-based moisture source analyses subjectively predetermined potential moisture source regions using prior knowledge, manually drew “control boxes,” and then calculated their contributions. The control boxes can identify where moisture comes from (moisture sources) but cannot indicate its transport pathway (moisture pathways). In the present study, moisture path clustering and moisture source attribution are combined together to quantitatively identify major moisture pathways of the Sanjiangyuan region and quantify their importance. The workflow of the combination method applied to the EP events is shown in Fig. 3. The moisture paths are fed to the HDBSCAN algorithm, where some of them are selected to form several path groups/clusters. These groups are treated as the major path groups (i.e., the major moisture transport pathways) in the EP events. Moisture paths belonging to

the same major path group are fed to the moisture source attribution method. Suppose we are dealing with the moisture paths (N paths in total) from path group i (denoted as G_i). For each path, we obtain Δq_m^0 , ΔL_m , and ΔP_m for each identified moisture uptake location m and the total moisture release amount R_{total} . The total unweighted moisture uptake amount $\Delta Q_{G_i}^0$, total attributed moisture loss L_{G_i} , total attributed moisture release P_{G_i} , and total moisture release $R_{\text{total}}^{G_i}$ of path group G_i can be obtained through the following summation:

$$\begin{aligned}
 \Delta Q_{G_i}^0 &= \sum_{n=1}^N \sum_{m=1}^{M_n} (\Delta q_m^0)^{(n)}, \\
 L_{G_i} &= \sum_{n=1}^N \sum_{m=1}^{M_n} (\Delta L_m)^{(n)}, \\
 P_{G_i} &= \sum_{n=1}^N \sum_{m=1}^{M_n} (\Delta P_m)^{(n)}, \\
 R_{\text{total}}^{G_i} &= \sum_{n=1}^N (R_{\text{total}})^{(n)}, \quad (8)
 \end{aligned}$$

where M_n is the number of uptake locations of path n and N is the number of paths in path group G_i . By repeating the calculation process for all path groups, the

contribution of path group G_i to the total moisture release by all path groups is calculated by the following equation:

$$C_{G_i} = \frac{P_{G_i}}{\sum_{i=1}^N R_{G_i}^{total}}, \quad (9)$$

where N is the number of path group and C_{G_i} is the contribution of path group G_i . Moreover, to acquire geographic information of moisture sources, Δq_m^0 , ΔL_m , and ΔP_m of each moisture uptake location of each moisture path from each path group are gridded onto grid cells of a $1^\circ \times 1^\circ$ global grid by aggregating the Δq_m^0 , ΔL_m , and ΔP_m of moisture paths when they pass through a certain grid cell.

3. Major moisture transport pathways

In this section, we first conduct a frequency analysis on the moisture paths to obtain an overview of the high-occurrence areas for the three precipitation types (section 3a). The results of the major moisture path group identifications are shown in section 3b and compared with the results of the frequency analysis.

a. Overview of moisture paths

The frequency of moisture paths passing over a grid cell (i, j) F_{ij} ($1^\circ \times 1^\circ$ grid) is calculated by the following equation:

$$F_{ij} = \frac{K}{N}, \quad (10)$$

where N is the total number of moisture paths and K is the total number of moisture paths passing over the cell (i, j) . The moisture path frequencies of different precipitation types are shown in Figs. 4a–c. In general, moisture paths that enter the target region are from four regions: the large area of Eurasia northwest of the target region, the Arabian Sea/Indian peninsula region controlled by the cross-equatorial flows, the Bay of Bengal, and southern and southeastern China (including the South China Sea). In the EP events, the frequencies of moisture paths over the Arabian Sea/Indian peninsula and the Bay of Bengal are rather high; notably, a high-frequency band occurs from the high latitudes, and it passes over northwest China and finally reaches the target region. A thin band also occurs over southern China. In the MP events, several high-frequency bands occur north and northwest of the target region. A long and thin band occurs over the Arabian Sea/Indian peninsula, and a rather short but marked band exists over inland South China. In the EA

events, the moisture mainly comes from west and central Asia and is controlled by the westerlies. A small branch from southeast China also occurs. Figures 4d–f show the spatial distribution of the mean specific humidity of moisture paths gridded onto each grid cell and weighted by the frequency of each grid cell. In the EP and MP events, the Arabian Sea/Indian peninsula band always carries abundant moisture. In the EP events, the Bay of Bengal, northwestern area, and southern China bands also carry abundant moisture. In the MP events, the condition is similar to that of the EP events, but the Bay of Bengal band is absent. Unsurprisingly, all high-frequency bands carry limited moisture in the EA events.

b. Clustering analysis of moisture paths

When applying the HDBSCAN algorithm to identify major moisture path groups, the primary parameters to be adjusted are the “minimum samples” and “minimum cluster size.” The two parameters are determined by first maximizing the mean probability of the training samples, a measure of the strength with which each sample is a member of its assigned cluster [refer to the HDBSCAN Python module documentation developed by McInnes et al. (2017) for the details of this measure]. Then, adjustments are made to meet the requirements of cluster coherence. The final parameters used to run the clustering algorithm are shown in Table 2. The ordinary Euclidean distance is used as the clustering distance metric.

