Climate warming-driven changes in the cryosphere and their impact on groundwater–surface water interactions in the Heihe River Basin

Amanda Triplett[1], Laura E. Condon[1]

[1]Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, 85705, The United States of America

5   *Correspondence to*: Amanda Triplett (aktriplett@arizona.edu)

Abstract. The Heihe River Basin in Northwest China depends heavily on both manmade and natural storage (e.g., surface reservoirs, rivers, and groundwater) to support economic and environmental functions. The Qilian Mountain cryosphere in the upper basin is integral to recharging these storage supplies. It is well established that climate warming is driving major shifts in high elevation water storage through loss of glaciers and permafrost. However, the impacts on groundwater–surface water

10   interactions and water supply in corresponding lower reaches are less clear. We built an integrated hydrologic model of the middle basin, where most water usage occurs, in order to explore the hydrologic response to the changing cryosphere. We simulate watershed response to loss of glaciers (*Glacier* scenario), advanced permafrost degradation (*Permafrost* scenario), both responses (*Combined* scenario) and projected temperature increases in the middle basin (*Warming* scenario) by altering streamflow inputs to the model to represent cryosphere melting processes, as well as by increasing the temperature of the

15   climate forcing data. Net losses to groundwater storage in the *Glacier* and net gains in *Permafrost* and *Combined* scenarios show the potential of groundwater exchanges to mediate streamflow shifts. The result of the *Combined* scenario also shows that permafrost degradation has more of an impact on the system than glacial loss. Seasonal differences in groundwater–surface water partitioning are also evident. The *Glacier* scenario has the highest fraction of groundwater in streamflow in early spring. The *Permafrost* and *Combined* scenarios meanwhile have the highest fraction of streamflow infiltration in late spring and

20   summer. The *Warming* scenario raises the temperature of the *Combined* scenario by 2 °C. This results in net groundwater storage loss, a reversal from the *Combined* scenario. Large seasonal changes in evapotranspiration and stream network connectivity relative to *Combined* show the potential for warming to overpower changes resulting from streamflow. Our results demonstrate the importance of understanding the entire system of groundwater–surface water exchanges to assess water resources under changing climatic conditions. Ultimately, this analysis can be used to examine the cascading impact of climate

25   change in the cryosphere on the resilience of water resources in arid basins downstream of mountain ranges globally.

1 Introduction

Mountains are an important source of freshwater for arid regions around the world (Qin et al., 2013; Viviroli and Weingartner, 2004; Wu et al., 2015). The cryosphere (i.e., water in mountainous, alpine regions stored as glaciers, snow, permafrost, and rain) plays a critical role in moderating water availability to downstream watersheds (Gao et al., 2018). It temporally

30 redistributes winter precipitation to higher demand periods like the spring and summer (Viviroli et al., 2011) and reduces the variability of flow (Wang and Cheng, 2000).

High latitude, cold regions have greater sensitivity to global warming (Chen et al., 2018; Jones and Rinehart, 2010; Zhang et al., 2020). The warming rate in the Tibetan Plateau, the largest and highest mountain-range in the world, is twice the global rate (You et al., 2020). This accelerated warming of the cryosphere has substantially altered water cycles and streamflow

35 (Chen et al., 2018; Wu et al., 2015; Xu et al., 2015; Zhang et al., 2016). Alterations in the quantity of cryosphere water storage and timing of discharge can change downstream water availability and how it is allocated (Chen et al., 2018; Xu et al., 2015). However, the impact of cryosphere melting on downstream systems is not fully understood.

The Heihe River Basin is an example of a system that has been impacted by the warming climate. It is a semi-arid, agriculturally important region located in Northwest China (Fig. 1). The Qilian Mountain cryosphere in the upper basin is the

40 region's primary water source (Wang and Cheng, 2000; Li et al., 2016). The movement of water from the high precipitation upper reaches to the arid valley floor has been critical for downstream development (Liu, Y. et al., 2019). It has allowed for the expansion of irrigated agriculture which accounts for over 90 % of water usage in the middle basin (Chen et al., 2005; Deng and Zhao, 2015; Sun et al., 2016). However, this reliance on water from the upper reaches makes the middle basin more vulnerable to warming induced changes in the cryosphere than other areas with higher local precipitation (Kang et al., 1999).

