Ecohydrological Model for Grasslands Lacking Historical Measurements II: Confluence Simulations Based on Dynamic Channel Parameters

Mingyang Li[1], Tingxi Liu[1], Limin Duan[1], Long Ma[1], Qiusheng Wu[2], Yixuan Wang[1], Guoqiang Wang[3], Huimin Lei[4], Vijay P Singh[5], Sinan Wang[6], and Junfang Liu[3]

[1]Inner Mongolia Agricultural University
[2]University of Tennessee at Knoxville
[3]Beijing Normal University
[4]Tsinghua University
[5]Texas A&M University
[6]Inner Mongolia Agricultural Univericity

November 24, 2022

Abstract

Technology has greatly promoted ecohydrological model development, but runoff generation and confluence simulations have fallen behind in ecohydrological model development due to limited innovations. To fully understand ecohydrological processes and accurately describe the coupling between ecological and hydrological processes, a distributed ecohydrological model was constructed by integrating multisource information into MYEH. We mainly describe runoff generation and convergence modules. Based on the improved HBV model and degree-3 hour factor method, runoff generation and snow routines were constructed for semiarid grassland basins. In view of meandering and variable steppe river channels and steep hydrological relief characteristics, a confluence module was constructed; the 1-km bend radius equivalent concept was innovatively proposed to unify river channel bend degrees. The daily runoff simulation validation results obtained using two datasets were R^2=0.947 and 0.932, NSE=0.945 and 0.905, and KGE=0.029 and 0.261. In the 3-hour flood simulations, the MYEH model could better restore small long-distance water flows than the confluence method that did not consider actual river lengths or bend energy losses; the MYEH model more accurately simulated the flood peak arrival time than the confluence method that did not consider overflow. The simulated mainstream overflow frequency increased by 0.84/10 years, and significant interaction periods of 10 to 13 years occurred with local precipitation, ecological status and global climate change. An approximately 2-year lag occurred in the global climate change response. This study helps us further understand and reveal the ecohydrological processes of steppe rivers in semiarid regions.

Ecohydrological Model for Grasslands Lacking Historical Measurements II: Confluence Simulations Based on Dynamic Channel Parameters

Mingyang Li[1], Tingxi Liu[1*], Limin Duan[1], Long Ma[1], Qiusheng Wu[2], Yixuan Wang[1], Guoqiang Wang[3], Huimin Lei[4], Vijay Singh[5], Sinan Wang[1], Junfang Liu[3]

([1] *Inner Mongolia Water Resource Protection and Utilization Key Laboratory; Water Conservancy and Civil Engineering College, Inner Mongolia Agricultural University, Hohhot 010018, China;*
[2] *Department of Geography, University of Tennessee, Knoxville, TN 37996-0925, USA;*
[3] *College of Water Sciences, Beijing Normal University, Beijing 100875, China;*
[4] *State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China;*
[5] *Department of Biological and Agricultural Engineering & Zachry Department of Civil Engineering, Texas A&M University, College Station, TX 77843, USA*)

Key Points:

- Using actual river length, bend radius equivalent and overflow data helps improve the steppe river confluence process
- The MYEH confluence module simulates the river diversion effect on the confluence before and after a flood
- Precipitation, ecological status and climate change significantly interact with OFN, and the overflow response to climate has a 2-year lag

Abstract

Technology has greatly promoted ecohydrological model development, but runoff generation and confluence simulations have fallen behind in ecohydrological model development due to limited innovations. To fully understand ecohydrological processes and accurately describe the coupling between ecological and hydrological processes, a distributed ecohydrological model was constructed by integrating multisource information into MY ecohydrological (MYEH) model. We mainly describe runoff generation and convergence modules. Based on the improved HBV model and degree-3 hour factor method, runoff generation and snow routines were constructed for semiarid grassland basins. In view of meandering and variable steppe river channels and steep hydrological relief characteristics, a confluence module was constructed; the 1-km bend radius equivalent concept was innovatively proposed to unify river channel bend degrees. The daily runoff simulation validation results obtained using two datasets were R^2=0.947 and 0.932, NSE=0.945 and 0.905, and KGE=0.029 and 0.261. In the 3-hour flood simulations, the MYEH model could better restore small long-distance water flows than the confluence method that did not consider actual river lengths or bend energy losses; the MYEH model more accurately simulated the flood peak arrival time than the confluence method that did not consider overflow. The simulated mainstream overflow frequency increased by 0.84/10 years, and significant interaction periods of 10 to 13 years occurred with local precipitation, ecological status and global climate change. An approximately 2-year lag occurred in the global climate change response. This study helps us further understand and reveal the ecohydrological processes of steppe rivers in semiarid regions.

1 Introduction

An ecohydrological model is a generalized expression of ecohydrological phenomena and processes created using mathematical language and physical processes (Svoray et al., 2015); these models help researchers describe the interactions between ecological and hydrological processes (Geng et al., 2020) and reveal the succession of ecological patterns and the synergy mechanisms involved in the hydrological cycle as it relates to ecological processes (Wu et al., 2021a). The results of many studies based on model designs and improvements have shown that ecohydrological models exhibit better simulation performances at their respective target scales and ecosystems (Sun et al., 2020; Yan et al., 2021). China's temperate grassland area covers a region spanning 1.68 × 10^6 km^2, accounting for 11.2% of the total global grassland area, concentrated in the semihumid and semiarid areas of northeastern China (Wu et al., 2021b). Grassland ecosystems have suffered degradation due to climate change, excessive grazing, and irrational development (Goenster-Jordan et al., 2021; Yin et al., 2018). Suitable models have been established based on ecohydrology for ecosystems with abundant water in humid and subhumid areas (Zha et al., 2020), alpine mountains (Tong et al., 2021), wetlands (Lou et al., 2019), and deserts (Yin et al., 2021). However, ecohydrological models that are specifically applicable to arid and semiarid steppe regions have rarely been reported. Semiarid grassland ecosystems are relatively barren, although their corresponding vegetation communities are rich and diverse. The evolution of ecohydrological processes, coupling mechanisms, and mutual feedback effects have strong regional characteristics in these regions that cannot be accurately described with existing models. Therefore, in the context of global ecological governance and protection, developing and debugging an ecohydrological model that is specifically applicable to arid and semiarid steppe regions is of great scientific significance (Ma et al., 2019).

Under the action of gravity, water from precipitation or icemelt flows into river networks from ground-surface and underground sources; the water that flows out of the basin outlet section becomes runoff (Betson, 1964; Cadle et al., 1987; Chang & Yeh, 2018; Poiani & Johnson, 1993; Young & Liu, 2015; Zhang & Singh, 2014). The runoff formation process can be generalized into runoff-generating and confluence processes (Gentry & Lopez-Parodi, 1980; Muzik, 1992; Xiong & Guo, 2004). Runoff simulations involve rainfall loss simulations and can be divided into two parts: evaporation and infiltration (processes such as plant interception and hollow depressions filling with dammed-up water resulting in the loss of water into the atmosphere through evaporation or the eventual infiltration process into the soil; here, these processes are not listed separately) (Asdak et al., 1998; David et al., 2005; Jakeman & Hornberger, 1993; Maniquiz et al., 2012). Confluence analyses include the calculation of confluence within a given hydrological response unit and the calculation of river confluence (flood calculus) (Moore & Grayson, 1991; Osborn & Lane, 1969; Vassova, 2013; Wendi et al., 2019). At present, a large number of studies have been carried out on evapotranspiration and infiltration in combination with the rapid development of remote sensing technologies and easily operated field experiments (den Besten et al., 2021; Dunne & Black, 1970; Li et al., 2020; Qiu et al., 2006; Yang et al., 2015). Due to the difficulty of obtaining spatially and temporally continuous confluence process observations, the many influencing factors, and the difficulty of solving partial differential equations of flood waves (David et al., 2019; Hassini & Guo, 2017; Yamanaka & Ma, 2017), both the understanding of the convergence process and the related research are far from sufficient (Hood et al., 2007; Song et al., 2020; Tanaka et al., 2005; Zoccatelli et al., 2019); these inadequacies are even more obvious in semiarid steppe watersheds where rivers meander and are changeable and floods rise and fall steeply.

Two major equations are used to calculate unsteady flow in open channels, the continuity equation and momentum equation; these equations are the basis of the Saint Venant equations (Carraro et al., 2018; Ding & Wang, 2005; Strelkoff, 1970; Wang et al., 2003). By simplifying the continuity equation to the water balance equation and the dynamic equation to the water tank storage relationship in the analyzed reach, the widely used Muskingum method can be deduced for the confluence calculation (Al-Humoud & Esen, 2006; Bozorg-Haddad et al., 2015; Choudhury et al., 2002; Gill, 1978; Tung, 1985). The key to the application of the Muskingum method is determining how to reasonably calculate the k and x parameters, that is, the average propagation time of the analyzed reach and the weight used to measure the effects of inflow and outflow on river storage (Al-Humoud & Esen, 2006; Bozorg-Haddad et al., 2019; David et al., 2015). However, traditionally utilized hydrological variables, such as the average propagation time, are no longer applicable to today's severely degraded steppe rivers. The current models have difficulties when (or are even incapable of) simulating the confluence processes of steppe rivers due to river characteristics such as instantaneous and rapidly changing discharge, sandy soils with low water-storage capacities, and the irregular and easy migration of river patterns (Birkhead & James, 2002; Bozorg-Haddad et al., 2019; Hamedi et al., 2016).

In view of the above existing problems and river characteristics, we constructed a runoff-generating and confluence module in the MY ecohydrological (MYEH) model (Figure 1). We improved the Hydrologiska Byrans Vattenbalansavdelning (HBV) model, which is applicable as a runoff-generating model in arid and semiarid regions, used the degree-3 hour factor instead of the degree-day factor to calculate the snowmelt and accumulation processes, and innovatively proposed a river network confluence module based on dynamic river length, river bend, 3-hour scale unit flood peak duration, and river overflow during flood transit information. Specifically, our objective was to (1) dynamically simulate and depict river flows, river types and other parameter change processes in grasslands; (2) explore and verify the applicability of the MYEH model to different input data sources and determine and explain the physical meaning of each process parameter; (3) compare the advancement of the convergence module with the existing convergence calculation method and explore the resulting space for improvement opportunities; (4) and simulate and explore the responses of river overflows to regional meteorological and ecological conditions and global climate change to further understand and reveal the unique ecohydrological processes of typical steppe regions.

Figure 1. Schematic diagram of natural processes such as the flow convergence, actual river lengths, and channel turns of grassland rivers. Note: The river network shown in the figure does not correspond to the real modelled river network resolution.

2 Method

2.1 Study area

The study area is located in the Xilin River basin (XRB) in the Inner Mongolia Autonomous Region, China (43°30”–44°4″ N, 115°37″–117°30″ E) and is characterized by a continental climate in the middle temperate zone. The annual average temperature in the study area is 2.6 °C, the annual evapotranspiration (ET) is significant, and sunshine is intense. Overall, the terrain is high in the southeast and low in the north, with elevations ranging from 977 to 1620 m (Figure 2a). In the southeastern part of the study area, there is a multilevel platform with a high elevation and a high number of gullies. Many fixed dunes are distributed in the middle of the tributary and the mainstream region. Several of these dunes are semifixed with notable wind erosion. More than 90% of the vegetation is natural foliage, including *Leymus chinensis* Tzvel., *Stipa grandis* P. Smirn., and *Stipa krylovii* Roshev. A certain amount of *Achnatherum splendens* Nevski vegetation can be found in the degraded wetlands and surrounding valleys. Many shrubs, such as *Stipa baicalensis* Roshev. and *Caragana microphylla* Lam., can be found in the higher arid steppe regions. The desert landscape in the central part of the study area is mainly composed of *Ulmus pumila* Linn., whereas *Picea asperata* Mast. and *Betula platyphylla* Suk. are distributed in the northeast region.

According to incomplete statistics, historical measured data in the XRB are relatively scarce. Only one Chinese National Hydrological Station and one Chinese National Meteorological Station had been built in 1964; these stations are located in an urban area and thus have little significance in reflecting the meteorological conditions of the studied grasslands in the historical period. To more accurately monitor the hydrometeorological conditions in the

XRB, we set up 3 sets of automatic velocity and flow monitoring stations, 1 set of Bowen ratio weather stations, 6 sets of automeasuring rain stations and 7 manual flow monitoring stations in the research area. The specific location of each station can be seen in Figure 2b and Table 1 lists the specific station information.

Figure 2. Location, vegetation types (a), topography and stations (b) in the XRB. SBG: *S. baicalensis* Roshev. grassland; LCG: *L. chinensis* (Trin.) Tzvel. grassland; SKG: *S. krylovii* Roshev. grassland; SGG: *S. grandis* P.A. Smirn. grassland; ASG: *A. splendens* (Trin.) Nevski grassland; CMG: *C. microphylla* Lam grassland; AFG: *Artemisia frigida* Willd. grassland; PAG: *P. asperata* Mast. grassland; FSG: *Filifolium sibiricum* (L.) Kitam. grassland; and WCG: weed community grassland.

Table 1. Information of measurement stations in the XRB.

2.2 Model

MYEH model is a bidirectional coupling eco-hydrological model for (but not limited to) steppe inland river basins in arid and semi-arid regions, which is driven by meteorological data and developed by Dr. Mingyang Li and Prof. Tingxi Liu. MY means “my”, which will be released as open source and gradually optimized and updated to get more support from researchers and better improve the model. The MYEH model mainly includes evapotranspiration, runoff, confluence, grazing disturbance, carbon and nitrogen cycle, etc. It absorbs the advantages of various existing ecological models, hydrological models, as well as the framework and algorithm of eco-hydrological models.

The runoff generation and convergence processes are reflected in the MYEH model with two modules: the simulation module (Sim module), which was improved based on the hydrological model (HYMOD) and HBV models (BERGSTRÖM, 1975; Kollat et al., 2012; Moore, 2007; Seibert, 2000), and the self-developed flow confluence module (FLC module). The function of the Sim module is to calculate the flow yield of each grid cell in the basin in units of time using input data such as temperature, precipitation, actual evapotranspiration (calculated by the Eva module in the MYEH model) and grid area data. The FLC module calculates all grid-simulated runoff in the basin according to the river direction generated using basin elevation, river width, river length, roughness and other characteristic data based on the runoff yield and upstream inflow calculated by the Sim module. The Monte Carlo method is used to calibrate the model; this method can not only eliminate any deviation in the calibration process but can also obtain the optimal parameter set. Table 2 lists the parameters, units and rate-setting ranges used by the Sim and FLC modules.

Figure 3. (a) Schematic diagram of the MYEH model simulation (Sim) module; (b) schematic diagram of the MYEH model flow confluence (FLC) module. The full names of the variables shown in Figure 3a can be seen in Table 2. DEM: digital elevation model; RS: remote sensing;

1-km RBRE: 1-km river bend radius equivalent; FTL: flow time length; FFTL: fixed flow time length; and RDacc: accumulated runoff depth.