Figures 5–7 show the clustering results. The clusters grouped by the HDBSCAN method match well with the high-frequency bands in Figs. 4a–c. In the EP events, cluster EP1 corresponds with the Bay of Bengal band. Moisture paths of EP1 start from the Arabian Sea, pass over the eastern edge of India and the Bay of Bengal, rise over the eastern Himalayas, and reach the target region through the southern boundary. EP2 is the cross-equatorial band. Moisture paths of EP2 start from the southern Indian Ocean, cross the equator, pass over the Arabian Sea and mainland India, rise over the Himalayas, and reach the target region from the southern boundary. EP4 is a branch of the South China Sea and Indochina Peninsula and moisture paths of EP4 enter the target region through the eastern boundary. EP3 and EP5 match the band from central Asia (mainly from Kazakhstan or western Siberia). Moisture paths of EP3 and EP5 pass over northwest China and reach the target region through the northern boundary. In the MP events, MP1, MP2, and MP4 are the bands from central Asia and northwest China. MP3 and MP6 cover the short inland moisture paths from central and southeast China. MP5 is the cross-equatorial band. In the EA

FIG. 4. (left) Frequency of the moisture paths and (right) mean specific humidity (g kg^{-1}) weighted by frequency of moisture paths in the (a),(d) EP, (b),(e) MP, and (c),(f) EA events. White colors indicate areas through which no moisture paths pass.

events, EA1, EA3, EA6, and EA7 are the long branches of moisture paths controlled by the westerlies. Paths of these clusters mainly come from North Africa and West Asia and then move straight to the east and change direction near the northern boundary of the target region. EA5 and EA8 represent moisture paths from western Siberia that pass northwest China and finally enter the target region through the northern boundary. EA2 and EA4 represent short paths that come from inland south China or its offshore areas and enter the target region through the eastern or southern boundary.

The corresponding specific humidity and pressure changes along the paths of each cluster are shown in Figs. 8–10. In the EP events, patterns of specific humidity and pressure of EP1 and EP2 (two Indian monsoon branches) indicate that particles with high moisture content (over 15 g kg^{-1}) and low altitude (pressure

approximately 1000 hPa) endure a sudden moisture loss and pressure decrease after $t = -80$. The patterns of EP3 and EP5 (two northwest branches) show that particles with low moisture content lose moisture and ascend slowly along the paths to the target region. The EP4 pattern (South China Sea branch) indicates a sudden loss of moisture around $t = -80$. In the MP events, patterns of MP1, MP2, and MP4 represent moisture paths with a slowly decreasing moisture content trend and a slowly decreasing pressure trend. The pattern of

TABLE 2. Final parameters used in the HDBSCAN algorithm.

	Minimum samples	Minimum cluster size of clusters
EP events	10	600
MP events	15	500
EA events	40	400

FIG. 5. HDBSCAN algorithm results for the EP events. Clusters are labeled as EP1, EP2, . . . , EP5. Values in parentheses indicate the numbers of paths from the same cluster and their percentages in total moisture paths (the same holds in Figs. 6 and 7).

MP5 is similar to the pattern of EP2. Patterns of MP3 and MP6 represent another low-level moisture path group with relatively high moisture contents. In the EA events, the moisture content of all clusters is low, mostly below 5 g kg^{-1} , except for certain paths of EA1, EA2, and EA8. The pressure patterns indicate that most (except some of EA1, EA2, and EA8) of the moisture paths are located in the middle troposphere (approximately 500 hPa).

According to Figs. 4–10, moisture path clusters of different precipitation types identified by the HDBSCAN algorithm can be further grouped as the major moisture path groups. These groups are as follows: 1) the Arabian Sea/Indian peninsula cross-equatorial group (the A.I. group), 2) the Bay of Bengal group (the B.B. group), 3) the northwest group (mainly covering western Siberia, central Asia, and northwest China; the N.W. group), 4) the south and southeast China group (inland south and southeast China and adjacent offshore area; the S.S. China group), and 5) the farther western Asia group (the F.W. Asia group). Table 3 shows the results of distributing each cluster into one of the five path groups. The path group boundaries are shown in Fig. 11, which shows that there are some overlaps between different path groups

and that the boundaries of the same group are not fixed during different precipitation events.

Apparently, the A.I. and B.B. groups are controlled by the Indian monsoon. Both groups carry abundant moisture during the initial stage but undergo a great loss of moisture and ascend abruptly before entering the target region. The A.I. group occurs in both the EP and MP events, while the B.B. group becomes a major group only in the EP events and acts as a moisture supplement for extreme precipitation events. The N.W. group and S.S. China group are the two major path groups for all three precipitation types. In the EP events, the N.W. group is a concentrated band, while in the MP and EA events this group covers a broader range. The position of the S.S. China group in the EP events is farther south than the positions in the MP and EA events, covering the area of the South China Sea and Indochina Peninsula. In the MP events, air parcels in this group move slowly, and it takes 10 days for these parcels to travel from central China to the target region. In the EA events, this group covers the major terrestrial area of southern and southeastern China. The F.W. Asia group, within which air parcels move very quickly, becomes a major path group only in the EA events, and the groups's

TABLE 3. Results of assigning each cluster to a certain path group. Values in parentheses are the percentages of the major moisture path groups in the total moisture paths.

	Arabian Sea/ Indian peninsula cross-equatorial group	Bay of Bengal group	Northwest group	South and Southeast China group	Farther West Asia group
EP events	EP2 (7.35%)	EP1 (14.96%)	EP3, EP5 (12.75%)	EP4 (6.00%)	
MP events	MP5 (6.17%)		MP1, MP2, MP4 (23.44%)	MP3, MP6 (12.03%)	
EA events			EA5, EA8 (5.61%)	EA2, EA4 (9.79%)	EA1, EA3, EA6, EA7 (17.42%)

high percentage indicates that the westerlies are dominant in the EA events.