45 The upper basin is expected to undergo significant changes in glacier volume, permafrost coverage and precipitation due to climate change. Future projections for northern Asia, where the Heihe River Basin is located, indicate precipitation will likely increase (Shi et al., 2006; Zhang et al., 2016). However, estimates for the timing and volume of future precipitation in high mountain areas are variable (IPCC, 2014). Increasing warming trends on the other hand are essentially certain (IPCC, 2014). Thus, in this study we focus on processes resulting from increased temperature alone, such as glacial melt and

50 permafrost degradation.

Glacial contribution to streamflow is of particularly high importance in arid basins (Viviroli et al., 2011). Glaciers can stabilize flows, especially during hot or dry years (Chen et al., 2015; Qin et al., 2014). The ability of glaciers to buffer streamflow depends on glacial volume, melt rate, and the balance with evapotranspiration (ET). Under climate warming, it is estimated that glaciers in the upper basin may disappear entirely by the middle of the 21st century (Chen et al., 2018; Wu et

55 al., 2015). In this case, the glacial contribution to flow, and its moderating effect in warmer months, will eventually vanish.

Warming temperatures have also caused significant permafrost degradation in alpine regions around the world, including the Qilian Mountain cryosphere (Gao et al., 2018; Ma et al., 2019; Niu et al., 2010; Song et al., 2019). Permafrost acts as an impermeable boundary to infiltrating water. For this reason, permafrost dominated catchments tend to have higher peak, and lower base flows, with primarily short, lateral groundwater flow paths (Carey and Woo, 2001; Niu et al., 2010; Ye

60 et al., 2009). When permafrost degrades, hydraulic conductivity increases and water can infiltrate to deeper depths and take longer flow paths (Ma et al., 2019; Niu et al., 2010). This results in lower peak flows, as more water infiltrates instead of running off, and higher base flows as more groundwater enters streams (Carey and Woo, 2001; Ma et al., 2019). There is also

an increase in the volume of ground-ice meltwater, which is the release of water stored as ice within permafrost (Ma et al., 2019).

65   Many studies have examined the contribution of glacial melt water to streamflow in the upper Heihe River Basin. These estimates range from 3 %, up to about 10 % for the Heihe River (Chen et al., 2018; Chen et al., 2015; Gao et al., 2018; Li et al., 2018; Niu et al., 2010; Wu et al., 2015). Cryosphere meltwater contribution to smaller rivers, like the Hulugou (Qilian mountains), may be as high as 32 % (Li et al., 2014). This contribution only occurs during the thawing season, which is from April to October (Gao et al., 2018). There is not full agreement on how much these glaciers contribute to total flow, and what
70   streamflow may look like after they disappear.

       Previous work has also quantified the impact of permafrost degradation on streamflow. Increasing winter streamflow trends in alpine regions can be attributed to permafrost degradation processes as there are very few alternate sources of water at this time (Gao et al., 2018; Ma et al., 2019; Niu et al., 2010). This increase is often only significant in basins with initially high permafrost coverage (Niu et al., 2010; Ye et al., 2009) such as the Heihe River Basin. The estimated increase in runoff in
75   the freezing season from permafrost degradation in the upper Heihe River Basin is around 50 % from 1971 to 2010, associated with an 8.8 % loss in permafrost area (Gao et al., 2018). While the change in flow is measured during the freezing season, degrading permafrost could impact baseflow in all seasons (Jones and Rinehart, 2010; Walvoord and Striegl, 2007).