Table 2. Summary of parameters used in the Sim module and FLC module within the MYEH model.

2.2.1 Sim module

The Sim module mainly includes the snow routine, soil moisture units and flow generation units (Figure 3a). We refer to the degree-day method concept holistically in this module (BERGSTRÖM, 1975). To adapt to the confluence time scale, the 3-hour unit is used to replace the number of days, and the model is improved to a degree-3 hour factor method to improve the simulation accuracy of the diurnal flow generation process. These processes are explained and described below.

2.2.1.1 Snow accumulation & melting routine

The snow routine is a subprogram used to describe the accumulation and ablation of snow, as water is fed into the soil moisture zone through these processes. We treat the snowmelt water in the soil in the same way as we treat rainfall, whereas snowfall on lakes is not treated using snowfall procedures because the pressure effect this snow has on lake ice has the same effect as rainfall on an ice-free lake (BERGSTRÖM, 1975).

The first step is to determine whether precipitation accumulates as snow or directly enters the soil moisture zone as liquid water. A physically correct snowmelt model should consider the entire energy balance of a snowpack, including consideration of sensible and latent heat fluxes, radiation, energy exchanges with the ground, the contribution of precipitation, and the thermal mass of snow itself (Kollat et al., 2012). In view of the uncertainty of the available data and the desire to avoid unreasonable complexity, we adopt the degree-3 hour factor method, representing an improvement from the degree-day factor method.

Temperature is selected as a representative index affecting snow melt. We set a temperature threshold parameter (Ts) to judge the temperature boundary, whether precipitation falls in the form of rain or snow, and whether fallen snow accumulates or melts. Additionally, snowbanks are assumed to retain meltwater, which is expressed as a fraction of their total water storage in terms of the corresponding water holding capacity (CWH) of the snow parameter. Meltwater contained within a snowpack can also be refrozen according to the refreezing parameter (CFR), which is expressed as a fraction of the degree-3 hour factor (CFMAX). See Hamilton et al. (2000) for more details on the formula of the degree-daily snowfall module.

2.2.1.2 Soil moisture accounting routine

The soil water unit calculation performed in the Sim module uses the storage capacity distribution function of a given storage unit. In this module, the storage elements of the analyzed watershed are distributed according to the probability density function defined by the maximum soil water storage and soil water storage distribution. The maximum soil water storage (Cmax) represents the maximum soil water storage capacity, while the shape parameter (BETA) describes the degree of spatial variability in the soil water storage (Wagener et al., 2004).

In contrast from the process involved in HYMOD, in this study, the soil water storage evaporation rate is calculated using the entity views attachment (EVA) module in the MYEH model. After the evaporation fraction is removed, surplus rainfall and snowmelt are used to fill the soil water reserves, and excess rainfall is sent to the flow-producing unit. In addition, we define the soil water storage limit (LP) when potential evaporation occurs. For soil water storage measurements between 0 and LP, the ratio of actual evaporation to potential evaporation changes linearly. For soil water storage measurements greater than or equal to LP, the actual evaporation is equal to the potential evaporation.

2.2.1.3 Runoff generating routine

Similar to the process applied in the HBV model, the flow generating unit of the Sim module involves the conversion of excess rainfall from the soil moisture storage module to the runoff module. The excess rainfall and snowmelt remaining after evaporation, as well as the filled soil water stores, are channeled into an upper response reservoir (UZ). Runoff is divided into three outlets from this upper response reservoir: near-surface flow, confluence and seepage to the base flow. The flows at these three outlets are defined by the near-surface flow regression coefficient (K0), middle flow regression coefficient (K1) and seepage rate (PERC). The threshold parameter (L) defines the runoff height at which near-surface flow occurs in the upper response reservoir. The runoff flowing into the lower response reservoir (LZ) is released according to the base flow regression coefficient (K2). A triangular distribution (MaxBas) is used to convert the runoff released from the reservoir from the top to the bottom, and finally, the runoff producing depth generated by the grid per unit time is obtained.

2.2.2 FLC module

The main work of the FLC module involves summarizing and calculating the runoff producing depth and upstream inflow of each grid cell in the studied basin in units of time according to the flow direction of the river; this work can be mainly divided into three units: inputs, process variable calculations and operation outputs (Figure 3b).

2.2.2.1 FLC module input unit

The input unit mainly includes elevation data obtained by using a digital elevation model (DEM) to calculate the grid flow direction and watershed boundaries, using remote sensing data to extract river features, and runoff producing depth time series calculated by the Sim module.

The flow direction is calculated by inputting the watershed boundaries and grid DEM into the model. According to the extreme value selection principle of, we can obtain the flow direction of the water in each grid cell in the analyzed watershed. While this method can be used to solve most cases, when there are depressions, occlusive lakes or other unique terrains in the basin, the flow direction can form a dead cycle that obviously cannot be satisfied by such a calculation method. Different from the depression-filling tools of the ArcHydro or Soil and Water Assessment Tool (SWAT) model, the idea constructed herein to solve such problems involves initially setting up the outlet of the basin and then determining the flow path of each grid cell to this established outlet. When the path is detected to enter a dead cycle, the module determines the shape of the depression according to the cycle characteristics, looks for the discharge mouth of the depression, and then directs the flow to the mainstream. Through high-resolution remote sensing images and field measurement data, we extracted and prepared the characteristic river quantitative data, including the actual river length, average river width, river bend angle and radius, river roughness, slope and other factors inside each grid cell.

2.2.2.2 FLC module variable-processing unit

The variable-processing unit in the FLC module is mainly used to calculate the channel state and hydraulic parameters of each grid cell during the flow generation period; this unit can be used to debug and perform aggregation calculations at the output unit. At each calculation step for each grid cell, we first calculate the river discharge, flow velocity and river depth using the runoff depth and grid area:

$$RD = Q\Delta t / 1000 A_G \quad (1)$$

$$Q = A_S \times v = W_R \times H_R \times v \quad (2)$$

where RD is the runoff depth (mm); Q is the average flow discharge ($m^3\ dt^{-1}$) in units of time (Δt); A_G is the grid area (km^2); A_S is the sectional area (m^2); v is the flow velocity; and W_R and H_R are the river width and runoff height, respectively.

In the general phase (Figure 3b), we assume that the water flow represents uniform flow in open channels. According to the law of energy conservation, the actual liquid element flow energy equation of rivers in grids should be as follows:

$$z_1 + \frac{p_1}{\rho g} + \frac{v_1^2}{2g} = z_2 + \frac{p_2}{\rho g} + \frac{v_2^2}{2g} + h_w \quad (3)$$

where z_1 and z_2 are the position heads of the inlet and outlet, respectively (m); p_1 and p_2 are the air pressures at the inlet and outlet, respectively ($kN\ m^{-2}$); $\rho = 1000$ is the density of water ($kg\ m^{-3}$); $g = 9.81$ is the gravitational constant ($m\ s^{-2}$); v_1 and v_2 are the initial and end velocities, respectively ($m\ s^{-1}$); and h_w is the total head loss (m). The total head loss can be divided into the frictional head loss (h_f) and local head loss (h_j) as follows:

$$h_w = \sum h_f + \sum h_j \quad (4)$$

$$h_f = \sum \lambda \frac{L_R}{4R} \frac{v^2}{2g}, \lambda = \frac{24}{Re} \quad (5)$$

$$h_j = \sum \zeta \frac{v^2}{2g}, \zeta = \frac{2gL_b}{C^2 R}\left(1 + \frac{3}{4}\sqrt{\frac{b}{r}}\right) \quad (6)$$

$$C = \frac{1}{n} R^{1/6} \quad (7)$$

$$n = (n_0 + n_1 + n_2 + n_3 + n_4) \times m_5 \quad (8)$$

where λ is the frictional head loss coefficient, which can be calculated using an empirical formula including the Reynolds number (Re); R is the hydraulic radius (m); ζ is the local head loss coefficient; L_b, b and r are the length (m), width (m) and bend radius (degree) of the river curve, respectively; C is the Chezy coefficient ($m^{1/2}\ s^{-1}$); and n is the channel roughness, which can be calculated using Eq. 8. In Eq. 8, n_0 to n_4 represent the basic roughness of natural channels, the influence of irregular water surfaces, the influence of changes in the channel cross section shape and size, the influence of water-blocking substances and the influence of plants, respectively; and m_5 is the river-winding coefficient, which is equal to 1 in our research.

Since the river bend degree is not similar to the other variables, it is difficult to unify the variables related to river bends, so we proposed the concept of the 1-km bend radius equivalent and converted the length sum of each river bend to the same magnitude to unify the river bend degree in the analyzed basin. Therefore, the total bending length of 1km bending radius equivalent L'_b in the grid can be expressed by the bending radius R_b and bending angle r of each bend:

$$L'_b = \sum \frac{r}{360°} \times 2\pi \times \frac{R_b}{1km} \quad (9)$$

To more realistically reflect the characteristics of grassland rivers, we set the overflow coefficient to determine whether overflow occurs when a flood peak passes according to the real-time river depth. When a flood phase occurs (Figure 3b), the raster channel is reset to a state with no bend and a base river length. After the flood passes (as represented by the recovery phase in Figure 3b), the river gradually begins to bend with the influence of the geostrophic deflection force and other factors; that is, the river length gradually recovers to the actual river length, and curved reaches reappear. The river length, curve length and bending angle of the three periods can be expressed as:

$$f(L_R, L'_b, r) = \begin{cases} Max(L_R, L'_b, r), & General\ phase \\ Min(L_R, L'_b, r), & Flood\ phase \\ {}^{t_m}/_{t_n}\ Max(L_R, L'_b, r), & Recovery\ phase \end{cases} \quad (10)$$

where, t_m and t_n are respectively the time from the last overflow to the present and the total time it took for the river to recover to bend.

2.2.2.3 FLC module operation and output unit

The operation and output unit summarizes the parameters calculated by the first two units at each moment, calculates the time and amount of flowing water moving to the next grid cell, and iteratively describes the flow situation of each section of the basin in the whole simulation period layer by layer. First, through the flow direction, we can calculate the number of grid layers j needed for each grid point to flow to the drainage outlet of the basin. If the row and column numbers of the watershed grid points are set as m and n, respectively, then the grid point layer being processed can be expressed as $m(j)$ and $n(j)$, respectively. The flow in a given grid cell at moment t can be expressed as $Q(t)_{m(j),n(j)}$, and the time (Δt) of the runoff flow to the next grid point at this moment can be calculated as follows:

$$\Delta t = \frac{L_R}{\bar{v}} = \frac{L_R}{0.5 \times (v_1 + v_2)} \quad (11)$$

where L_R is the river length and $\bar{v}$ is the average discharge velocity. Since our unit time is 3 hours, when the runoff time is not an integer, we divide the flow according to the integer time so that the flow out of the grid at time t is $q(t)_{m(j),n(j)}$:

$$q\big(t + fix(\Delta t)\big)_{m(j),n(j)} = Q(t)_{m(j),n(j)} \times \frac{fix(\Delta t)}{\Delta t} + Q(t-1)_{m(j),n(j)} \times \frac{\Delta(t-1) - fix\big(\Delta(t-1)\big)}{\Delta(t-1)} \quad (12)$$

where fix is a downward rounding function.

The above equations represent the case in which upstream grid inflow is not considered. When upstream grid inflow is present, we first calculate the initial flow obtained by the grid cell as follows:

$$Q(t)_{m(j),n(j)} = Qsim(t)_{m(j),n(j)} + \sum_1^{dir} q(t)_{m(j+1),n(j+1)} \quad (13)$$

where $Qsim(t)$ is the flow rate in each grid cell calculated by the flow generation module and $dir = 1\ to\ 7$ represents 1 to 7 upstream convergence directions. Notably, a given grid cell has a total of eight possible directions: north, northeast, east, southeast, south, southwest, west, and northwest. In the confluence process, if the water flowing from all eight directions flows into the central point, we regard this grid cell as a depression. When the water surface exceeds the lowest surrounding elevation within the grid, it is discharged in this direction; please refer to section 2.2.2.1 for details.

2.3 Validation

2.3.1 Verification system

To verify the accuracy and applicability of the MYEH model, we adopted dual-drive data source adaptation, traditional model comparison and measured data inspection methods. The dual-drive data source adaptation used in this study refers to the drive data generated under two different observation systems, the China Meteorological Driven Data Set (CMFD) and Global Land Data Assimilation System Noah Land Surface Model L4 (GLDAS-Noah); these datasets are brought into the MYEH model to calculate and simulate the ecohydrological process of the XRB. The comparison with the traditional model is mainly reflected in the confluence model calculation, in which the length of each grid cell is fixed and the head loss of the river is not included. The measured data test includes a comparison and verification of the daily discharge data collected at the Chinese National Hydrological Station in the basin, the discharge data recorded at the self-built automatic detection hydrological stations at three river sections and the artificially measured real-time discharge data obtained through value simulations (Figure 2b).

2.3.2 Multiobjective calibration

To evaluate the MYEH model simulations in the studied semiarid grassland more comprehensively, this paper selects several evaluation indexes. We used the coefficient of determination (R^2), Nash-Sutcliffe efficiency coefficient (NSE) (Nash & Sutcliffe, 1970), bias between the simulated and measured values (Bias), transform root mean square error (TRMSE), mean absolute error (MAE), and Kling-Gupta efficiency (KGE) to quantify the mismatches between the simulated and tested data. These metrics can be expressed as follows:

$$R^2 = 1 - \frac{\sum_{t=1}^{N}\left(Q_{s,t} - \bar{Q}_{o,t}\right)^2}{\sum_{t=1}^{N}\left(Q_{o,t} - \bar{Q}_{o,t}\right)^2} \quad (14)$$

where $Q_{s,t}$ and $Q_{o,t}$ are the simulated and observed runoff, respectively, at time t and $\bar{Q}_{o,t}$ is the mean of the observed and predicted data over the calibration period.

NSE compares the predicted values to the 1:1 line between the measured and predicted values rather than the best regression line through the points. NSE values range from 1 (optimal) to $-\infty$, and this metric been frequently used as a hydrologic model calibration objective:

$$NSE = 1 - \frac{\sum_{t=1}^{N}\left(Q_{o,t} - Q_{s,t}\right)^2}{\sum_{t=1}^{N}\left(Q_{o,t} - \bar{Q}_{o,t}\right)^2} \quad (15)$$

$$Bias = \frac{1}{N}\sum_{t=1}^{N} \bar{Q}_{s,t} - E\left(Q_{o,t}\right) \quad (16)$$

where $E\left(Q_{o,t}\right)$ is the expected observed value.