4. Quantitative importance of major moisture transport pathways

Figures 8–10 indicate that many moisture paths undergo a large moisture loss before reaching the target region; therefore, it cannot be concluded that moisture paths with higher moisture content contribute more to the precipitation in the target region. Because of the overlap of different path groups in the same precipitation type and the boundary variation of the same path group in the different precipitation types, we do not subjectively mark rectangular polygons on the map, treat them as potential moisture source regions,

and then perform the calculations. Instead, the relative importance of each major moisture path group to the target region is determined using the moisture source attribution method described in section 2c(3). The spatial distributions of unweighted moisture uptake, attributed moisture loss, and attributed moisture release of moisture uptake locations are also shown in this section to reveal the spatial distribution patterns of moisture source locations of major moisture path groups.

a. Contributions of different moisture path groups

The total attributed moisture release P_G of each moisture path group and total moisture release $\sum_i R_{\text{total}}^{G_i}$ by all major moisture path groups for each precipitation type are given in Fig. 12. Moisture uptake locations belonging to

FIG. 6. HDBSCAN algorithm results for the MP events.

FIG. 7. HDBSCAN algorithm results for the EA events.

each path group are divided into two categories: locations outside the target region (nonlocal moisture uptake locations) and locations inside the target region (local moisture uptake locations). The amount of total moisture release by all major moisture path groups for the three precipitation types is 680.9, 362.1, and 112.0 $\text{g kg}^{-1} (240 \text{ h})^{-1}$, matching well with the precipitation types. Generally, the nonlocal moisture uptake locations are dominant on rainy events (EP and MP events), and it is unnecessary to analyze the

moisture sources for EA events. Therefore, nonlocal moisture uptake locations for the EP and MP events are mainly emphasized in the following analysis.

For the EP events, the A.I. group is of limited importance (4.9%; all contributed by nonlocal moisture uptake locations) while the N.W. group (18.4%; 17.7% by nonlocal moisture uptake locations; 0.7% by local moisture uptake locations) and S.S. China group (25.9%; 22.9% by nonlocal moisture uptake locations;

FIG. 8. (a)–(e) Specific humidity q (g kg^{-1}) and (f)–(j) pressure p (hPa) time series of each cluster for the EP events. The abscissa is the time t (h) from the release point of the moisture path; $t = -240$ is the end point of moisture paths, and $t = 0$ is the release time point. The black solid line shows the 50th percentiles of p or q at the certain time point. The orange area indicates the data between the 25th and 75th percentiles. The cyan area indicates the data between the 5th and 25th percentiles or between the 75th and 95th percentiles. The vertical gray line represents the time point when air parcels enter the target region (vertical gray solid line: the 50th percentile; vertical gray dashed lines: 25th and 75th percentiles). The styles of the figures remain the same in Figs. 9 and 10.

FIG. 9. (a)–(f) Specific humidity (g kg^{-1}) and (g)–(l) pressure (hPa) time series of each cluster for the MP events.

3.0% by local moisture uptake locations) play leading roles. The role of the B.B. group (14.9%; 14.3% by nonlocal moisture uptake locations; 0.6% by local moisture uptake locations) is also nonnegligible. The situation remains the same for the MP events, where the N.W. group (32.2%; 28.1% by nonlocal moisture uptake locations; 4.1% by local moisture uptake locations) and

S.S. China group (28.5%; 26.7% by nonlocal moisture uptake locations; 1.8% by local moisture uptake locations) are dominant. In addition, oceanic moisture sources make limited contributions to both the EP and MP events. Surprisingly, the two Indian monsoon groups are not the dominant moisture path groups for the precipitation in the target region. However, this

FIG. 10. (a)–(h) Specific humidity (g kg^{-1}) and (i)–(p) pressure (hPa) time series of each cluster for the EA events.

FIG. 11. Boundary of each moisture path group for (a) EP, (b) MP, and (c) EA events. Label 1: S.S. China group; label 2: N.W. group; label 3: A.I. group; label 4: B.B. group; label 5: F.W. Asia group.

result matches the findings of Sun and Wang (2014) that the contribution of moisture from the Arabian Sea and the Bay of Bengal can almost be ignored for the eastern TP during the warm season. The total moisture uptake and loss amount of each group (Fig. 13) indicate that for the EP events, the B.B. group obtains abundant moisture en route, but the moisture loss amount before reaching the target region is considerable. The A.I. group gains a similar amount of moisture as the N.W. and the S.S. China groups in the EP events but also undergoes great loss. For the MP events, the moisture uptake amount of the A.I. group is far smaller than that of the N.W. and S.S. China groups, and nearly all the moisture is released outside the target region.

b. Spatial distribution of moisture sources

Figure 14 shows the spatial distribution of unweighted moisture uptake, attributed moisture loss, and attributed moisture release of the nonlocal moisture uptake locations belonging to the major moisture path groups in the EP and MP events.

For the EP events, the majority of moisture sources of each path group can obtain abundant moisture (Fig. 14a). However, the amount of attributed moisture loss of the two Indian monsoon path groups is considerable (Fig. 14c), which indicates that a large part of the moisture acquired from the Arabian Sea, mainland India, and the Bay of Bengal is lost before the air parcels enter the target region. Nevertheless, moisture sources located in northeast India and the southern TP contribute to precipitation in the target region (Fig. 14e). Although moisture sources of the S.S. China group also undergo a great loss en route (Fig. 14c), moisture sources near the eastern boundary of the target region still have residual moisture for the target region (Fig. 14e). For the N.W. group, the moisture sources located at the border between China and Kazakhstan are mainly responsible for

the moisture loss en route (Fig. 14c), and the sources located close to the northern boundary of the target region provide moisture for the target region (Fig. 14e).

As for the MP events, the moisture paths of the A.I. group acquire little moisture from the ocean and only gain some moisture in front of the southern foot of the Himalayas (Fig. 14b), which is released en route (Fig. 14d). As is the case for the EP events, the moisture sources located close to the eastern boundary of the target region lose considerable moisture en route (Fig. 14d) but also meet the moisture demand of the target region. The moisture sources close to the northern region of the target region also make contributions to the target region, although these contributions are smaller than those near the eastern boundary (Fig. 14f).