       Numerical, process-based hydrologic models have been used previously to study the Heihe River Basin. Cryosphere response to global warming in the upper basin was studied by Chen et al. (2018) and Gao et al. (2018). Models have also been
80   used to examine a wide range of water resource issues in the middle and lower reaches of the Heihe River Basin. For example, the model HEIFLOW (Hydrological-Ecological Integrated Watershed-scale Flow) has been used to simulate groundwater–surface water interactions, agricultural operations, ecohydrological response, and reservoir impacts amongst other topics (Han et al., 2021; Li et al., 2018; Sun et al., 2018; Tian et al., 2018; Tian et al., 2015a; Tian et al., 2015b; Yao et al., 2018; Yao et al., 2015a; Yao et al., 2015b). However, to our knowledge, no studies have examined the impact of changes in upper basin
85   streamflow due to cryosphere processes on both groundwater and surface water in the middle basin.

       We address this gap by modeling the middle basin response to cryosphere changes using the integrated hydrologic model ParFlow-CLM. ParFlow-CLM is designed to capture interactions between groundwater, surface water and land surface fluxes (Jones and Woodward, 2001; Kollet and Maxwell, 2006; Kollet and Maxwell, 2006; Maxwell et al., 2015). It is thus, well suited to examine evolving watershed dynamics. Using this approach, we explore how groundwater–surface water
90   interactions and water storage in the middle basin evolve as a result of changing streamflow coming from the cryosphere, how these processes vary seasonally, and how projected warming in the middle basin can shift this response.

2 Data and methods

2.1 The study area

The Heihe River Basin is a semi-arid catchment with an area of approximately 130 000 km^2 in Northwest China (Fig. 1). It is located in the Hexi corridor, one of the most arid regions in the world (Lu et al., 2015). The basin decreases in elevation and increases in temperature and aridity moving from south to north. Elevation varies from about 5600 to 900 m (Yao et al., 2018), long-term average temperature ranges from –4 °C to 10 °C (Liu, Y. et al., 2019), precipitation from 800 mm to below 50 mm (Liu, Y. et al., 2019) and potential ET from 700 mm (Zhang, L et al., 2015) up to 2300 mm (Deng and Zhao, 2015).

The Heihe basin has three principal sections: the upper, middle and lower basin. The upper basin is located on the northern edge of the Tibetan Plateau and contains the Qilian Mountains and the headwaters of the Heihe River, the largest river in the basin. This is the primary runoff generation area for the rest of the basin, contributing about 70 % of total river runoff in the lower reaches (Liu et al., 2019; Yang et al., 2014). The middle Heihe is a flat oasis area where most human settlement and economic activity is located. The middle basin uses an estimated 80–95 % of the available fresh water (Deng and Zhao, 2015; Liu et al., 2009; Sun et al., 2016; Yang et al., 2015; Yao et al., 2015b). Of this, 80–90 % is consumed by irrigated agriculture (Chen et al., 2005). The lower basin is primarily Gobi Desert and has little human development. It contains the two terminal lakes of the Heihe River.
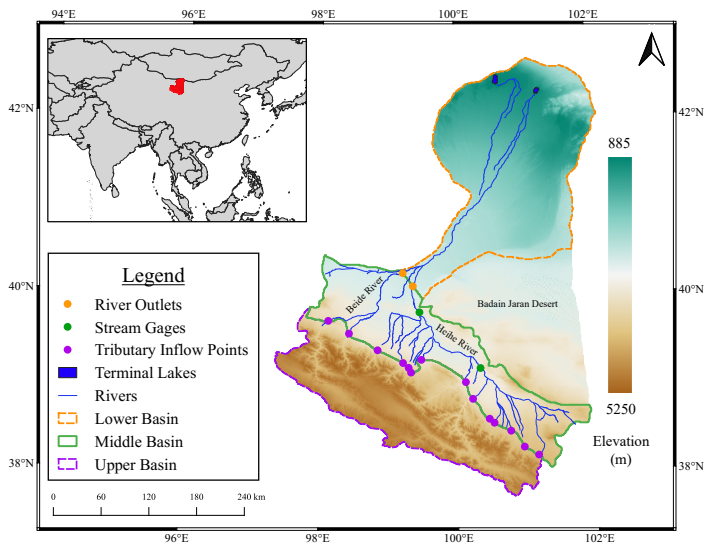
Figure 1. The location of the Heihe River Basin in red in the inset to provide its geographic context in Asia. The upper basin is outlined in purple, the middle basin in green and the lower basin in orange. The Badain Jaran Desert is labelled and is hydrologically connected to the basin. The gaged inlets between the upper and middle basin are in purple. The two river gages on the main stem of the Heihe River are in green and were used for calibration. The location of the Beide and Heihe River outlets are shown in orange. The two terminal lakes are colored in blue at the end the river network. Elevation ranges from ~ 5250 m to 885 m and is shown in the legend to the right.