Following prior studies (Misirli et al., 2003; Tang et al., 2007), one of the objectives analyzed herein emphasizes low flow errors using the Box-Cox (Box & Cox, 1964) TRMSE, as shown in equation (12):

$$\text{TRMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{Q}_{s,t} - \hat{Q}_{o,t}\right)^2}\ , where\ \hat{Q} = \frac{(1+Q)^{\lambda} - 1}{\lambda} \quad (17)$$

where $\hat{Q}_{s,t}$ is the Box-Cox-transformed simulated runoff at time *t* and $\hat{Q}_{o,t}$ is the Box-Cox-transformed observed runoff at time step *t*. The summation is performed from time step 1 through the number of time steps in the calibration period (N). $\hat{Q}$ represents the Box-Cox-transformed runoff value *Q*, where $\lambda = 0.3$. The Box-Cox transformation, in addition to emphasizing low flow periods, also serves to reduce the impacts of heteroscedasticity in the RMSE calculation.

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|Q_{o,t} - Q_{s,t}\right| \quad (18)$$

The KGE (Eq. 17) (Gupta et al., 2009; Knoben et al., 2019) is based on the decomposition of NSE into its constitutive components (correlation, variability bias and mean bias), addresses several perceived shortcomings in NSE (although there are still opportunities to improve the KGE metric and to explore alternative ways to quantify model performances) and is increasingly used for model calibrations and evaluations:

$$KGE = \sqrt{(1-\gamma)^2 + (1-\alpha)^2 + (1-\beta)^2} \quad (19)$$

where γ, α, and β are the linear correlation coefficients, the ratio of their standard deviations and the ratio of the mean values of the simulated and measured values, respectively. Additionally, P values were used to test the sample variance in the measured and simulated values, and the significance level was set to 0.01. When the p value was less than 0.001, there was a highly significant difference.

2.4 Overflow frequency analysis

To further understand the steppe river overflow phenomenon, the frequency and locations of overflow events and the vegetation status of the basin during the whole study period are assessed. We use a cross-wavelet analysis to study the periodic interactions between monthly overflow times and precipitation, the leaf area index (LAI) to reflect the vegetation status, and the southern oscillation index (SOI) and sea surface temperature (SST) to reflect climate change in the NINO3.4 region. The Morlet wavelet is selected as the wavelet type, and the confidence level is set to 95%. For detailed introductions of the wavelet transform, see Sang (2013) and Nourani et al. (2014).

3 Data

The data used in this paper can be divided into product data and measured data, which mainly include meteorological data, remote sensing data and verification data. The spatial resolution and time span of these data are shown in Table 3.

The meteorologically driven data include CMFD and GLDS-NOAH data, among which CMFD mainly includes 2-m temperature (T), precipitation (P), relative humidity, 10-m wind speed, longwave and shortwave radiation and air pressure data (Yang et al., 2010). Since the temporal coverage of the CMFD does not include 2019 or 2020, the meteorologically driven data representing these two years are obtained by spatial interpolation using data from the self-built stations. The GLDAS-NOAH data are obtained from NASA's Global Land Data Assimilation System (Beaudoing & Rodell, 2020; Rodell et al., 2004). To match the simulation time, we used GLDAS-2.0 data from 1980 to 2000 and GLDAS-2.1 data from 2001 to 2020. For the remote sensing data, Leaflet through the open-source JavaScript library and high-resolution Google historical satellite images downloaded for interactive mapping were used. The image tile level was 17, and the spatial resolution was 2.15 m. The verification data included data recorded at China's national stations, self-built hydrological weather stations and measured artificial river flow data (see Figure 2b and Table 1 for the specific location and information of each station). The vegetation data were obtained using Global Land Surface Satellite (GLASS) product data statistics (Liang et al., 2013a; Liang et al., 2013b); this product is spatially and temporally continuous, without gaps or missing values, and the wideband longwave emissivity product is the first product in the world with an 8-day temporal resolution and 1-km spatial resolution (Liang et al., 2020). The SOI and SST data were provided by the official website of the Bureau of Meteorology of the Commonwealth of Australia 2021 and Weatherzone based on data from the Bureau of Meteorology.

Table 3. Characteristics of the two meteorological datasets.

4 Results and discussion

4.1 Simulation validation of the MYEH runoff generation and confluence processes

The meteorologically driven data contained in the CMFD and GLDAS-NOAH datasets were introduced into the MYEH model to simulate the runoff generation and confluence processes in the XRB from 1980 to 2020, and the river discharge simulated by the model was verified using the station data shown in Figure 2b. The results are shown in Figure 4 and Table 4.

Figure 4b is a Q-Q plot analysis showing the daily river discharge measured at the XRB National Hydrographic Station section. According to the kurtosis and skewness of the daily discharge data, it is not difficult to see that the XRB river discharge presents skewed distribution characteristics, indicating that the discharge at this section is far below the mean value. The daily discharge measurements below the mean value are much higher than the expected value of 0.554 $m^3\ s^{-1}$, indicating that slow surface runoff is normal in the XRB and reflecting the characteristics of a trickling grassland river in a nonflood period (Coe et al., 2011; Metivier et al., 2016).

The MYEH model performs well when simulating river runoff using two kinds of meteorologically driven data. From the perspective of evaluation indexes, the R^2 and NSE values were both greater than 0.9, and the KGE value was less than 0.3, indicating that the MYEH model performed well when controlling the trend of the overall ecohydrological process (Figure 4a). The daily discharge simulation results obtained using the two meteorological driven datasets show that the discharge simulated using CMFD was slightly better than that simulated using the GLDAS-NOAH dataset in the daily flood peak simulations; specifically, some deviation occurred in the maximum daily flood peak outputs within the year. The test results of the three automatic hydrological stations show that the simulated results were relatively accurate (Figure 4c-e), and the corresponding R^2 and NSE values were slightly lower than the test results obtained at the national hydrological station section. The manual flow measurement results show that the NSE value of the simulated and observed flow values is high, while the corresponding R^2 value is low. The distribution of the scatter diagram is relatively convergent, and the linear fitting and 1:1 line are also relatively consistent, indicating that the overall results are reliable and the process simulation error is small.

Figure 4. (a) Comparison of the simulated and measured daily sectional discharge at a national hydrological station obtained using two meteorologically driven datasets with the MYEH model; (b) Q-Q plot of daily discharge in the XRB; (c-e) comparison of the simulated and measured daily sectional discharge at three automatic hydrological stations using two meteorologically driven datasets with the MYEH model; (f) comparison of the simulated and measured daily sectional discharge at seven measuring sites using the CMFD; and (g) comparison of the simulated and measured daily sectional discharge at seven measuring sites using the GLDAS-NOAH dataset. The green points in (a) to (e) are the values observed at the national hydrological station.

Table 4. Six evaluation value of simulated runoff in XRB using two data sources.

The results show that the MYEH model combined with the evapotranspiration and runoff generation and confluence modules can effectively simulate the runoff process of each grid section representing the grassland river channel. Combined with the river location of the three inspection sections, the downstream flow simulations conducted in the basin are also better than those in the upstream flow-producing area; this result can be mainly summarized as the influence of different flow-producing methods and measurement accuracy insufficiencies. The runoff generation modes of steppe rivers can be mainly divided into two types: runoff generation on the mountain slopes and groundwater outcropping in front of mountains (da Silva et al., 2018; Gupta et al., 2019). Most of the runoff generation models in the basin conform to the Sim precipitation and soil water storage module (Zhang et al., 2021), and only a few grid points contain outcropping groundwater (although such areas can also be simulated using the soil water storage principle) (Liang et al., 2012; Wagener et al., 2004); however, some deviations exist (Li et al., 2015; Lopes & Canfield, 2004). On the other hand, the grassland river characteristics in the upstream flow-producing areas mostly represent wetlands (floodplains) with soil water contents close to or at saturation (Tang et al., 2020; Wang et al., 2014). In the actual flow measurements collected in such areas, although we selected river sections in wetlands (floodplains) that met the flow measurement standards, the verification data did not contain mid-soil flow information, further leading to verification errors (Bendjoudi et al., 2002; Wagener et al., 2004).

Compared with the validation results of the data recorded at the two hydrological stations, the tested accuracy of the manual flow measurements was the lowest; this result can be summarized with two reasons. Firstly, due to the rich data of hydrology station and automatic flow measurements, the model will be inclined to the site with rich data when scaling parameters, while there are only a few measured data of manual flow measurement, so the weight of the data will be reduced and the error will become larger. The other reason is also related to the characteristics of steppe rivers; during the flood period, sandy riverbeds do not easily maintain stable shapes (Staudt et al., 2019), and semiarid grasslands experience high wind speeds in both the spring and summer rainy seasons (Li et al., 2021), thus affecting the river flow measurements.

4.2 Parameter optimization and sensitivity analysis

Model parameters can be defined as quantities that are used to represent the physical or ecohydrological characteristics of a watershed and remain constant during the simulation process (Melsen & Guse, 2019; Pfannerstill et al., 2015; Qi et al., 2019). The optimization of parameters in the MYEH model can make automatic adjustments using a variety of evaluation indexes to ensure that the simulated and observed runoff values match well (Song et al., 2012). Figure 5a-c describes model parameter optimization process through the use of three evaluation indexes (due to a large number of iterations, only partial results are shown in the figure). The results showed that the related variables of some runoff production modules dominated by soil water migration still showed convergence trends even under different rating indexes; this directly reflected the characteristics of the basin among the parameter values (Huang et al., 2015a; Yokoo & Kazama, 2012).

The parameters that control soil, snowmelt and river channels are all important input variables in ecohydrological models, and subtle changes in these parameters directly affect the stability of the models. Therefore, it is particularly important to discuss the influence of the parameters utilized in each module on the practical applications of the model (Guse et al., 2016; Pfannerstill et al., 2015). The parameters considered in the Sim module and snow routine unit mainly affect the change in yield over time (Croke & Jakeman, 2004; Huang et al., 2015b), and the FLC module parameters directly affect runoff collection (Reaney et al., 2014). All three of these parameter groups alter the flood propagation process to a certain extent. Therefore, the average simulation results obtained for these two modules and one routine unit were increased or decreased by 1, 2, 5, 7.5, 10, 12.5, 15, 20, 25, 30% and no change, and a total of 21 conditions were analyzed, respectively. The results are shown in Figure 5d-f.

The parameter sensitivity analysis results are all within the acceptable range. Among them, changes in the snow routine parameters had the smallest impact on the simulated runoff. When the change range was greater than 5%, an increase in the snow process parameters had a greater impact on the simulation accuracy than a decrease in the snow process parameters. The parameters of the Sim and FLC modules were much more sensitive than those of the snow routine unit. When the variation range of the parameters of the above two modules exceeded 5%, the simulated R^2 and NSE values dropped to approximately 0.7 and 0.55, respectively (Figure 5d-e). The KGE index shows that when the variation range exceeded 10, the model accuracy significantly decreased (Figure 5f).

Figure 5. Parameter optimization (a-c) and parameter sensitivity analysis (d-f) results obtained for the MYEH model.

4.3 Applicability analysis

The applicability analysis conducted in this paper focuses on the universality of different driving datasets in the XRB. The results show that the runoff simulation results obtained using two different meteorological driving datasets were basically distributed on the 1:1 line between the R^2 and NSE values, and the TRMSEs of most simulated values were less than 0.6 (Figure 6). As a test index considering correlation, variability bias and mean bias, the KGE values remained between 0 and 0.4 on the whole, and the closer the R^2 and NSE values of the two simulated values were, the better the KGE evaluation result was.

In the simulation comparison, the errors resulting from data sources increased when the lower runoff or base flow were simulated (Balin et al., 2010; Faramarzi et al., 2015; Sikorska et al., 2015). In addition, collapse phenomena with high R^2 values but poor NSE, KGE and TRMSE values occurred rarely in both the non-icebound period and the icebound period but occurred slightly more frequently in the non-icebound period than in the icebound period. These outliers indicate that although the results conform to the change rule in the whole time series, there is a certain deviation. The main reason for this phenomenon is that there a certain difference exists in the precipitation data between the two meteorologically driven datasets (Renard et al., 2011; Schoups & Nasseri, 2021), resulting in consistent flood peak occurrence times (consistent with the temporal rules) in the runoff generation simulations but deviating runoff (base flow) flood peak values.

Figure 6. Comparison of the MYEH model-simulated runoff discharge during the nonfreezing period (a) and base flow during the freezing period (b) in the XRB as determined using the CMFD and GLDAS-NOAH data sources. In this figure, NSE and R^2 are plotted on the *X* and *Y* axes, respectively, KGE is plotted in color, and TRMSE is plotted using the size of the markers. The black arrow points in the direction of decreasing flow or base flow. The red arrow indicates the tendency of both data-source simulations to collapse. NSE: Nash-Sutcliffe efficiency; KGE: Kling-Gupta efficiency; and TRMSE: Box-Cox transformed root mean square error.

4.4 Flood process

In view of the good applicability and strong stability of the MYEH model in the XRB, we further investigated the confluence mode of the FLC module (runoff in this mode is referred to as Qs) and two common confluence modes (we called the runoff discharge in these two confluence modes Qs1 and Qs2). The confluence model that does not consider the actual river length, river bend or overflow and the confluence model that considers the actual river length or river bend but does not consider overflow were compared and analyzed in their simulation of the four flood modes. First, we selected two 1/20-year frequency floods and two 1/50-year frequency floods in the simulation period. Two driving datasets and three confluence modes were used to simulate the flood process at the 3-hour scale. Yellow and red five-pointed stars were used to indicate the initial times at which overflows started in the tributaries and in both the tributaries and mainstream. The flow data measured at the XRB National Hydrological Station section are shown in Figure 7.

The results show that the flood peak times simulated by the two datasets at the diurnal scale are basically the same, and only the peak flood value differs slightly; this is consistent with the previous results regarding the universality of the XRB for different driving datasets. On the whole, Qs1, which does not consider the actual river length, river bend or flood, resulted in the fastest flood arrival time and the shortest flood duration. Qs2, which considers the actual river length or river bend, resulted in the latest flood arrival time and the longest flood duration. Qs, as simulated by FLC, had outputs in the middle of the two results described above (Figure 7).

Both floods that occurred in 1987 were triggered by single heavy rain events. The rainfall event that occurred on August 11th was short but intense, while the rainfall event that occurred on August 26th was light but prolonged. Accordingly, on August 11th, runoff reached the flood peak within 3 to 6 hours, while the runoff peak measured on August 26th was not as urgent as the former (Figure 7a). In 1998, many basins in China experienced extensive regional floods, and precipitation in the XRB was abundant. The year 1998 mainly included four floods caused by continuous precipitation, among which two floods showed a bimodal pattern due to short interruptions in precipitation (Figure 7b). In 2004 and 2012, extremely rare heavy rains occurred and caused extreme flood events. The runoff simulation results of different confluence modes also showed similar differences in these years. The flood peak of the Qs1 mode was 3 to 6 hours earlier than that of the FLC mode, while the flood peak of the Qs2 mode was 3 to 6 hours later than that of the FLC mode. The flood waveform and numerical runoff characteristics simulated by the three modes were basically consistent. In particular, when river overflows occurred, the flood peak value simulated by the FLC mode was slightly higher than that simulated by the Qs1 mode (Figure 7c, d).