Thus, it can be concluded that for rainy events in the Sanjiangyuan region, the moisture sources located close to the eastern and northern boundary of the region serve as the major precipitation contributors of the Sanjiangyuan region, and for EP events, the moisture sources located south of the Sanjiangyuan region are also active in providing additional moisture. Remote moisture sources, especially oceanic sources, contribute a limited amount of moisture.

5. Discussion

a. Sensitivity tests, and issues involved in clustering

The final parameter settings used in the present study [10-day trajectories released from $1^\circ \times 1^\circ$ gridded release positions at 0000 UTC on the day after the last day of the precipitation event, $e_c = 0.2 \text{ g kg}^{-1} (6 \text{ h})^{-1}$] are based on the results of sensitivity tests, where different trajectory release positions [the center of the target region ($96^\circ\text{N}, 34^\circ\text{E}$); a $1^\circ \times 1^\circ$ grid clipped by the regional boundary], trajectory release time points (0000 UTC on

FIG. 12. Total amount of attributed moisture release [$\text{g kg}^{-1} (240 \text{ h})^{-1}$] of moisture uptake locations of each path group responsible for the precipitation in the target region for the (a) EP, (b) MP, and (c) EA events. Values in the box are the total moisture release amount of all major path groups. Values at the top of the bars are the contributions to the total moisture release. Bars in blue represent the amount of attributed moisture release of terrestrial moisture uptake locations, and bars in orange represent the amount of attributed moisture release of oceanic moisture uptake locations. The label “Non-local” indicates the total attributed moisture release of the nonlocal moisture uptake locations, and the label “Local” is for local moisture uptake locations.

the first day of the precipitation event; 0000 UTC on the day after the last day of the precipitation event), and e_c [$0.2 \text{ g kg}^{-1} (2 \text{ h})^{-1}$; $0.2 \text{ g kg}^{-1} (6 \text{ h})^{-1}$; $0.8 \text{ g kg}^{-1} (6 \text{ h})^{-1}$; $0.8 \text{ g kg}^{-1} (12 \text{ h})^{-1}$] are tested to check the robustness of the results. The robustness of the HDBSCAN to “minimum cluster size” and “minimum samples” is also tested by slightly changing the determined parameters in Table 2. Moreover, the pressures of the moisture paths from the Indian summer monsoon region (clusters EP1, EP2, and MP5) decline rapidly from approximately 1000 hPa to approximately 400 hPa in 80 h ($t = -80$ to $t = 0$), indicating seemingly unrealistic vertical velocities of air parcels. To ensure that the suspicious vertical velocities are not caused by the usage of NNR1 data, the

moisture paths in the EP events after 2005 are generated using $1^\circ \times 1^\circ$ meteorological data output from the Global Data Assimilation System (GDAS). Detailed information and the results of the parameter and data sensitivity tests can be found in the supplemental material.

The HDBSCAN clustering results in the present study are relatively conservative. Many of the moisture paths are not included in the major moisture path groups, which is mainly due to two reasons: 1) infrequency of certain moisture paths, which might be caused by the occurrence of real noise paths (paths with calculation errors even after the exclusion of paths with large integration errors) or by the natural characteristics of some paths, and 2) the HDBSCAN parameter selection strategy. To further illustrate the two reasons, we apply the k -means algorithm to the moisture paths in the EP events. Through the k -means clustering results, we can verify the occurrence of some infrequent moisture path clusters and the conservation of the HDBSCAN parameter selection strategy. Detailed descriptions of the k -means clustering results can be found in the supplemental material. Moreover, although the clustering results of the HDBSCAN algorithm are consistent with the frequency analysis results and the majority of clusters show good coherence, a few southerly paths occur in the cluster MP2 (one cluster of the N.W. group), and they lead to an approximate symmetrical structure of MP2 that is very stable during parameter tuning. We speculate that the performance of the HDBSCAN algorithm on separating this symmetrical structure might be limited in such a high-dimensional data space (a dimension of 482 in the present study). We think that this problem can be alleviated by introducing dimensionality reduction techniques (e.g., the principal component analysis) or adding additional information (e.g., the moisture content along the moisture paths) and consider it worthy of further efforts.

b. Circulation differences between the EP and MP events, and the importance of the southwesterly moisture pathway

Figures 4a and 4b and Fig. 11 show that the EP and MP events share three major moisture pathways: the N.W. path group, the S.S. path group, and the A.I. path group. However, two distinct differences occur in the moisture pathways and moisture sources for the EP and MP events. One is the significant increase in the attributed moisture release amount and the spatial extension of moisture sources of the S.S. China path group in the EP events. The other is the occurrence of the B.B. path group and moisture sources located in northeast India and the southern TP in the EP events. To identify the reasons for these differences, composite anomalies of geopotential height and horizontal wind

FIG. 13. (top) Total unweighted moisture uptake amount, (middle) total attributed moisture loss, and (bottom) total attributed moisture release of the nonlocal moisture uptake locations of each moisture path group for the (a),(d),(g) EP, (b),(e),(h) MP, and (c),(f),(i) EA events [$\text{g kg}^{-1} (240 \text{ h})^{-1}$].