2.2 Hydrologic modeling approach

We elected to use the integrated hydrologic model ParFlow-CLM in this study. ParFlow-CLM is a fully integrated hydrologic modeling platform that simulates surface and subsurface processes together. It has been used extensively in hydrologic studies of groundwater–surface water interactions, the food-energy-water nexus and climate change in small to large sized basins, including the entire continental US (Condon et al., 2020; Condon and Maxwell, 2014; Ferguson and Maxwell, 2010; Hein et al., 2019). ParFlow-CLM has also been used to model the Central Valley in California, a semi-arid, mountain-valley agriculture system with many parallels to the Heihe River Basin (Gilbert and Maxwell, 2017, 2018; Thatch et al., 2020). In the subsurface, variably saturated flow is solved using the mixed form of Richards' equation. Overland flow is calculated by solving the kinematic wave approximation and Manning's equation (Kollet and Maxwell, 2006). Further details about the workings of ParFlow are provided in: (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006; Maxwell et al., 2015; Maxwell, 2013). ParFlow is coupled to the Common Land Model (CLM). CLM is a land surface model which handles the surface water–energy balance (Maxwell and Miller, 2005; Kollet and Maxwell, 2008).

There has been previous hydrologic model development in the Heihe basin. HEIFLOW is a GSFLOW (Coupled Ground-Water and Surface-Water Flow Model) based model (Markstrom et al., 2008). GSFLOW was developed by the USGS and couples MODFLOW (Modular Ground-Water Flow Model) and PRMS (Precipitation-Runoff Modeling System). The vadose zone, rivers, lakes and other components are defined ahead of time and handled by coupled packages. GSFLOW was enhanced across several studies to include modules which handle surface diversion and pumping (Tian et al., 2018; Tian et al., 2015a; Tian et al., 2015b), dynamic vegetation growth (Sun et al., 2018) and sub grid parameterization of soil and irrigation water (Han et al., 2021).

The primary difference between the ParFlow-CLM model we present here and HEIFLOW is that ParFlow solves variably saturated flow in all subsurface cells. Additionally, overland flow is fully integrated with the subsurface in ParFlow through a free surface overland flow boundary condition that allows rivers to form and disappear as moisture changes (Kollet and Maxwell, 2006). The approach used by ParFlow means that there is no need for a priori specification of saturated zone, vadose zone, river network, etc. prior to simulation. This approach allows for a dynamic evolution of groundwater–surface water interactions and accurate accounting of exchanges of water between surface and subsurface layers. This capability is of

135 importance in modeling the middle Heihe because of the high rate of conversion between surface and groundwater (Yao et al., 2015a; Wang and Cheng, 2000).

2.3 Model inputs

Unless otherwise stated, the data used in the model were originally obtained from the Heihe Program Data Management Center (HPDMC) (http://www.heihedata.org). If data were altered, the publication which details that alteration is given as opposed to
140 the original data repository in Table 1. For example, some of the data underwent pre-processing and parametrization to be used in the construction of the HEIFLOW model for which details can be found in (Tian et al., 2018; Tian et al., 2015a; Tian et al., 2015b; Yao et al., 2015a; Yao et al., 2015b). These data served as our source data and will be referred to as such throughout the paper.