Figure 7. Simulations of the 3-hour flood process under three confluence modes using the CMFD and GLDAS-NOAH data sources. Figures 7(a) to 7(d) show monsoon floods in 1987, 1998, 2004 and 2012, respectively. Qs indicates the MYEH model confluence mode (FLC). Qs1 indicates the confluence mode in which the actual river length, river bending and overflow are not considered. Qs2 indicates the confluence mode in which the actual river length and river bending are considered but overflow is not considered. The orange and red stars represent the overflow of tributaries and the overflow of main streams and tributaries in a flood event, respectively.

Different confluence modes cause different flood arrival times, flood peak values and even flood waveforms (Gao et al., 2004; Wagener & Montanari, 2011). Through a comparison of the three confluence modes, it can be seen that considering the actual river length and river bend can result in more realistic simulations of steppe river network characteristics. However, if the overflow situation is not considered, the arrival time of the flood lags behind, and this situation is more obvious when the flood peak is larger. As a prominent feature of steppe rivers, channel overflow events not only advance the arrival times of flood peaks but also increase the flood peak values to a certain extent. To refine and decompose the impacts of overflow events, we divided the overflows into tributary overflows and mainstream overflows. Since the mainstream is wider and deeper than the tributaries, we found that overflows first occur in tributaries during the whole simulation period and then occur in the mainstream when the flood becomes sufficiently large.

First, we take 1987 and 1998 as examples to study the influence of tributary overflows on the confluence of steppe rivers. The two minor floods that occurred on August 7th, 1987, and June 2nd, 1998, showed that the difference between Qs and Qs1 was mainly that the overflow event occurred slightly earlier in Qs1 than in Qs; the two overflow values were basically consistent. When tributary overflow occurred, the Qs-derived peak value basically exceeded the Qs1-derived value, mainly reflecting the influence of river overflow on the runoff flood peak (Figure 7a, b). In the extreme flood events of 2004 and 2012, when overflows occurred in both the tributaries and mainstreams, not only did the Qs-derived peak exceed the Qs1-derived peak, but the slope of the simulated runoff also gradually increased, and the arrival time of the flood peak continually approached the Qs1-derived linear confluence value, reflecting the influence of river overflows on the arrival time and value of the flood peak (Figure 7c, d). In terms of the overflow process, the length of the flow path was shortened and the flow velocity was reduced by river bends, thus enabling the flow to converge more quickly to the downstream section (Cervantes et al., 2020; Knighton et al., 2014). A shorter river distance serves to reduce losses associated with evaporation, infiltration and other processes and improves the flood peak value compared with that derived using the mode that does not consider overflow (Krasnostein & Oldham, 2004).

4.5 Response analysis of overflow frequency to climate and ecology

Simply speaking, the overflow of river channels is a special situation in which an abundant inflow of water from the upstream region leads the river to overflow, thus disturbing the channel parameters and influencing the confluence of river networks. From the perspective of the hydrological function of a basin, river overflows are extremely destructive, as they lead not only to frequent riverbank collapses and diversions but also easily lead to extreme hydrological events such as decreased storage capacities, a steeply rising floods, and increased river sediment loads. In addition, from the perspective of vegetation ecology, river overflows can also lead to swamp conditions in valley wetlands and community succession in ecosystems dominated by plants and microorganisms through the resulting changes in soil moisture, ion concentrations and nutrient availabilities. Although it is difficult to directly define or judge the advantages and disadvantages of these succession processes, these processes represent another scientific question we hope to explore with the help of the proposed grassland watershed ecohydrological model.

River overflows are not only directly related to precipitation but are also related to the regional vegetation and river stability. First, according to our simulation of the ecohydrological process in the XRB, we created a diagram of the annual average overflow frequency in the study area and the correlation distribution between overflows and the regional vegetation status (Figure 8a). The overflow frequency results showed that the average annual overflow frequency was more than once a year in the upper reaches of the analyzed basin, especially in the river channel in the northeastern part of the study area and in the Hilltara wetland in the central and eastern parts of the study area, as the river channel in the upper reaches of the river and the wetland were relatively shallow and prone to overflowing (Bornette & Amoros, 1991). The northwestern part of the study area contains the Xilinhot Reservoir, which has low terrain and a large catchment area, so we did not analyze this region.

The correlation between overflows and ecological conditions showed that the overflow frequency in the mainstream (OFN) was strongly correlated with the vegetation conditions, and the correlation between LAI and the mainstream grid prone to overflowing exceeded 0.5. The correlations between OFN and the vegetation status in tributaries and nonmainstream streams were higher in the south and lower in the north. According to the precipitation trend analysis, from 1980 to 2020, LAI and OFN as well as precipitation (17.18 mm $10a^{-1}$) and OFN (0.84 $10a^{-1}$) showed obvious increasing trends (Figure 8b), while LAI, which represents the vegetation conditions, showed a slight decreasing trend (-0.04 $10a^{-1}$). Moderate overflow will improve the ecology of wetland vegetation in the valley. For example, snowmelt and ice-melt runoff in spring will increase the soil moisture content of wetland in the valley after overflow, making it easier for the wetland to turn green. It is believed that the real cause of wetland vegetation degradation in XRB valley should be the combination of riparian vegetation degradation and stunting caused by overgrazing. These results were consistent with the conclusion that the increase in OFN was related to the increase in precipitation and the degradation of vegetation analyzed in the previous section; further, these results are also consistent with the conclusion reached by Xu et al. (2009) in their study on the effect of the riparian vegetation ecological status on overflow events in the lower reaches of the Tarim River, another arid region.

To further study the response trend of OFN to the environmental changes that have occurred in the last 41 years, we conducted periodic analyses of OFN with precipitation, vegetation, SOI, and SST in the NINO3.4 region using cross-spectrum analysis techniques. The significant cross-wavelet energy results obtained between OFN and precipitation, LAI, SOI and SST were mainly distributed in periods from 5 to 7 years and from 10 to 13 years (Figure 8c-f), among which the 10-13-year period was the most significant, indicating that overflows in the XRB have a strong corresponding relationship with global climate change, such as El Niño changes, and that these two processes are closely related (He et al., 2015; Kundzewicz et al., 2010; Minville et al., 2010).

In addition, in some years, precipitation and LAI also had significant and strong interactions with OFN in periods ranging from 1 to 4 years and from 2 to 4 years, respectively, further indicating that overflow events strongly interacted with precipitation and LAI with short periods. In the strong interaction cycle lasting 10 to 13 years, the interactions between OFN and precipitation and between OFN and LAI were in the positive phase; that is, no lag effect was observed between the overflows and local meteorological or vegetation conditions. The phase difference between OFN and SST was approximately 30°, indicating that overflow events in the XRB have a lag period of approximately 2 years in response to global climate change.

Overall, the MYEH model that considers river overflow events helps us to understand that the special steppe river overflow phenomenon is closely related to the local precipitation, vegetation status, global climate change and other factors; further, this model helps reveal the unique ecohydrological processes and response mechanisms of typical steppe ecosystems.

Figure 8. (a) Average annual overflow frequency and the correlation between overflows and vegetation status. (b) Trend analyses of precipitation, LAI, and OFN from 1980 to 2020. (c-f) The cross-wavelet energy spectrum analyses of the OFN with precipitation, LAI, SOI and NINO3.4 SST. The 5% significance level against red noise is shown as a thick contour line. The relative phase relationships are shown as arrows (with in-phase relationships pointing right and anti-phase relationships pointing left).

4.6 Existing problems and uncertainty analysis

4.6.1 Refine the overflow process

Although we optimized the grassland river confluence process by setting the overflow coefficient and other methods, we still found that the simulated flow value at the maximum flood peak time was slightly higher, while the flood peak was slightly lower at the later time units (Figure 4); these results indicated that our confluence speed simulation results were still overestimated to some extent. A further subdivision of the overflow process (type) may be helpful for obtaining a more detailed optimization. First, do overflow events permanently reset channels? This question corresponds to the dynamic treatment of the recovery period following diffuse flow conditions. The second step is to distinguish overflow events into temporary and dam break overflow. These errors tend to focus on the flow increases caused by the summertime rainy season and the spring flood caused by springtime snowmelt. Small overestimations can be seen in both utilized datasets (Figures 4 and 7). In a more refined river channel description, it is critical to optimize the confluence process in the future to determine whether the overflow coefficients can be graded and the sensitivity of the analyzed reach can be graded, aiming to realize dynamic overflow simulations inside the river network grid cells for flood control in small grassland river basins. In addition, overflow will affect soil moisture and vegetation, so the ecological process simulation of a watershed, especially the simulation of wetland ecological process after overflow, is also the direction for further optimization and improvement.

4.6.2 Optimize the parameter selection system

The MYEH model constructed in this study adopts the overall parameter adjustment principle and comprehensively judges the results using multiple evaluation indexes to select the simulation results with the highest accuracies, smallest errors and most physically significant parameter combinations as effectively as possible. Such an evaluation system is considered comprehensive but still has room for improvement. For example, a variety of verification methods that aim to ensure data accuracy are included in the evaluation system;

the evapotranspiration and production confluence module parameters are adjusted simultaneously; and the simulation results are evaluated. In addition, simulation process parameters such as the soil moisture content could also be included in the evaluation system to improve the ecohydrological process simulation accuracy.

4.6.3 Subsequent module design

Through multiple ecological-hydrological process simulations, we found that a certain error still exists when simulating the water balance in grassland wetlands. On the one hand, due to the lack of frozen soil simulations, the water resulting from springtime snowmelt in the flood season cannot undergo large-scale penetration or flow into the soil, thus leading to evapotranspiration underestimations; on the other hand, due to this omission, groundwater recharge is not considered. For the subsequent expansion of the MYEH model, we plan to take two steps: one step involves improving the FLC module and building frozen soil and groundwater modules by summarizing and combining existing problems; and the other step involves considering more ecological-hydrological processes and building modules to represent the nutrient element cycle, plant growth, grazing disturbances and so on.

5 Conclusion

Aiming to represent runoff in a semiarid steppe basin with variable meandering rivers and steep flood flows, we simulated the 3-hour runoff process in the XRB from 1980 to 2020 by constructing runoff generation and convergence modules in the MYEH model to consider the dynamic actual river length, river bend and overflow characteristics and discussed the occurrence frequency and influencing factors of steppe river overflow events. The results show that the MYEH model has a high accuracy and stability when simulating the ecohydrological process and can also simulate changes in river overflows, flood peaks and arrival times caused by the passage of large flood events. With the use of an appropriate eco-hydrological model, it is helpful to further reveal the special phenomenon of the overflow of steppe rivers. Vegetation degradation caused by overgrazing and the increase of precipitation in the basin are the main reasons for the increase of the overflow number of XRB, and there is a two-year lag between the overflow number and global climate change factors.

Acknowledgments: This research was funded by the National Natural Science Foundation of China (Nos. 51939006, 51620105003 and 51909122), the Inner Mongolia Major Science and Technology Projects (Grant Nos. 2020ZD0009 and 2019ZD007), the Inner Mongolia Science and Technology Plan Project (Nos. 2020 and 2021GG0071), the Ministry of Education Innovative Research Team (No. IRT_17R60), the Innovation Team in Priority Areas Accredited by the Ministry of Science and Technology (No. 2015RA4013), the Inner Mongolia Industrial Innovative Research Team (No. 2012), and the Research and Innovation Funding Project for Graduate Students (BZ2020069). We are grateful to the principal investigators and the teams behind all the datasets used in this study. The CMFD datasets are available at http://data.tpdc.ac.cn/en/data/8028b944-daaa-4511-8769-965612652c49/. The GLDAS Noah

datasets are available at https://ldas.gsfc.nasa.gov/gldas/. Thanks for Leaflet, the leading open-source JavaScript library for mobile-friendly interactive maps, which is available at https://leafletjs.com/index.html. SOI dataset are available at http://www.bom.gov.au/climate/enso/soi/. SST dataset are available at https://www.weatherzone.com.au/climate/indicator_enso.jsp?c=nino34&p=monthly. All generated methods used in this study can be downloaded from https://github.com/myli1993/MYEH_model_ver1.0. And generated data can be downloaded from https://zenodo.org/record/5578630#.YW7PgRy-uHs. The authors declare no conflict of interest.

Author Contributions: M.L. developed the initial and final versions of this manuscript and analyzed the data. T.L., L.M., L.D., Q.W., Y.W., G.W., H.L., and V.S. contributed their expertise and insights to oversee the analysis. S.W. and J.L. helped complete the partial data preprocessing.