field relative to the climatology (June to September) at 850, 500, and 200 hPa during the EP and MP events and the corresponding differences between the EP and MP events are shown in Fig. 15. At 850 hPa, southerly wind anomalies are found in southeastern China (around 30°N , 110°E) during EP events (Fig. 15a), while slight northerly anomalies occur in similar regions during MP events (Fig. 15d). The difference map of geopotential height and wind field (Fig. 15g) presents a weak anticyclone anomaly (nonsignificant geopotential height difference) centered near Taiwan. At 500 hPa, in terms of the EP events, two obvious anticyclone anomalies are located at the central Asia (centered around 50°N , 65°E) and East China (centered around 35°N , 115°E), with one weak cyclone anomaly centered west of the Sanjiangyuan region. The “cyclone–anticyclone” anomaly structure around the Sanjiangyuan region induces strong southerly anomalies over the whole region. For the MP events, a weak cyclone anomaly is located over East China (centered around 40°N , 120°E) and the wind field anomalies near the Sanjiangyuan region are rather weak (Fig. 15e). The difference map at 500 hPa (Fig. 15h) confirms the significance of the two anticyclone

anomalies and the southerly wind anomalies over the Sanjiangyuan region. The patterns of the anticyclone anomalies at 200 hPa are similar to those at 500 hPa but with stronger amplitudes. Therefore, it can be confirmed that during EP events, a significant and deep anticyclone anomaly occurs over East China (centered around 35°N , 115°E). Anomalous anticyclones in similar locations have been found in previous works focusing on the dynamics and circulation patterns of the East Asian summer monsoon (EASM). Lau et al. (2000) revealed a strong positive correlation among the EASM index defined by them, the strong southerly winds over East and North China at 850 hPa, and an intensive anticyclone centered at 40°N , 120°E at 200 hPa. Wu (2002) identified a midlatitude Asian summer (MAS) pattern featuring two anomalous anticyclones centered at 37.5°N , 65°E and 42.5°N , 130°E . Cheng et al. (2008) developed a midlatitude land–sea thermal contrast index (LSI; 500-hPa geopotential height difference between 40° – 55°N , 75° – 90°E and 35° – 42.5°N , 140° – 150°E) to investigate the atmospheric thermal contrast between the midlatitude Asian continent and adjacent ocean. In the low LSI year, a broad anomalous anticyclone

FIG. 14. Spatial distribution of (top) unweighted moisture uptake, (middle) attributed moisture loss, and (bottom) attributed moisture release of nonlocal moisture uptake locations of major moisture path groups for the (left) EP and (right) MP events [$\text{g kg}^{-1} (240 \text{ h})^{-1}$ integrated over 10^4 km^2].

(centered near 40°N , 145°E) prevails over $35^\circ\text{--}50^\circ\text{N}$, $120^\circ\text{--}170^\circ\text{E}$. All these works revealed a positive correlation between anomalous anticyclones and precipitation over North China. Located at the western edges of the anomalous anticyclones, North China is controlled by southerly wind anomalies, which are

expected to bring warm and rich moisture. The anomalous anticyclone found in the present study is centered slightly southwest of the centers of those found by [Lau et al. \(2000\)](#), [Wu \(2002\)](#), and [Cheng et al. \(2008\)](#), and this finding indicates that the Sanjiangyuan region is under the control of the western edge of the anomalous

FIG. 15. Composite anomalies of the geopotential height (shading; gpm) and horizontal wind fields (vectors; m s^{-1}) at (left) 850, (center) 500, and (right) 200 hPa in the (a)–(c) EP and (d)–(f) MP events, and (g)–(i) corresponding differences between the EP and MP events. In the difference maps, shading and vectors are shown where the differences of anomalies in the EP and MP events meet the 95% confidence level (two-tailed Student's *t* test). Vectors are plotted if differences in the wind anomalies are significant in at least one direction (zonal or meridional).

anticyclone. This circulation pattern enables the southerly moisture pathways originating from the Asian summer monsoon region to reach into the target region. Consequently, in the EP events, the S.S. China path group is enhanced and extends to the South China Sea and the B.B. path group is activated, bringing abundant moisture into the target region. To further show the relationship between the precipitation over China and the anomalous anticyclone centered around 35°N, 115°E, correlation maps of daily precipitation with region-averaged (30°–40°N, 110°–120°E) geopotential height at 500 and 200 hPa during June to September (1960–2015) are shown in Fig. 16. A “positive/negative” correlation structure occurs between North China and South China, which is consistent with the findings of Lau et al. (2000), Wu (2002) and Cheng et al. (2008).

Notably, the positive pattern over North China is broad and reaches into the TP, covering the Sanjiangyuan region.

Another point we should note is that although the contribution of the southwesterly moisture pathway (i.e., the A.I. path group) to summer precipitation over the target region is limited, both in terms of the relative percentage and the absolute attributed moisture release amount, previous studies have revealed the important role of this moisture pathway in modulating the precipitation over the Sanjiangyuan region and the TP. The enhanced water vapor transport from the Indian Ocean during July and September is one of the primary reasons underlying the recent precipitation increase over the TP (Zhang et al. 2017). Sun and Wang (2018, 2019) performed a regression analysis between vertically

FIG. 16. Correlation of precipitation with region-averaged (30° – 40° N, 110° – 120° E) geopotential height at (a) 500 and (b) 200 hPa from June to September (1960–2015). Shadings are shown where the correlation is significant at the 95% confidence level.

integrated water vapor transport fluxes and the detrended and standardized precipitation time series of the Sanjiangyuan region and revealed close connections between the interannual variability of summer precipitation in the Sanjiangyuan region and the southwesterly moisture transport from the Indian peninsula. According to the findings of Zhang et al. (2017) and Sun and Wang (2018, 2019), the interannual variability of the contribution of the A.I. path group to precipitation over the Sanjiangyuan region deserves a more in-depth study.

6. Summary and conclusions

In this study, major moisture transport pathways for three types of consecutive precipitation events (EP, MP, and EA events) during the rainy season in the Sanjiangyuan region are identified using a clustering algorithm known as the HDBSCAN, and then the importance of the major moisture transport pathways and the spatial distributions of moisture sources belonging to these major moisture transport pathways are quantified and presented using a Lagrangian moisture source diagnostic method, which was originally developed by Sodemann et al. (2008) and modified by Sun and Wang (2014).