<table>
<tr><td colspan="6">Table 1. Source Data</td></tr>
<tr><td>Starting Variable</td><td>Data Soure</td><td>Original Units</td><td>Spatial Resolution</td><td>Time of Data</td><td>Model Input</td></tr>
<tr><td>Geolayers</td><td>Yao et al. (2015a)</td><td>m</td><td>1 km x 1 km</td><td>2000</td><td>Vertical Discretization</td></tr>
<tr><td>DEM</td><td>HPDMC</td><td>m</td><td>1 km x 1 km</td><td>2008</td><td>X/Y Slopes</td></tr>
<tr><td>Hydraulic Conductivity (K)</td><td>Tian et al. (2015a)</td><td>m day^{-1}</td><td>1 km x 1 km</td><td>2000</td><td>K</td></tr>
<tr><td>Specific Storage (SS)</td><td>Tian et al. (2015a)</td><td>m^{-1}</td><td>1 km x 1 km</td><td>2000</td><td>SS</td></tr>
<tr><td>Specific Yield (SY)</td><td>Tian et al. (2015a)</td><td>[]</td><td>1 km x 1 km</td><td>2000</td><td>Porosity (n)</td></tr>
<tr><td>Groundwater Boundary Condition (GWBC)</td><td>Tian et al. (2015a)</td><td>m day^{-1}</td><td>Boundary points</td><td>average from 2000–2012 Annual Data</td><td>GWBC</td></tr>
<tr><td>Surface Water Boundary Condition (SWBC)</td><td>HPDMC</td><td>m^3 s^{-1}</td><td>14 Stations</td><td>2000–2012 (daily/monthly)</td><td>SWBC</td></tr>
<tr><td>Landcover</td><td>HPDMC</td><td>NLUD-C groups</td><td>1 km x 1 km</td><td>2011</td><td>Landcover, Mannings values</td></tr>
</table>

145   ParFlow-CLM requires gridded inputs for hydraulic conductivity (K), specific storage and porosity. The source data for K were parameterized in Tian et al. (2015a) resulting in 92 unique values ranging from 0.001 to 5.625 m h^{-1}. We aggregated these values for our study into 15 soil and 15 geological units to facilitate calibration by hydrogeologic group. We also assigned a value of 0.001 m h^{-1} to regions in the vertical domain of the ParFlow-CLM model that had no source data. The intention is for this region to be considered bedrock.

150   The source data had information regarding specific yield but not the required variable, porosity. A majority of the domain acts as an unconfined aquifer, with only locally confining conditions (Yao et al., 2015a; Yao et al., 2015b). This was confirmed by the source data where the difference in conductivity between aquitard and aquifer layers was not noticeable

across much of the domain and was consistently much less than the decrease in conductivity with depth. For our model, we

155 assume that specific yield can be used as a reasonable estimate for porosity. This is of course an assumption, but it is likely that the difference between specific yield and porosity is less than the considerable spatial uncertainty in specific yield. We also simplified the values in the specific yield data from 17 unique values ranging from 0.05–0.35 as calibrated in Tian et al. (2015a) to three intervals of 0.1, 0.2 and 0.3 and 0.05 for bedrock. A lumped approach allowed for greater focus in the calibration process on variables the model was more sensitive to such as K.

160 The source data used two unique values for specific storage and was assigned by hydrogeologic unit. The spatial distribution of these values did not vary with depth, and we applied the data unchanged to our model layers. We categorized the porosity and specific storage variables into seven groups, six for each unique pair of porosity and specific storage values and a seventh group representing bedrock for regions in our vertical domain that had no source data. This group was assigned the lowest value from the source data for each respective variable. This corresponds to 0.05 for porosity and 1.0^{-4} m^{-1} for

165 specific storage. These values are corroborated by the literature as reasonable values for bedrock (Huntington and Niswonger, 2012).