References

Al-Humoud, J. M., & Esen, II. (2006). Approximate methods for the estimation of Muskingum flood routing parameters. *Water Resources Management, 20*(6), 979-990. doi: 10.1007/s11269-006-9018-2

Asdak, C., Jarvis, P. G., van Gardingen, P., & Fraser, A. (1998). Rainfall interception loss in unlogged and logged forest areas of Central Kalimantan, Indonesia. *Journal of Hydrology, 206*(3), 237-244. doi: 10.1016/S0022-1694(98)00108-5

Balin, D., Lee, H., & Rode, M. (2010). Is point uncertain rainfall likely to have a great impact on distributed complex hydrological modeling? *Water Resources Research, 46*. doi: 10.1029/2009wr007848

Beaudoing, H., & Rodell, M. (2020). GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1. *Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [2021-7-31].* doi: 10.5067/E7TYRXPJKWOQ

Bendjoudi, H., Weng, P., Guerin, R., & Pastre, J. F. (2002). Riparian wetlands of the middle reach of the Seine river (France): historical development, investigation and present hydrologic functioning. A case study. *Journal of Hydrology, 263*(1-4), 131-155. doi: 10.1016/s0022-1694(02)00056-2

BERGSTRÖM, S. (1975). THE DEVELOPMENT OF A SNOW ROUTINE FOR THE HBV-2 MODEL. *Hydrology Research, 6*(2), 73-92. doi: 10.2166/nh.1975.0006

Betson, R. P. (1964). What is watershed runoff? *Journal of Geophysical Research (1896-1977), 69*(8), 1541-1552. doi: 10.1029/JZ069i008p01541

Birkhead, A. L., & James, C. S. (2002). Muskingum river routing with dynamic bank storage. *Journal of Hydrology, 264*(1-4), 113-132. doi: 10.1016/s0022-1694(02)00068-9

Bornette, G., & Amoros, C. (1991). Aquatic vegetation and hydrology of a braided river floodplain. *Journal of Vegetation Science, 2*(4), 497-512. doi: 10.2307/3236032

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological), 26*(2), 211-243. doi: 10.1111/j.2517-6161.1964.tb00553.x|

Bozorg-Haddad, O., Abdi-Dehkordi, M., Hamedi, F., Pazoki, M., & Loaiciga, H. A. (2019). Generalized Storage Equations for Flood Routing with Nonlinear Muskingum Models. *Water Resources Management, 33*(8), 2677-2691. doi: 10.1007/s11269-019-02247-2

Bozorg-Haddad, O., Hamedi, F., Orouji, H., Pazoki, M., & Loaiciga, H. A. (2015). A Re-Parameterized and Improved Nonlinear Muskingum Model for Flood Routing. *Water Resources Management, 29*(9), 3419-3440. doi: 10.1007/s11269-015-1008-9

Cadle, S. H., Dasch, J. M., & Vande Kopple, R. (1987). Composition of snowmelt and runoff in northern Michigan. *Environmental science & technology, 21*(3), 295-299. doi: 10.1021/es00157a011

Carraro, F., Valiani, A., & Caleffi, V. (2018). Efficient analytical implementation of the DOT Riemann solver for the de Saint Venant-Exner morphodynamic model. *Advances in Water Resources, 113*, 189-201. doi: 10.1016/j.advwatres.2018.01.011

Cervantes, M. S., Hernando, A., Garcia-Abril, A., Valbuena, R., Saornil, J. V., & Manzanera, J. A. (2020). Simulation of overflow thresholds in urban basins: Case study in Tuxtla Gutierrez, Mexico. *River Research and Applications, 36*(7), 1307-1320. doi: 10.1002/rra.3642

Chang, C. M., & Yeh, H. D. (2018). Spectral analysis of temporal non-stationary rainfall-runoff processes. *Journal of Hydrology, 559*, 84-88. doi: 10.1016/j.jhydrol.2018.02.017

Choudhury, P., Shrivastava, R. K., & Narulkar, S. M. (2002). Flood routing in river networks using equivalent Muskingum inflow. *Journal of Hydrologic Engineering, 7*(6), 413-419. doi: 10.1061/(asce)1084-0699(2002)7:6(413)

Coe, M. T., Latrubesse, E. M., Ferreira, M. E., & Amsler, M. L. (2011). The effects of deforestation and climate variability on the streamflow of the Araguaia River, Brazil. *Biogeochemistry, 105*(1-3), 119-131. doi: 10.1007/s10533-011-9582-2

Croke, B. F. W., & Jakeman, A. J. (2004). A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environmental Modelling & Software, 19*(1), 1-5. doi: 10.1016/j.envsoft.2003.09.001

da Silva, R. M., Santos, C. A. G., & dos Santos, J. Y. G. (2018). Evaluation and modeling of runoff and sediment yield for different land covers under simulated rain in a semiarid region of Brazil. *International Journal of Sediment Research, 33*(2), 117-125. doi: 10.1016/j.ijsrc.2017.04.005

David, C. H., Famiglietti, J. S., Yang, Z. L., & Eijkhout, V. (2015). Enhanced fixed-size parallel speedup with the Muskingum method using a trans-boundary approach and a large subbasins approximation. *Water Resources Research, 51*(9), 7547-7571. doi: 10.1002/2014wr016650

David, C. H., Hobbs, J. M., Turmon, M. J., Emery, C. M., Reager, J. T., & Farniglietti, J. S. (2019). Analytical Propagation of Runoff Uncertainty Into Discharge Uncertainty Through a Large River Network. *Geophysical Research Letters, 46*(14), 8102-8113. doi: 10.1029/2019gl083342

David, J. S., Valente, F., & Gash, J. H. (2005). Evaporation of Intercepted Rainfall. In *Encyclopedia of Hydrological Sciences*.

den Besten, N., Steele-Dunne, S., de Jeu, R., & van der Zaag, P. (2021). Towards Monitoring Waterlogging with Remote Sensing for Sustainable Irrigated Agriculture. *Remote Sensing, 13*(15). doi: 10.3390/rs13152929

Ding, Y., & Wang, S. S. Y. (2005). Identification of Manning's roughness coefficients in channel network using adjoint analysis. *International Journal of Computational Fluid Dynamics, 19*(1), 3-13. doi: 10.1080/10618560410001710496

Dunne, T., & Black, R. D. (1970). An Experimental Investigation of Runoff Production in Permeable Soils. *Water Resources Research, 6*(2), 478-490. doi: 10.1029/WR006i002p00478

Faramarzi, M., Srinivasan, R., Iravani, M., Bladon, K. D., Abbaspour, K. C., Zehnder, A. J. B., & Goss, G. G. (2015). Setting up a hydrological model of Alberta: Data discrimination analyses prior to calibration. *Environmental Modelling & Software, 74*, 48-65. doi: 10.1016/j.envsoft.2015.09.006

Gao, Y. H., Lu, S. H., & Cheng, G. D. (2004). Simulation of rainfall-runoff and watershed convergence process in the upper reaches of Heihe River Basin, July 2002. *Science in China Series D-Earth Sciences, 47*, 1-

8. doi: 10.1360/04zd0001

Geng, X. J., Zhou, X. C., Yin, G. D., Hao, F. H., Zhang, X., Hao, Z. C., et al. (2020). Extended growing season reduced river runoff in Luanhe River basin. *Journal of Hydrology, 582*, 9. doi: 10.1016/j.jhydrol.2019.124538

Gentry, A. H., & Lopez-Parodi, J. (1980). Deforestation and increased flooding of the upper Amazon. *Science (New York, N.Y.), 210*(4476), 1354-1356. doi: 10.1126/science.210.4476.1354

Gill, M. A. (1978). Flood routing by the Muskingum method. *Journal of Hydrology, 36*(3), 353-363. doi: 10.1016/0022-1694(78)90153-1

Goenster-Jordan, S., Ingold, M., Jannoura, R., Buerkert, A., & Joergensen, R. G. (2021). Soil microbial properties of subalpine steppe soils at different grazing intensities in the Chinese Altai Mountains. *Scientific Reports, 11*(1), 8. doi: 10.1038/s41598-021-81120-y

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology, 377*(1-2), 80-91. doi: 10.1016/j.jhydrol.2009.08.003

Gupta, S. K., Tyagi, J., Sharma, G., Jethoo, A. S., & Singh, P. K. (2019). An Event-Based Sediment Yield and Runoff Modeling Using Soil Moisture Balance/Budgeting (SMB) Method. *Water Resources Management, 33*(11), 3721-3741. doi: 10.1007/s11269-019-02329-1

Guse, B., Pfannerstill, M., Gafurov, A., Fohrer, N., & Gupta, H. (2016). Demasking the integrated information of discharge: Advancing sensitivity analysis to consider different hydrological components and their rates of change. *Water Resources Research, 52*(11), 8724-8743. doi: 10.1002/2016wr018894

Hamedi, F., Bozorg-Haddad, O., Orouji, H., Fallah-Mehdipour, E., & Loaiciga, H. A. (2016). Nonlinear Muskingum Model for Flood Routing in Irrigation Canals Using Storage Moving Average. *Journal of Irrigation and Drainage Engineering, 142*(5). doi: 10.1061/(asce)ir.1943-4774.0001000

Hamilton, A., Hutchinson, D., & Moore, R. (2000). Estimating winter streamflow using conceptual streamflow model. *Journal of Cold Regions Engineering, 14*(4), 158-175. doi: 10.1061/(ASCE)0887-381X(2000)14:4(158)

Hassini, S., & Guo, Y. P. (2017). Derived flood frequency distributions considering individual event hydrograph shapes. *Journal of Hydrology, 547*, 296-308. doi: 10.1016/j.jhydrol.2017.02.003

He, R., Zhang, J., Bao, Z., Yan, X., Wang, G., & Liu, C. (2015). Response of runoff to climate change in the Haihe River basin. *Advances in Water Science, 26*(1), 1-9. doi: 10.14042/j.cnki.32.1309.2015.01.001

Hood, M. J., Clausen, J. C., & Warner, G. S. (2007). Comparison of stormwater lag times for low impact and traditional residential development. *Journal of the American Water Resources Association, 43*(4), 1036-1046. doi: 10.1111/j.1752-1688.2007.00085.x

Huang, C. C., Wang, G. Q., Zheng, X. G., Yu, J. S., & Xu, X. Y. (2015a). Simple Linear Modeling Approach for Linking Hydrological Model Parameters to the Physical Features of a River Basin. *Water Resources Management, 29*(9), 3265-3289. doi: 10.1007/s11269-015-0996-9

Huang, J. B., Wen, J. W., Wang, B., & Zhu, S. J. (2015b). Numerical analysis of the combined rainfall-runoff process and snowmelt for the Alun River Basin, Heilongjiang, China. *Environmental Earth Sciences, 74*(9), 6929-6941. doi: 10.1007/s12665-015-4694-y

Jakeman, A. J., & Hornberger, G. M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resources Research, 29*(8), 2637-2649. doi: 10.1029/93WR00877

Knighton, J., White, E., Lennon, E., & Rajan, R. (2014). Development of probability distributions for urban hydrologic model parameters and a Monte Carlo analysis of model sensitivity. *Hydrological Processes, 28*(19), 5131-5139. doi: 10.1002/hyp.10009

Knoben, W. J., Freer, J. E., & Woods, R. A. (2019). Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences, 23*(10), 4323-4331. doi: 10.5194/hess-23-4323-2019

Kollat, J., Reed, P., & Wagener, T. (2012). When are multiobjective calibration trade‐offs in hydrologic models meaningful? *Water Resources Research, 48*(3). doi: 10.1029/2011WR011534

Krasnostein, A. L., & Oldham, C. E. (2004). Predicting wetland water storage. *Water Resources Research, 40*(10). doi: 10.1029/2003wr002899

Kundzewicz, Z. W., Hirabayashi, Y., & Kanae, S. (2010). River floods in the changing climate—observations and projections. *Water Resources Management, 24*(11), 2633-2646. doi: 10.1007/s11269-009-9571-6

Li, J., Liu, C. M., Wang, Z. G., & Liang, K. (2015). Two universal runoff yield models: SCS vs. LCM. *Journal of Geographical Sciences, 25*(3), 311-318. doi: 10.1007/s11442-015-1170-2

Li, M. Y., Liu, T. X., Duan, L. M., Luo, Y. Y., Ma, L., Wang, Y. X., et al. (2020). Scale transfer and simulation of the infiltration in chestnut soil in a semi-arid grassland basin. *Ecological Engineering, 158*. doi: 10.1016/j.ecoleng.2020.106045

Li, M. Y., Liu, T. X., Duan, L. M., Ma, L., Wang, Y. X., Zhou, Y. J., et al. (2021). Hydrologic gradient changes of soil respiration in typical steppes of Eurasia. *Science of the Total Environment, 794*. doi: 10.1016/j.scitotenv.2021.148684

Liang, S., Cheng, C., Jia, K., Jiang, B., Liu, Q., Xiao, Z., et al. (2020). The Global LAnd Surface Satellite (GLASS) products suite. *Bulletin of the American Meteorological Society*. doi: 10.1175/BAMS-D-18-0341.1

Liang, S., Zhang, X., Xiao, Z., Cheng, J., Liu, Q., & Zhao, X. (2013a). Global LAnd Surface Satellite (GLASS) products: Algorithms, validation and analysis. *Springer*. doi:

Liang, S. L., Zhao, X., Liu, S. H., Yuan, W. P., Cheng, X., Xiao, Z. Q., et al. (2013b). A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies. *International Journal of Digital Earth, 6*, 5-33. doi: 10.1080/17538947.2013.805262

Liang, Z. M., Wang, J., Li, B. Q., & Yu, Z. B. (2012). A statistically based runoff-yield model coupling infiltration excess and saturation excess mechanisms. *Hydrological Processes, 26*(19), 2856-2865. doi: 10.1002/hyp.8357

Lopes, V. L., & Canfield, H. E. (2004). Effects of watershed representation on runoff and sediment yield modeling. *Journal of the American Water Resources Association, 40*(2), 311-319. doi: 10.1111/j.1752-1688.2004.tb01031.x

Lou, H. Z., Yang, S. T., Hao, F. H., Jiang, L. M., Zhao, C. S., Ren, X. Y., et al. (2019). SMAP, RS-DTVGM, and in-situ monitoring: Which performs best in presenting the soil moisture in the middle-high latitude frozen area in the Sanjiang Plain, China? *Journal of Hydrology, 571*, 300-310. doi: 10.1016/j.jhydrol.2018.12.023

Ma, F., Ye, A. Z., & Duan, Q. Y. (2019). Seasonal drought ensemble predictions based on multiple climate models in the upper Han River Basin, China. *Climate Dynamics, 53*(12), 7447-7460. doi: 10.1007/s00382-017-3577-1

Maniquiz, M. C., Kim, L. H., Lee, S., & Choi, J. (2012). Flow and mass balance analysis of eco-bio infiltration system. *Frontiers of Environmental Science & Engineering, 6*(5), 612-619. doi: 10.1007/s11783-012-0448-1

Melsen, L. A., & Guse, B. (2019). Hydrological Drought Simulations: How Climate and Model Structure Control Parameter Sensitivity. *Water Resources Research, 55*(12), 10527-10547. doi: 10.1029/2019wr025230

Metivier, F., Devauchelle, O., Chauvet, H., Lajeunesse, E., Meunier, P., Blanckaert, K., et al. (2016). Geometry of meandering and braided gravel-bed threads from the Bayanbulak Grassland, Tianshan, P. R. China. *Earth*

Surface Dynamics, 4(1), 273-283. doi: 10.5194/esurf-4-273-2016

Minville, M., Brissette, F., & Leconte, R. (2010). Impacts and uncertainty of climate change on water resource management of the Peribonka River System (Canada). *Journal of Water Resources Planning and Management, 136*(3), 376-385. doi: 10.1061/(ASCE)WR.1943-5452.0000041

Misirli, F., Gupta, H. V., Sorooshian, S., & Thiemann, M. (2003). Bayesian recursive estimation of parameter and output uncertainty for watershed models. *Calibration of Watershed Models, Water Sci. Appl. Ser, 6*, 113-124. doi: 10.1029/WS006p0113

Moore, I. D., & Grayson, R. B. (1991). Terrain-based catchment partitioning and runoff prediction using vector elevation data. *Water Resources Research, 27*(6), 1177-1191. doi: 10.1029/91WR00090

Moore, R. (2007). The PDM rainfall-runoff model. *Hydrology and Earth System Sciences, 11*(1), 483-499. doi: 10.5194/hess-11-483-2007

Muzik, I. (1992). Derivation of unit and flood hydrographs using a gis. *Environmental monitoring and assessment, 23*(1-3), 45-56. doi: 10.1007/bf00406951

Nash, J. E., & Sutcliffe, J. E. (1970). River flow forecasting through conceptual models. Part 1 - A discussion of principles. *Journal of Hydrology, 10*, 282-290. doi: 10.1016/0022-1694(70)90255-6