The results of the moisture transport pathway identification and moisture source diagnostic suggest that the N.W. path group and the S.S. China path group are the main and stable moisture transport pathways during the EP and MP events in the Sanjiangyuan region [precipitation contribution: N.W. path group, 18.4% (EP), 32.2% (MP); S.S. path group, 25.9% (EP), 28.5% (MP)]. Consequently, moisture sources located near the eastern and northern boundary of the region serve as major contributors to the regional precipitation in the EP and MP events. During the EP events, the Sanjiangyuan region is located at the western edge of a significant and

deep anticyclone anomaly located over East China (centered around 35° N, 115° E), which enables the B.B. path group to reach into the target region and become an important supplementary contributor (14.9%) to the precipitation of the EP events. Consequently, moisture sources located near the southern boundary of the target region are active in providing additional moisture. The A.I. group is also active in both the EP and MP events, but the precipitation contribution is limited [4.9% (EP); 5.6% (MP)]. In the EA events, the F.W. path group controlled by the westerlies plays a dominant role and all major moisture path groups (F.W., N.W., and S.S. path groups) carry limited moisture to the target region.

Acknowledgments. This research was supported by the National Natural Science Foundation of China (Grant 91547204) and the National Key Research and Development Program of China (Grant 2017YFC0404303).

REFERENCES

- Campello, R. J. G. B., D. Moulavi, and J. Sander, 2013: Density-based clustering based on hierarchical density estimates. *Advances in Knowledge Discovery and Data Mining*, J. Pei et al., Eds., Lecture Notes in Computer Science, Vol. 7819, Springer, 160–172, https://doi.org/10.1007/978-3-642-37456-2_14.
- , —, A. Zimek, and J. Sander, 2015: Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Trans. Knowl. Discov. Data*, **10**, 1–51, <https://doi.org/10.1145/2733381>.
- Cao, L., and S. Pan, 2014: Changes in precipitation extremes over the “Three-River Headwaters” region, hinterland of the Tibetan Plateau, during 1960–2012. *Quat. Int.*, **321**, 105–115, <https://doi.org/10.1016/j.quaint.2013.12.041>.
- Chen, B., X. D. Xu, S. Yang, and W. Zhang, 2012: On the origin and destination of atmospheric moisture and air mass over the Tibetan Plateau. *Theor. Appl. Climatol.*, **110**, 423–435, <https://doi.org/10.1007/s00704-012-0641-y>.