The digital elevation model (DEM) (Table 1) was processed to ensure adequate surface drainage of every cell in the domain to either the stream network or domain boundary. The process was accomplished using PriorityFlow, an open-source R package which is a modified priority flood and global slope enforcement algorithm (Condon and Maxwell, 2019). The result

170 is a smoothed, fully draining DEM which was used to produce the X and Y slope files which are required input for ParFlow-CLM. The processed DEM was also used to calculate drainage areas and stream orders which were later used to define Manning's roughness parameters.

The land cover dataset used is the NLUD-C (National Land Use / Cover Database of China) for 2011 (Table 1). Land cover has not been static over the period of simulation. For example, farmland, forest and built-up land have all increased due

175 to the expansion of agriculture and other economic activity in the basin while grasslands, water bodies, wetlands and desert have all decreased, likely converted to the previous land types (Hu et al., 2015). However, land conversion slowed considerably after the year 2000, and most natural oases in the basin had already been converted to farmland by 1975 (Lu et al., 2015). In addition, future land-use patterns are not expected to be appreciably different from the year 2000 (Zhang, L. et al., 2015). For these reasons, we made the decision to use the 2011 land cover map for our simulations. The land cover map was converted

180 from NLUD-C to IGBP (International Geosphere-Biosphere Programme) classification as that is the categorization required by CLM. NLUD-C categories were matched with the closest IGBP group based on descriptions. In some cases, CLM parameters such as LAI or canopy height were altered to better match with the NLUD-C land cover categories. The result is an 18 category IGBP land cover map that matches the 2011 NLUD-C map. Finally, the IGBP land cover map and the stream order map of the domain produced by the topographic processing workflow were used to create a spatially variable Manning's

185 roughness value grid. The conversions for land cover type and stream order to Manning's roughness were obtained from Foster and Maxwell (2018) and the 2015 WRF Hydro User Guide version 3.0 (Gochis et al., 2015).

The climate forcing variables required to run CLM are long and shortwave radiation, precipitation, atmospheric pressure, specific humidity, and u and v wind components. The input climate dataset used is CMFD (China Meteorological Forcing Dataset) detailed in He et al. (2020). It has a temporal resolution of 3-h and a spatial resolution of 0.1 degrees or ~ 10

190   km. Although there are several other climate forcing datasets available, this one was selected as it had almost all the variables required to run CLM, was available for the entire simulation period, and had good spatial and temporal resolution. In order to fit the data to our 1 km modeling grid, the climate data from CMFD were extracted and resampled. Then, the 3 h time step was divided into a 1 h time step where each span of three hours contains the same average data. The CMFD data only contained total wind as opposed to the u and v wind vectors required by CLM. To address this issue, we used wind direction data

195   generated by a high-resolution regional climate model specifically designed for the Heihe River Basin (Xiong and Yan, 2013). This was used to derive the wind direction angle which was then applied to the CMFD wind magnitude to obtain u and v wind components.

2.4 Natural flow state modeling

Our simulations focus on the natural state of the Heihe (i.e., ignoring anthropogenic activities such as groundwater pumping

200   and irrigation). Natural flow models have often been used to isolate anthropogenic contribution to flow regime changes, quantify water available to managers and regulators, and to study catchment response to climatic change, even in heavily managed systems (Terrier et al., 2020). Following this approach, we exclude surface water diversion and groundwater pumping processes from our simulations. These water uses are significant in the Heihe River Basin (Li et al., 2018; Tian et al., 2018). However, the complexity of addressing climate and water-use change simultaneously can make it difficult to assess

205   purely climatic impacts (Terrier et al., 2020). As the broader goal of this paper is to establish trends in water availability in the Heihe in response to future climate warming, a natural flow state model was deemed appropriate.