Nourani, V., Baghanam, A. H., Adamowski, J., & Kisi, O. (2014). Applications of hybrid wavelet-Artificial Intelligence models in hydrology: A review (vol 514, pg 358, 2014). *Journal of Hydrology, 517*, 1189-1189. doi: 10.1016/j.jhydrol.2014.06.024

Osborn, H. B., & Lane, L. (1969). Precipitation-runoff relations for very small semiarid rangeland watersheds. *Water Resources Research, 5*(2), 419-425. doi: 10.1029/WR005i002p00419

Pfannerstill, M., Guse, B., Reusser, D., & Fohrer, N. (2015). Process verification of a hydrological model using a temporal parameter sensitivity analysis. *Hydrology and Earth System Sciences, 19*(10), 4365-4376. doi: 10.5194/hess-19-4365-2015

Poiani, K. A., & Johnson, W. C. (1993). A Spatial Simulation Model of Hydrology and Vegetation Dynamics in Semi-Permanent Prairie Wetlands. *Ecological applications : a publication of the Ecological Society of America, 3*(2), 279-293. doi: 10.2307/1941831

Qi, W., Zhang, C., Fu, G. T., Sweetapple, C., & Liu, Y. L. (2019). Impact of robustness of hydrological model parameters on flood prediction uncertainty. *Journal of Flood Risk Management, 12*. doi: 10.1111/jfr3.12488

Qiu, G. Y., Shi, P. J., & Wang, L. M. (2006). Theoretical analysis of a remotely measurable soil evaporation transfer coefficient. *Remote Sensing of Environment, 101*(3), 390-398. doi: 10.1016/j.rse.2006.01.007

Reaney, S. M., Bracken, L. J., & Kirkby, M. J. (2014). The importance of surface controls on overland flow connectivity in semi-arid environments: results from a numerical experimental approach. *Hydrological Processes, 28*(4), 2116-2128. doi: 10.1002/hyp.9769

Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G., & Franks, S. W. (2011). Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation. *Water Resources Research, 47*. doi: 10.1029/2011wr010643

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society, 85*(3), 381-394. doi: 10.1175/BAMS-85-3-381

Sang, Y. F. (2013). A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric Research, 122*, 8-15. doi: 10.1016/j.atmosres.2012.11.003

Schoups, G., & Nasseri, M. (2021). GRACEfully Closing the Water Balance: A Data-Driven Probabilistic Approach Applied to River Basins in Iran. *Water Resources Research, 57*(6). doi:

10.1029/2020wr029071

Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences, 4*(2), 215-224. doi: 10.5194/hess-4-215-2000

Sikorska, A. E., Montanari, A., & Koutsoyiannis, D. (2015). Estimating the Uncertainty of Hydrological Predictions through Data-Driven Resampling Techniques. *Journal of Hydrologic Engineering, 20*(1). doi: 10.1061/(asce)he.1943-5584.0000926

Song, C. L., Wang, G. X., Mao, T. X., Dai, J. C., & Yang, D. Q. (2020). Linkage between permafrost distribution and river runoff changes across the Arctic and the Tibetan Plateau. *Science China-Earth Sciences, 63*(2), 292-302. doi: 10.1007/s11430-018-9383-6

Song, X. M., Zhan, C. S., & Xia, J. (2012). Integration of a statistical emulator approach with the SCE-UA method for parameter optimization of a hydrological model. *Chinese Science Bulletin, 57*(26), 3397-3403. doi: 10.1007/s11434-012-5305-x

Staudt, F., Mullarney, J. C., Pilditch, C. A., & Huhn, K. (2019). Effects of grain-size distribution and shape on sediment bed stability, near-bed flow and bed microstructure. *Earth Surface Processes and Landforms, 44*(5), 1100-1116. doi: 10.1002/esp.4559

Strelkoff, T. (1970). Numerical Solution of Saint-Venant Equations. *Journal of the Hydraulics division, 96*(1), 223-252. doi: doi:10.1061/JYCEAJ.0002262

Sun, R. C., Hernandez, F., Liang, X., & Yuan, H. L. (2020). A Calibration Framework for High-Resolution Hydrological Models Using a Multiresolution and Heterogeneous Strategy. *Water Resources Research, 56*(8), 26. doi: 10.1029/2019wr026541

Svoray, T., Assouline, S., & Katul, G. (2015). Introduction to a special section on ecohydrology of semiarid environments: Confronting mathematical models with ecosystem complexity. *Water Resources Research, 51*(11), 8677-8683. doi: 10.1002/2015wr018131

Tanaka, Y., Matsukura, Y., Batnasan, N., & Tuvshinjargal, D. (2005). Distinct runoff processes in granite and sandstone drainage basins near Ulaanbaatar, Mongolia. *Environmental Geology, 47*(5), 640-646. doi: 10.1007/s00254-004-1189-7

Tang, Y., Leon, A. S., & Kavvas, M. L. (2020). Impact of Size and Location of Wetlands on Watershed-Scale Flood Control. *Water Resources Management, 34*(5), 1693-1707. doi: 10.1007/s11269-020-02518-3

Tang, Y., Reed, P. M., & Kollat, J. B. (2007). Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications. *Advances in Water Resources, 30*(3), 335-353. doi: 10.1016/j.advwatres.2006.06.006

Tong, R., Parajka, J., Salentinig, A., Pfeil, I., Komma, J., Szeles, B., et al. (2021). The value of ASCAT soil moisture and MODIS snow cover data for calibrating a conceptual hydrologic model. *Hydrology and Earth System Sciences, 25*(3), 1389-1410. doi: 10.5194/hess-25-1389-2021

Tung, Y. K. (1985). River Flood Routing by Nonlinear Muskingum Method. *Journal of Hydraulic Engineering, 111*(12), 1447-1460. doi: doi:10.1061/(ASCE)0733-9429(1985)111:12(1447)

Vassova, D. (2013). Comparison of Rainfall-Runoff Models for Design Discharge Assessment in a Small Ungauged Catchment. *Soil and Water Research, 8*(1), 26-33. doi: 10.17221/36/2012-swr

Wagener, T., & Montanari, A. (2011). Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resources Research, 47*. doi: 10.1029/2010wr009469

Wagener, T., Wheater, H., & Gupta, H. V. (2004). *Rainfall-runoff modelling in gauged and ungauged catchments*: World Scientific.

Wang, G. T., Chen, S., & Boll, J. (2003). A semianalytical solution of the Saint-Venant equations for channel flood routing. *Water Resources Research, 39*(4). doi: 10.1029/2002wr001690

Wang, R., Li, R., Li, J., & Hu, C. M. (2014). A hydraulics-based analytical method for artificial water replenishment in wetlands by reservoir operation. *Ecological Engineering, 62*, 71-76. doi: 10.1016/j.ecoleng.2013.10.026

Wendi, D., Merz, B., & Marwan, N. (2019). Assessing Hydrograph Similarity and Rare Runoff Dynamics by Cross Recurrence Plots. *Water Resources Research, 55*(6), 4704-4726. doi: 10.1029/2018wr024111

Wu, Y. Y., Huang, L., Zhao, C. W., Chen, M. H., & Ouyang, W. (2021a). Integrating hydrological, landscape ecological, and economic assessment during hydropower exploitation in the upper Yangtze River. *Science of the Total Environment, 767*, 12. doi: 10.1016/j.scitotenv.2021.145496

Wu, Z. J., Zhang, J. H., Deng, F., Zhang, S., Zhang, D., Xun, L., et al. (2021b). Fusion of GF and MODIS Data for Regional-Scale Grassland Community Classification with EVI2 Time-Series and Phenological Features. *Remote Sensing, 13*(5), 19. doi: 10.3390/rs13050835

Xiong, L. H., & Guo, S. L. (2004). Effects of the catchment runoff coefficient on the performance of TOPMODEL in rainfall-runoff modelling. *Hydrological Processes, 18*(10), 1823-1836. doi: 10.1002/hyp.1449

Xu, H., Ye, M., & Li, J. (2009). The ecological characteristics of the riparian vegetation affected by river overflowing disturbance in the lower Tarim River. *Environmental Geology, 58*(8), 1749-1755. doi: 10.1007/s00254-008-1674-5

Yamanaka, T., & Ma, W. C. (2017). Runoff prediction in a poorly gauged basin using isotope-calibrated models. *Journal of Hydrology, 544*, 567-574. doi: 10.1016/j.jhydrol.2016.12.005

Yan, H. X., Sun, N., Fullerton, A., & Baerwalde, M. (2021). Greater vulnerability of snowmelt-fed river thermal regimes to a warming climate. *Environmental Research Letters, 16*(5), 13. doi: 10.1088/1748-9326/abf393

Yang, H., Piao, S. L., Zeng, Z. Z., Ciais, P., Yin, Y., Friedlingstein, P., et al. (2015). Multicriteria evaluation of discharge simulation in Dynamic Global Vegetation Models. *Journal of Geophysical Research-Atmospheres, 120*(15), 7488-7505. doi: 10.1002/2015jd023129

Yang, K., He, J., Tang, W. J., Qin, J., & Cheng, C. C. K. (2010). On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau. *Agricultural and Forest Meteorology, 150*(1), 38-46. doi: 10.1016/j.agrformet.2009.08.004

Yin, S. Q., Zhu, Z. Y., Wang, L., Liu, B. Y., Xie, Y., Wang, G. N., & Li, Y. S. (2018). Regional soil erosion assessment based on a sample survey and geostatistics. *Hydrology and Earth System Sciences, 22*(3), 1695-1712. doi: 10.5194/hess-22-1695-2018

Yin, X. W., Feng, Q., Zheng, X. J., Zhu, M., Wu, X., Guo, Y., et al. (2021). Spatio-temporal dynamics and eco-hydrological controls of water and salt migration within and among different land uses in an oasis-desert system. *Science of the Total Environment, 772*, 12. doi: 10.1016/j.scitotenv.2021.145572

Yokoo, Y., & Kazama, S. (2012). Numerical investigations on the relationships between watershed characteristics and water balance model parameters: searching for universal relationships among regional relationships. *Hydrological Processes, 26*(6), 843-854. doi: 10.1002/hyp.8299

Young, C. C., & Liu, W. C. (2015). Prediction and modelling of rainfall-runoff during typhoon events using a physically-based and artificial neural network hybrid model. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 60*(12), 2102-2116. doi: 10.1080/02626667.2014.959446

Zha, X. N., Xiong, L. H., Guo, S. L., Kim, J. S., & Liu, D. D. (2020). AR-GARCH with Exogenous Variables as a Postprocessing Model for Improving Streamflow Forecasts. *Journal of Hydrologic Engineering, 25*(8), 16. doi: 10.1061/(asce)he.1943-5584.0001955

Zhang, J. P., Zhang, H., Xiao, H. L., Fang, H. Y., Han, Y. P., & Yu, L. (2021). Effects of rainfall and runoff-yield conditions on runoff. *Ain Shams Engineering Journal, 12*(2), 2111-2116. doi: 10.1016/j.asej.2020.10.010

Zhang, L., & Singh, V. P. (2014). Joint and Conditional Probability Distributions of Runoff Depth and Peak Discharge Using Entropy Theory. *Journal of Hydrologic Engineering, 19*(6), 1150-1159. doi: 10.1061/(asce)he.1943-5584.0000906

Zoccatelli, D., Marra, F., Armon, M., Rinat, Y., Smith, J. A., & Morin, E. (2019). Contrasting rainfall-runoff characteristics of floods in desert and Mediterranean basins. *Hydrology and Earth System Sciences, 23*(6), 2665-2678. doi: 10.5194/hess-23-2665-2019

Figures

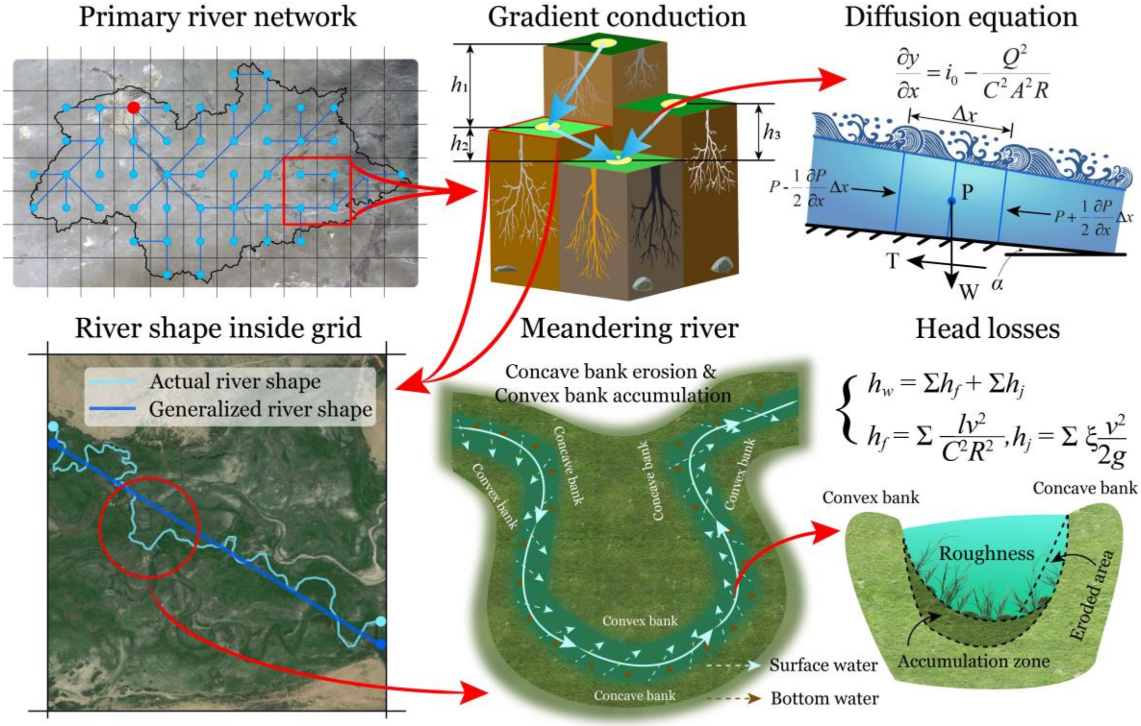


Figure 1. Schematic diagram of natural processes such as the flow convergence, actual river lengths, and channel turns of grassland rivers. Note: The river network shown in the figure does not correspond to the real modelled river network resolution.