- Chen, Y., and P. M. Zhai, 2013: Persistent extreme precipitation events in China during 1951–2010. *Climate Res.*, **57**, 143–155, <https://doi.org/10.3354/cr01171>.
- Cheng, H., T. Wu, and W. Dong, 2008: Thermal contrast between the middle-latitude Asian continent and adjacent ocean and its connection to the East Asian summer precipitation. *J. Climate*, **21**, 4992–5007, <https://doi.org/10.1175/2008JCLI2047.1>.
- Dominguez, F., P. Kumar, X. Liang, and M. Ting, 2006: Impact of atmospheric moisture storage on precipitation recycling. *J. Climate*, **19**, 1513–1530, <https://doi.org/10.1175/JCLI3691.1>.
- , G. Miguez-Macho, and H. Hu, 2016: WRF with water vapor tracers: A study of moisture sources for the North American monsoon. *J. Hydrometeor.*, **17**, 1915–1927, <https://doi.org/10.1175/JHM-D-15-0221.1>.
- Draxler, R., and G. Hess, 1997: Description of the HYSPLIT_4 modeling system. NOAA Tech. Memo. ERL ARL-224, 24 pp.
- , and —, 1998: An overview of the HYSPLIT_4 modelling system for trajectories, dispersion, and deposition. *Aust. Meteor. Mag.*, **47**, 295–308.
- , B. Stunder, G. Rolph, A. Stein, and A. Taylor, 2018: HYSPLIT4 user's guide. NOAA Tech. Memo. ERL ARL-230, 254 pp.
- Ester, M., H.-P. Kriegel, J. Sander, and X. Xu, 1996: A density-based algorithm for discovering clusters in large spatial databases with noise. *Proc. Second Int. Conf. on Knowledge Discovery and Data Mining*, Portland, OR, AAAI, 226–231.
- Feng, L., and T. J. Zhou, 2012: Water vapor transport for summer precipitation over the Tibetan Plateau: Multidata set analysis. *J. Geophys. Res.*, **117**, D20114, <https://doi.org/10.1029/2011JD017012>.
- Gariel, M., A. N. Srivastava, and E. Feron, 2011: Trajectory clustering and an application to airspace monitoring. *IEEE Trans. Intell. Transp. Syst.*, **12**, 1511–1524, <https://doi.org/10.1109/TITS.2011.2160628>.
- Hu, Q., D. Jiang, and X. Lang, 2018: Sources of moisture for different intensities of summer rainfall over the Chinese Loess Plateau during 1979–2009. *Int. J. Climatol.*, **38**, e1280–e1287, <https://doi.org/10.1002/joc.5416>.
- Hua, L., L. Zhong, and Z. Ke, 2017: Characteristics of the precipitation recycling ratio and its relationship with regional precipitation in China. *Theor. Appl. Climatol.*, **127**, 513–531, <https://doi.org/10.1007/s00704-015-1645-1>.
- Huang, W., T. Qiu, Z. Yang, D. Lin, J. S. Wright, B. Wang, and X. He, 2018a: On the formation mechanism for wintertime extreme precipitation events over the southeastern Tibetan Plateau. *J. Geophys. Res. Atmos.*, **123**, 12 692–12 714, <https://doi.org/10.1029/2018JD028921>.
- , X. S. He, Z. F. Yang, T. P. Qiu, J. S. Wright, B. Wang, and D. Y. Lin, 2018b: Moisture sources for wintertime extreme precipitation events over South China during 1979–2013. *J. Geophys. Res. Atmos.*, **123**, 6690–6712, <https://doi.org/10.1029/2018JD028485>.
- Huang, Y., and X. Cui, 2015: Moisture sources of an extreme precipitation event in Sichuan, China, based on the Lagrangian method. *Atmos. Sci. Lett.*, **16**, 177–183, <https://doi.org/10.1002/asl2.562>.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).
- Kim, J., and H. S. Mahmassani, 2015: Spatial and temporal characterization of travel patterns in a traffic network using vehicle trajectories. *Transp. Res. Procedia*, **9**, 164–184, <https://doi.org/10.1016/j.trpro.2015.07.010>.
- Lau, K.-M., K.-M. Kim, and S. Yang, 2000: Dynamical and boundary forcing characteristics of regional components of the Asian summer monsoon. *J. Climate*, **13**, 2461–2482, [https://doi.org/10.1175/1520-0442\(2000\)013<2461:DABFCO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2461:DABFCO>2.0.CO;2).
- Li, L., and A. J. Dolman, 2016: A synoptic overview and moisture trajectory analysis of the “7.21” heavy rainfall event in Beijing. *J. Meteor. Res.*, **30**, 103–116, <https://doi.org/10.1007/s13351-016-5052-z>.
- , —, and Z. Xu, 2016: Atmospheric moisture sources, paths, and the quantitative importance to the Eastern Asian monsoon region. *J. Hydrometeor.*, **17**, 637–649, <https://doi.org/10.1175/JHM-D-15-0082.1>.
- Li, S. C., D. L. Li, P. Zhao, and G. Q. Zhang, 2009: The climatic characteristics of vapor transportation in rainy season of the origin area of three rivers in Qinhai-Xizang Plateau (in Chinese). *Acta Meteor. Sin.*, **67**, 591–598.
- Martinez, J. A., and F. Dominguez, 2014: Sources of atmospheric moisture for the La Plata River basin. *J. Climate*, **27**, 6737–6753, <https://doi.org/10.1175/JCLI-D-14-00022.1>.
- McGregor, G. R., 2016: Climate variability and change in the Sanjiangyuan region. *Landscape and Ecosystem Diversity, Dynamics and Management in the Yellow River Source Zone*. G. J. Brierley et al., Eds., Springer, 35–57.
- McInnes, L., J. Healy, and S. Astels, 2017: hdbscan: Hierarchical density based clustering. *J. Open Source Softw.*, **2**, 205, <https://doi.org/10.21105/joss.00205>.
- Numaguti, A., 1999: Origin and recycling processes of precipitating water over the Eurasian continent: Experiments using an atmospheric general circulation model. *J. Geophys. Res.*, **104**, 1957–1972, <https://doi.org/10.1029/1998JD200026>.
- Pan, C., B. Zhu, J. Gao, H. Kang, and T. Zhu, 2019: Quantitative identification of moisture sources over the Tibetan Plateau and the relationship between thermal forcing and moisture transport. *Climate Dyn.*, **52**, 181–196, <https://doi.org/10.1007/s00382-018-4130-6>.
- Quan, C., B. Chen, T. L. Zhao, B. R. Zhou, and Y. X. Han, 2016: Application of Lagrange water vapor source diagnosis method to the three river source area (in Chinese). *J. Appl. Meteor. Sci.*, **27**, 688–697.
- Ren, Z. H., F. L. Zou, Y. Yu, G. A. Wang, Z. F. Zhang, S. H. Fan, Z. Q. Zhang, and C. Sun, 2012: Dataset of daily climate data from Chinese surface stations, version 3.0. National Meteorological Information Center, accessed 5 January 2018, https://data.cma.cn/data/cdcdetail/dataCode/SURF_CLI_CHN_MUL_DAY_V3.0.html.
- Simmonds, I., D. Bi, and P. Hope, 1999: Atmospheric water vapor flux and its association with rainfall over China in summer. *J. Climate*, **12**, 1353–1367, [https://doi.org/10.1175/1520-0442\(1999\)012<1353:AWVFAI>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1353:AWVFAI>2.0.CO;2).
- Sodemann, H., and A. Stohl, 2013: Moisture origin and meridional transport in atmospheric rivers and their association with multiple cyclones. *Mon. Wea. Rev.*, **141**, 2850–2868, <https://doi.org/10.1175/MWR-D-12-00256.1>.
- , C. Schwiertz, and H. Wernli, 2008: Interannual variability of Greenland winter precipitation sources: Lagrangian moisture diagnostic and North Atlantic Oscillation. *J. Geophys. Res.*, **113**, D03107, <https://doi.org/10.1029/2007JD008503>.
- , H. Wernli, and C. Schwiertz, 2009: Sources of water vapour contributing to the Elbe flood in August 2002—A tagging study in a mesoscale model. *Quart. J. Roy. Meteor. Soc.*, **135**, 205–223, <https://doi.org/10.1002/qj.374>.
- Stohl, A., and P. James, 2004: A Lagrangian analysis of the atmospheric branch of the global water cycle. Part I: Method