The inflow to the middle Heihe (our study domain) is almost entirely natural as there is very little water use upstream. However, the stream gauges and wells inside the domain are within areas which have been heavily impacted by surface diversion and pumping for decades (Li et al., 2018; Lu et al., 2015; Tian et al., 2018). To correct for this, we apply a

210   naturalization method to compare model and observed streamflow timeseries. We used a water balance method outlined in Zhang, A. et al. (2015) which assumes that water lost between an upstream and downstream gage roughly corresponds to the water diverted between stations. We quantified the flow lost between the inlet and the HRB2 gage (Fig. 1, closest to outlet) and added that quantity of water back to the streamflow time series for HRB2. It should be noted that this is an approximation, as this method does not take into account the delayed and indirect impact groundwater withdrawals have on river flows (Terrier

215   et al., 2020). Additionally, other tributaries which connect to the Heihe River inside the domain are ungauged. This means, that although the flows in these tributaries are also impacted by diversion and pumping, we cannot adequately correct for them. Thus, despite naturalization, we still expect the flows in our model to be higher than the observed data. For this reason, we focused on matching winter flows when there is little pumping or diversion, as well as the timing of flow rather than magnitude.

2.5 Model configuration and initialization

The modeling domain selected is the middle Heihe as shown in Fig. 1. The horizontal resolution of the model is 1 km with nx = 360 and ny = 270. This is the resolution of most of the source data (Table 1). The domain was divided into 14 vertical layers of varying thickness (note there is no lateral variation in thickness) as follows: 0.1, 0.3, 0.6, 1.0, 10.0, 10.0, 30.0, 30.0, 30.0, 30.0, 30.0, 100.0, 100.0 and 100.0 m. The top four layers correspond to soil and the bottom 10 to geologic layers. This results in a total depth of 472 m. The thickness of the bottom 10 layers was selected to capture the variability seen in the hydrogeologic data of the basin. The HEIFLOW model had five vertical layers that vary in thickness laterally and correspond to the shallow unconfined aquifer, the first aquitard, the shallow confined aquifer, the second aquitard and the deep confined aquifer (Yao et al., 2015b), with a maximum depth of 2094 m (Yao et al., 2015a; Tian et al., 2015b). However, a no-flow boundary was imposed in the model across most of the domain at a much shallower depth. For example, only 8 % of the domain contains data past a depth of 1000 m. We selected a thickness of 472 m for our model as it retains most of the spatial variability of the input data, allows for the resolution of groundwater flow paths on the time order of simulation (11 years), and maximizes model performance.

We applied a constant flux boundary condition along the border between the upper and middle Heihe and no flow boundaries along the rest of the subsurface. The flux across the boundary with the upper basin was calibrated for use in the HEIFLOW model (Tian et al., 2015b; Tian et al., 2018) (Table 1). We adjusted this flux for our modeling domain by subtracting the flow entering below our model depth of 472 m. The remaining flux was applied evenly to all non-bedrock cells along the boundary (i.e., K greater than 0.004 m h^{-1}) (Gleeson et al., 2014; Huntington and Niswonger, 2012). This resulted in a flux value of 1.7^{-4} m h^{-1}applied to all non-bedrock cells on the southern boundary. While groundwater fluxes are likely to change seasonally, there were no available data to support intra-annual values. Additionally, the groundwater flux only makes up approximately 5 % of the average annual water input to the model. Further, during calibration, values between +/– 75 % of the original values were tested and had minimal impact on model output. As seasonal variation is unlikely fall outside these values, we determined a constant boundary condition was sufficient.