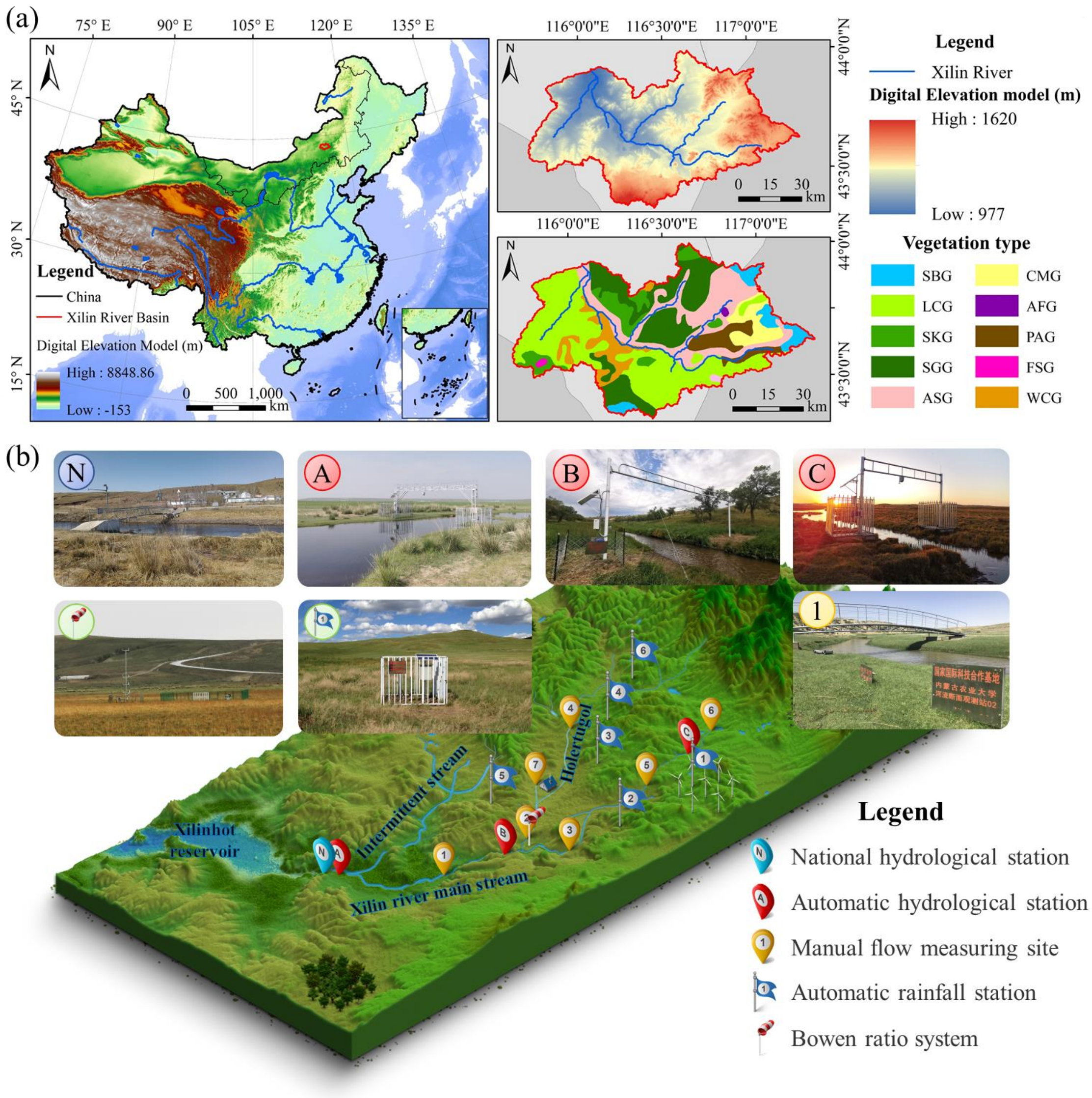


Figure 2. Location, vegetation types (a), topography and stations (b) in the XRB. SBG: *S. baicalensis* Roshev. grassland; LCG: *L. chinensis* (Trin.) Tzvel. grassland; SKG: *S. krylovii* Roshev. grassland; SGG: *S. grandis* P.A. Smirn. grassland; ASG: *A. splendens* (Trin.) Nevski grassland; CMG: *C. microphylla* Lam grassland; AFG: *Artemisia frigida* Willd. grassland; PAG: *P. asperata* Mast. grassland; FSG: *Filifolium sibiricum* (L.) Kitam. grassland; and WCG: weed community grassland.

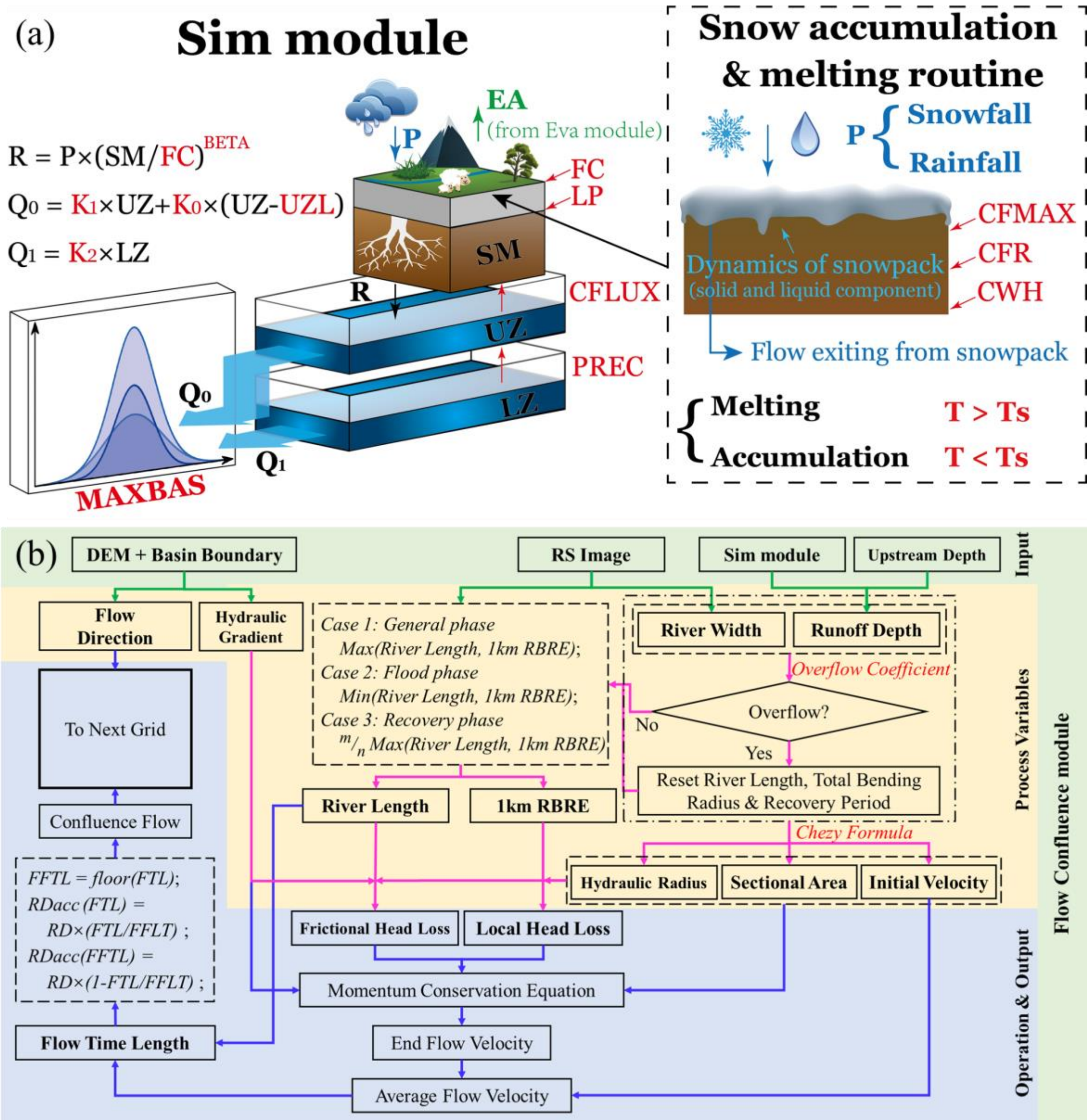


Figure 3. (a) Schematic diagram of the MYEH model simulation (Sim) module; (b) schematic diagram of the MYEH model flow confluence (FLC) module. The full names of the variables shown in Figure 3a can be seen in Table 2. DEM: digital elevation model; RS: remote sensing; 1-km RBRE: 1-km river bend radius equivalent; FTL: flow time length; FFTL: fixed flow time length; and RDacc: accumulated runoff depth.

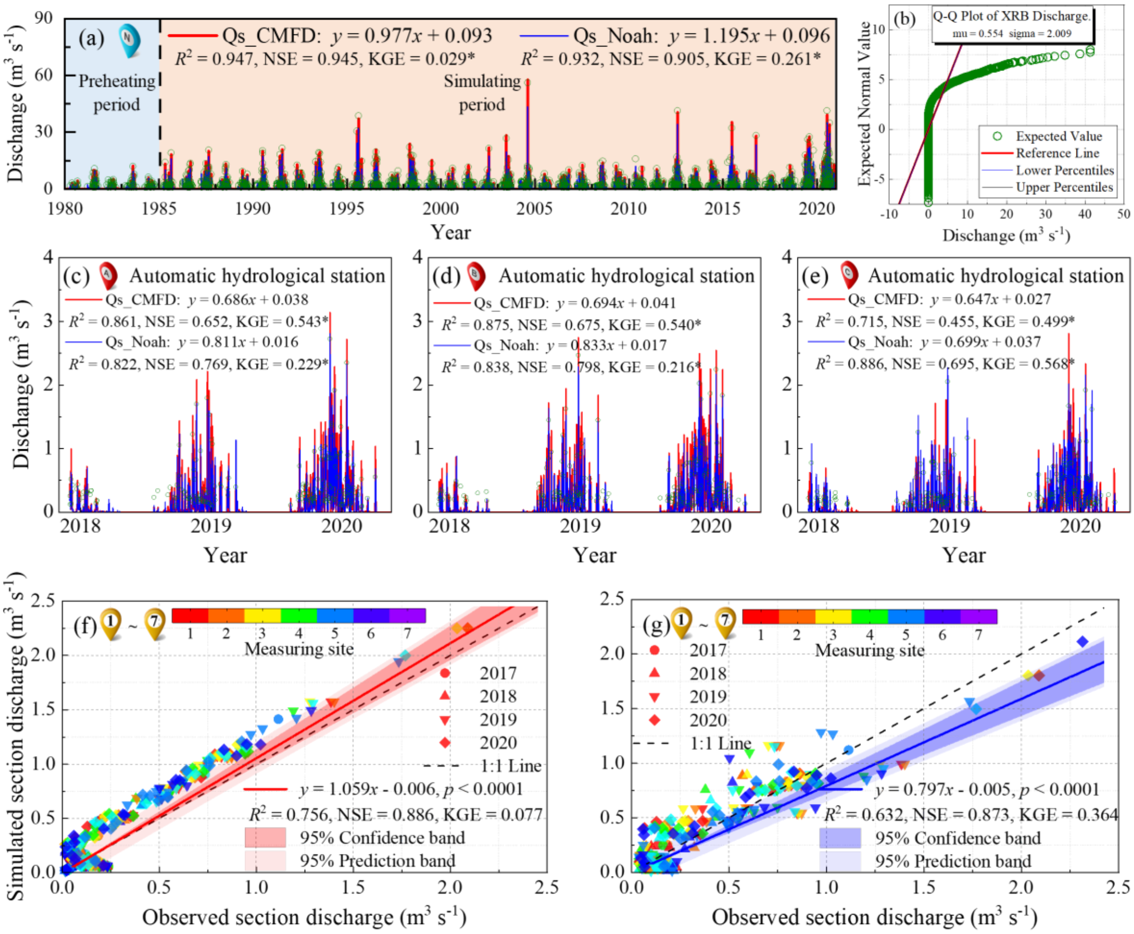
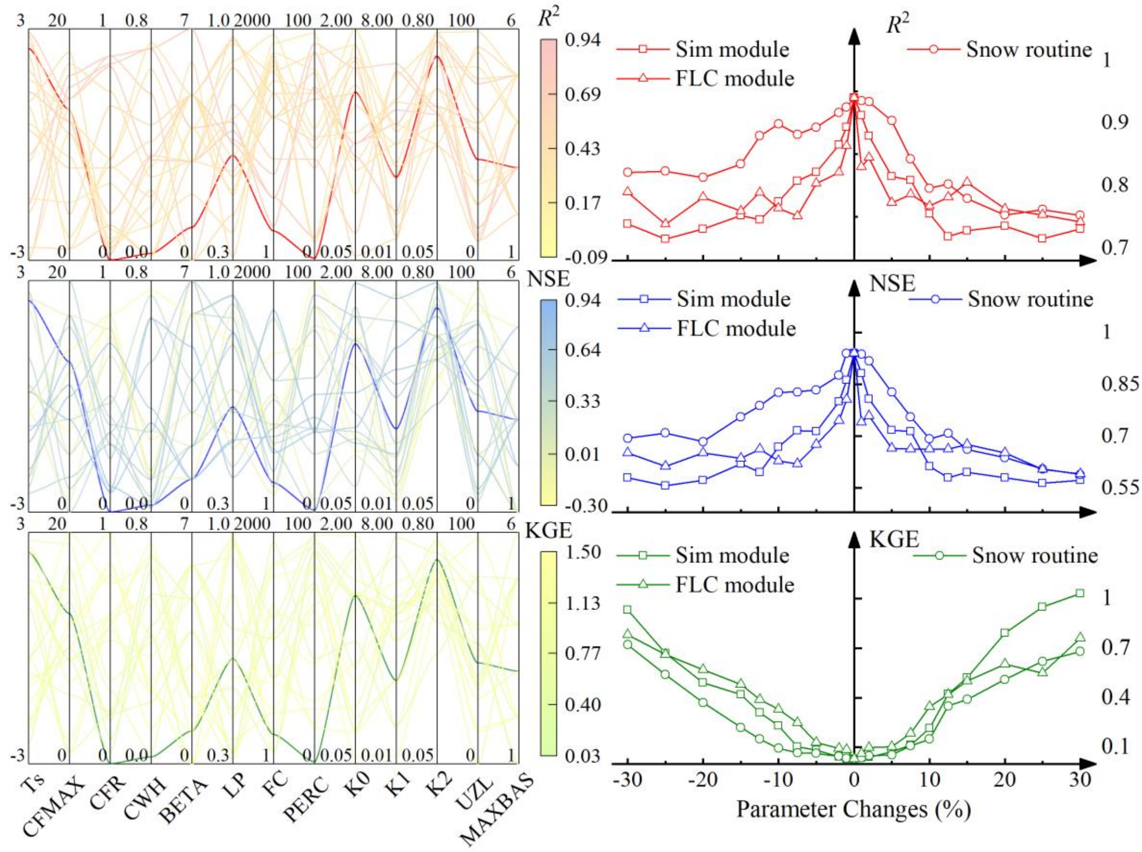


Figure 4. (a) Comparison of the simulated and measured daily sectional discharge at a national hydrological station obtained using two meteorologically driven datasets with the MYEH model; (b) Q-Q plot of daily discharge in the XRB; (c-e) comparison of the simulated and measured daily sectional discharge at three automatic hydrological stations using two meteorologically driven datasets with the MYEH model; (f) comparison of the simulated and measured daily sectional discharge at seven measuring sites using the CMFD; and (g) comparison of the simulated and measured daily sectional discharge at seven measuring sites using the GLDAS-NOAH dataset. The green points in (a) to (e) are the values observed at the national hydrological station.