- description, validation, and demonstration for the August 2002 flooding in central Europe. *J. Hydrometeorol.*, **5**, 656–678, [https://doi.org/10.1175/1525-7541\(2004\)005<0656:ALAOA>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0656:ALAOA>2.0.CO;2).
- , and —, 2005: A Lagrangian analysis of the atmospheric branch of the global water cycle. Part II: Moisture transports between Earth's ocean basins and river catchments. *J. Hydrometeorol.*, **6**, 961–984, <https://doi.org/10.1175/JHM470.1>.
- Sun, B., and H. Wang, 2014: Moisture sources of semiarid grassland in China using the Lagrangian particle model FLEXPART. *J. Climate*, **27**, 2457–2474, <https://doi.org/10.1175/JCLI-D-13-00517.1>.
- , and —, 2015: Analysis of the major atmospheric moisture sources affecting three sub-regions of east China. *Int. J. Climatol.*, **35**, 2243–2257, <https://doi.org/10.1002/joc.4145>.
- , and —, 2018: Interannual variation of the spring and summer precipitation over the Three River Source Region in China and the associated regimes. *J. Climate*, **31**, 7441–7457, <https://doi.org/10.1175/JCLI-D-17-0680.1>.
- , and —, 2019: Enhanced connections between summer precipitation over the Three-River-Source region of China and the global climate system. *Climate Dyn.*, **52**, 3471–3488, <https://doi.org/10.1007/s00382-018-4326-9>.
- Tan, X., T. Gan, and Y. D. Chen, 2018: Moisture sources and pathways associated with the spatial variability of seasonal extreme precipitation over Canada. *Climate Dyn.*, **50**, 629–640, <https://doi.org/10.1007/s00382-017-3630-0>.
- Tang, Y., H. Pang, W. Zhang, Y. Li, S. Wu, and S. Hou, 2015: Effects of changes in moisture source and the upstream rainout on stable isotopes in precipitation—A case study in Nanjing, eastern China. *Hydrol. Earth Syst. Sci.*, **19**, 4293–4306, <https://doi.org/10.5194/hess-19-4293-2015>.
- Tao, S., and Y. Ding, 1981: Observational evidence of the influence of the Qinghai-Xizang (Tibet) Plateau on the occurrence of heavy rain and severe convective storms in China. *Bull. Amer. Meteor. Soc.*, **62**, 23–30, [https://doi.org/10.1175/1520-0477\(1981\)062<0023:OEOTIO>2.0.CO;2](https://doi.org/10.1175/1520-0477(1981)062<0023:OEOTIO>2.0.CO;2).
- Tong, L., X. Xu, Y. Fu, and S. Li, 2014: Wetland changes and their responses to climate change in the “three-river headwaters” region of China since the 1990s. *Energies*, **7**, 2515–2534, <https://doi.org/10.3390/en7042515>.
- Trenberth, K. E., 1998: Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change. *Climatic Change*, **39**, 667–694, <https://doi.org/10.1023/A:1005319109110>.
- , 1999: Atmospheric moisture recycling: Role of advection and local evaporation. *J. Climate*, **12**, 1368–1381, [https://doi.org/10.1175/1520-0442\(1999\)012<1368:AMRROA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1368:AMRROA>2.0.CO;2).
- van der Ent, R. J., H. H. Savenije, B. Schaefli, and S. C. Steele-Dunne, 2010: Origin and fate of atmospheric moisture over continents. *Water Resour. Res.*, **46**, W09525, <https://doi.org/10.1029/2010WR009127>.
- , O. A. Tuinenburg, H. R. Knoche, H. Kunstmann, and H. H. G. Savenije, 2013: Should we use a simple or complex model for moisture recycling and atmospheric moisture tracking? *Hydrol. Earth Syst. Sci.*, **17**, 4869–4884, <https://doi.org/10.5194/hess-17-4869-2013>.
- Wang, K., G. Cheng, Y. Ding, Y. Shen, and H. Jiang, 2006: Characteristics of water vapor transport and atmospheric circulation for precipitation over the source regions of the Yellow and Yangtze Rivers (in Chinese). *J. Glaciol. Geocryol.*, **1**, 8–14.
- Wang, N., X. M. Zeng, W. D. Guo, C. H. Chen, W. You, Y. Q. Zheng, and J. Zhu, 2018: Quantitative diagnosis of moisture sources and transport pathways for summer precipitation over the mid-lower Yangtze River basin. *J. Hydrol.*, **559**, 252–265, <https://doi.org/10.1016/j.jhydrol.2018.02.003>.
- Wilson, A. T., M. D. Rintoul, and C. G. Valicka, 2016: Exploratory trajectory clustering with distance geometry. *10th Int. Conf. on Augmented Cognition*, Toronto, ON, Canada, HCI International, 263–274.
- Winschall, A., S. Pfahl, H. Sodemann, and H. Wernli, 2014: Comparison of Eulerian and Lagrangian moisture source diagnostics—The flood event in eastern Europe in May 2010. *Atmos. Chem. Phys.*, **14**, 6605–6619, <https://doi.org/10.5194/acp-14-6605-2014>.
- Wu, R., 2002: A mid-latitude Asian circulation anomaly pattern in boreal summer and its connection with Indian and East Asian summer monsoons. *Int. J. Climatol.*, **22**, 1879–1895, <https://doi.org/10.1002/joc.845>.
- Xu, X., T. Zhao, C. Lu, Y. Guo, B. Chen, R. Liu, Y. Li, and X. Shi, 2014: An important mechanism sustaining the atmospheric “water tower” over the Tibetan Plateau. *Atmos. Chem. Phys.*, **14**, 11 287–11 295, <https://doi.org/10.5194/acp-14-11287-2014>.
- Xue, Z., 1980: A case study of a shear line producing heavy rainfall on Qinghai–Xizang Plateau (in Chinese). *Acta Meteor. Sin.*, **38**, 142–149.
- Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, 2012: APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull. Amer. Meteor. Soc.*, **93**, 1401–1415, <https://doi.org/10.1175/BAMS-D-11-00122.1>.
- Zhang, C., Q. Tang, and D. Chen, 2017: Recent changes in the moisture source of precipitation over the Tibetan Plateau. *J. Climate*, **30**, 1807–1819, <https://doi.org/10.1175/JCLI-D-15-0842.1>.
- Zhang, X., X. Yao, J. Ma, and Z. Mima, 2016: Climatology of transverse shear lines related to heavy rainfall over the Tibetan Plateau during boreal summer. *J. Meteor. Res.*, **30**, 915–926, <https://doi.org/10.1007/s13351-016-6952-7>.