1159

1160 **Figure 5.** Parameter optimization (a-c) and parameter sensitivity analysis (d-f) results obtained
1161 for the MYEH model.

1162

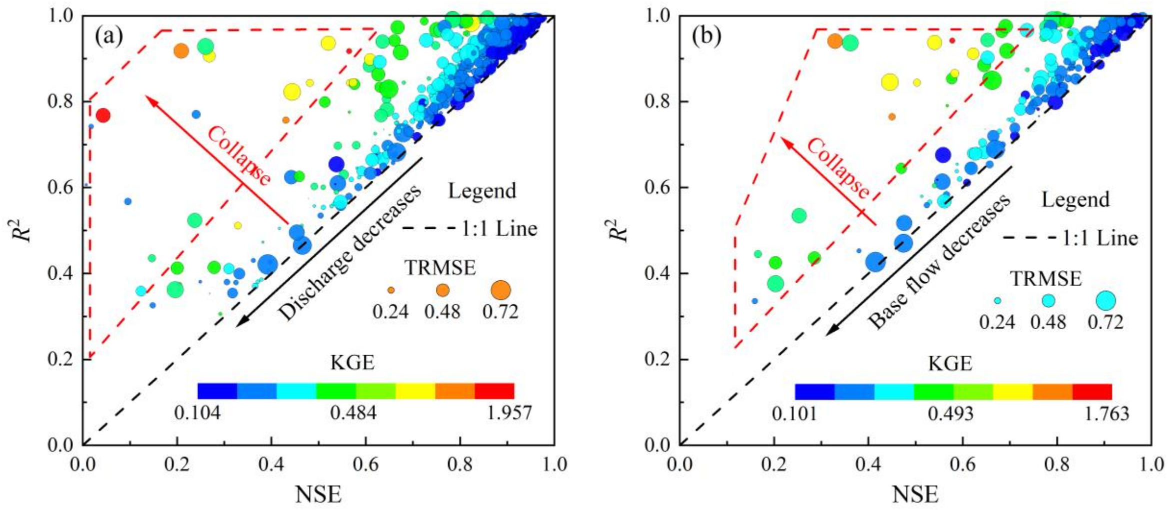


Figure 6. Comparison of the MYEH model-simulated runoff discharge during the nonfreezing period (a) and base flow during the freezing period (b) in the XRB as determined using the CMFD and GLDAS-NOAH data sources. In this figure, NSE and R^2 are plotted on the *X* and *Y* axes, respectively, KGE is plotted in color, and TRMSE is plotted using the size of the markers. The black arrow points in the direction of decreasing flow or base flow. The red arrow indicates the tendency of both data-source simulations to collapse. NSE: Nash-Sutcliffe efficiency; KGE: Kling-Gupta efficiency; and TRMSE: Box-Cox transformed root mean square error.

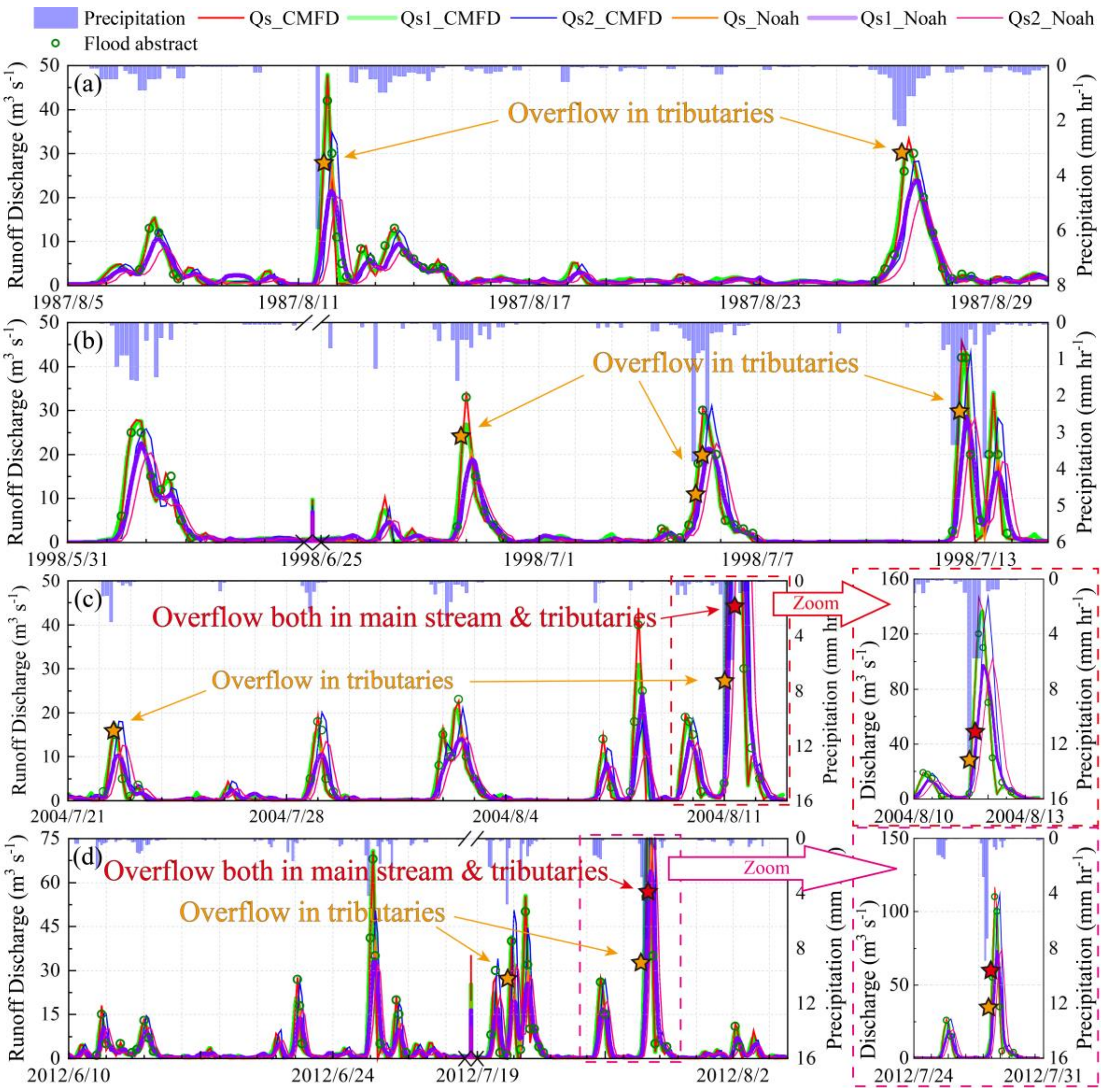


Figure 7. Simulations of the 3-hour flood process under three confluence modes using the CMFD and GLDAS-NOAH data sources. Figures 7(a) to 7(d) show monsoon floods in 1987, 1998, 2004 and 2012, respectively. Qs indicates the MYEH model confluence mode (FLC). Qs1 indicates the confluence mode in which the actual river length, river bending and overflow are not considered. Qs2 indicates the confluence mode in which the actual river length and river bending are considered but overflow is not considered. The orange and red stars represent the overflow of tributaries and the overflow of main streams and tributaries in a flood event, respectively.

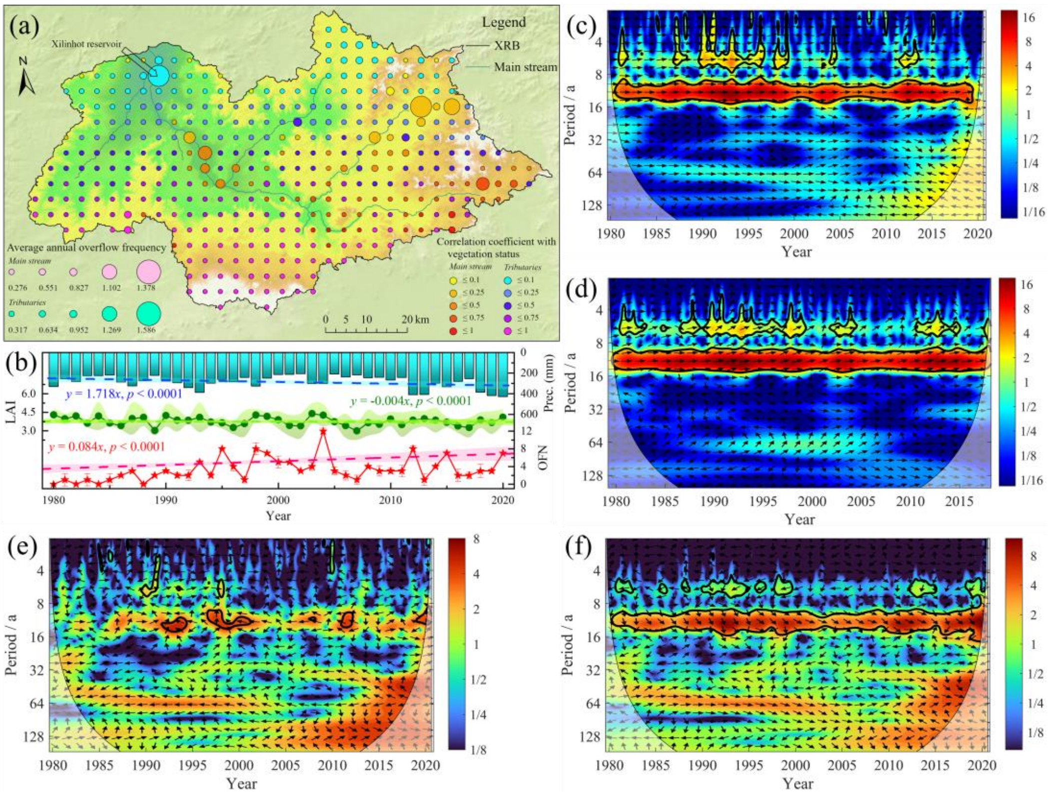


Figure 8. (a) Average annual overflow frequency and the correlation between overflows and vegetation status. (b) Trend analyses of precipitation, LAI, and OFN from 1980 to 2020. (c-f) The cross-wavelet energy spectrum analyses of the OFN with precipitation, LAI, SOI and NINO3.4 SST. The 5% significance level against red noise is shown as a thick contour line. The relative phase relationships are shown as arrows (with in-phase relationships pointing right and anti-phase relationships pointing left).

Tables

Table 1. Information of measurement stations in the XRB.

Name	Collector	Monitoring indicator	Frequency	Data length
National hydrological station	/	Flow discharge	1 day	1964/1/1-2020/12/31
Automatic hydrological station	RQ-30 radar sensor (Sommer GmbH, Austria)	Water level, flow velocity, flow discharge	1 min	2018/8/15-2020/12/31
Bowen ratio system	CR1000 (Campbell Scientific Inc., Logan, UT, USA)	Air temperature*, humidity*, wind speed*, wind direction*, precipitation, total radiation, soil heat flux, etc	1 min	2017/6/15-2020/12/31
Automatic rainfall station	RG600 tilting rain gauge (Global water, USA)	Precipitation	1 min	2016/6/30-2020/12/31
Manual flow measuring site	LS1206B propeller type flow sensor (Nanjing Nanshui Water Technology Company, PRC)	Flow velocity	7 days	Apr. to Oct. from 2017 to 2020

Note: * represents the monitoring indicator is located at a height of 2,3.5, 5,10 meters.

Table 2. Summary of parameters used in the Sim module and FLC module within the MYEH model.

Module	Parameter (Units)	Full name	Range	Module	Parameter (Units)	Full name
	Ts (°C)	Threshold temperature	-3 to 3		A_G (km^2)	Grid area
	CFMAX (mm °C^{-1})	Degree-3-hour factor	0 to 20		W_R (m)	River width
	CFR (-)	Refreezing factor	0 to 1		L_R (km)	River length
	CWH (-)	Water holding capacity of snow	0 to 0.8		H_R (m)	Runoff height
	BETA (-)	Exponential parameter in soil routine	0 to 7		dH (km)	Elevation difference
	LP (-)	Evapotranspiration limit	0.3 to 1		FD (-)	Flow direction
Sim	FC (mm)	Field capacity	1 to 2000	FLC	1km RBRE (degree)	1km river bending radius equivalent
	PERC (mm dt^{-1})	maximum flux from Upper to Lower Zone	0 to 100		v_1 (m dt^{-1})	Initial velocity
	K0 (dt^{-1})	Near surface flow coefficient (ratio)	0.05 to 2		v_2 (m dt^{-1})	End flow velocity
	K1 (dt^{-1})	Upper Zone outflow coefficient (ratio)	0.01 to 8		A_S (m^2)	Sectional area
	K2 (dt^{-1})	Lower Zone outflow coefficient (ratio)	0.05 to 0.8		FTL (dt)	Flow time length
	UZL (mm)	Near surface flow threshold	0 to 100		FFTL (dt)	Fixed flow time length
	MAXBAS (dt)	Flow routing coefficient	1 to 6		RD_{Acc} (mm)	Accumulated runoff depth

Note: In this table, dt represents the unit time.

Table 3. Characteristics of the two meteorological datasets.

Dataset	Version	Date used in study	Temporal resolution
CMFD	01.05.0016	1980.01.01-2018.12.31	3 hours
GLDAS-Noah	V2.0	1980.01.01-2000.12.31	3 hours
GLDAS-Noah	v2.1	2000.01.01-2020.12.31	3 hours

Note: CMFD: China meteorological forcing dataset, in which the temperature, pressure, specific humidity, wind speed, downward shortwave radiation, downward longwave radiation, and precipitation rate data are used in the study. NASA Global Land Data Assimilation System Version 2 (GLDAS-2) has three components: GLDAS-2.0, GLDAS-2.1, and GLDAS-2.2. GLDAS-2.0 is forced entirely with the Princeton meteorological forcing input data and provides a temporally consistent series from 1948 through 2014. GLDAS-2.1 is forced with a combination of model and observation data from 2000 to present.

Table 4. Six evaluation value of simulated runoff in XRB using two data sources.

	R^2	NSE	KGE	RMSE	BIAS	MAE
CMFD	0.947**	0.946	0.029	0.463	0.003	0.147
GLDAS-Noah	0.932**	0.905	0.262	0.616	0.096	0.191

Note: ** indicates that the increasing or decreasing trend is significant at $\alpha \leq 0.001$.