There are 14 gaged rivers entering the middle basin from the upper basin (Fig. 1). We injected water into the model according to the flow at these gages. While many of the stream gauges have daily data, others only have monthly data (Table 1). In cases where daily data for a gage do not exist, the daily fraction of monthly flow was calculated for the closest gage with daily data. These fractions were then multiplied by monthly flow to interpolate daily data for the target stream gage. Based on annual averages for streamflow, precipitation minus ET from the climate forcing, and groundwater influx from the boundary condition, the input water breakdown for the middle Heihe domain is approximately 75 % from streamflow, 20 % from precipitation, and 5 % from the groundwater boundary condition.

In order to accurately calibrate the model, an appropriate groundwater configuration was needed to serve as an initial pressure state. To do this, the model was run with a long-term recharge forcing at the land surface until groundwater storage as a percent of recharge was less than 1 %. The long-term recharge forcing was derived from the average difference between

precipitation and ET from the climate forcing data. This resulting pressure state was used as the starting point for all parameter calibration. For model calibration, streamflow observations from the 2011 WY at two gages, HRB1 and HRB2, as well as average water table depth (WTD) at 44 groundwater observations wells was used to assess model performance. Calibration was performed by manually adjusting the groundwater boundary condition, K, and Manning's roughness coefficients.

After final parameters were selected, a new model initialization was performed. This was done because groundwater systems move very slowly, and it was necessary to ensure that the model was once again in equilibrium before running the scenarios. During this final initialization, the model was first brought to steady state with the long-term recharge forcing and then with the climate forcing to bring the groundwater system as close as possible to the conditions at the start of the scenario period (2001 WY). We judged this as complete in both cases when subsurface storage change as a percent of recharge fell below 1 %. This final model pressure state was then used as the starting point for all scenarios outlined below.

2.6 Cryosphere melt scenarios

We designed five scenarios to model the middle basin response to future climate change. (1) A *Baseline* scenario, which uses historic climate and streamflow data to model an unaltered, natural flow state. Next, three scenarios explore the middle Heihe response to changing streamflow input from the cryosphere as a result of warming. (2) a *Glacier* scenario simulating the loss of the glacial contribution to streamflow, (3) a *Permafrost* scenario capturing increases in baseflow as a result of permafrost degradation (4) a *Combined* scenario, which models both the glacial and permafrost impacts on streamflow together. Finally, (5) a *Warming* scenario captures temperature increase in the middle basin.

The above scenarios are modeled using input data from the 2001 to the 2011 WY. This period was selected for three reasons. The first is data availability. The second is that we are interested in changes in groundwater storage. As groundwater is slow moving, if we want to capture the longer-term trends in storage, it is important to simulate for as many years as possible. Last, the period is representative of wet, normal and dry years which makes it ideal to examine climate impacts on hydrologic processes (Tian et al., 2018).

The *Baseline* scenario represents observed historical conditions. Here we apply daily streamflow from historic data at the 14 gage locations between the upper and middle basins (Fig. 1). Figure 2 shows idealized annual streamflow separated into components. For the *Baseline*, all three components are applied at their historic fraction throughout the year. The light-blue, precipitation component, which consists of rain and snowmelt, remains unaltered in all scenarios.

Figure 2. Conceptual model of streamflow with each of the flow components highlighted. Freeze and thaw refer to times of year where we expect water to melt in the cryosphere (thaw) or remain frozen (freeze). The precipitation fraction is light blue and consists of rain and snowmelt. This fraction is unaltered in the scenarios. The glacial fraction is dark blue. This fraction corresponds to 15 % in the scenarios and is removed from streamflow during the thawing season for *Glacier*, *Combined* and *Warming*. The orange fraction represents baseflow and is increased by 50 % in the *Permafrost*, *Combined* and *Warming* scenarios.

310   The *Glacier* scenario is designed to represent a future in which the glaciers in the upper basin have completely disappeared. To do this we remove the fraction of streamflow contributed by glacial melt during the thawing season (April to October). This is represented by the dark blue component in Fig. 2. Freezing and thawing periods were taken from Gao et al. (2018). The contribution of glacial melt to upper basin streamflow has been estimated by several studies as ~ 3–10 % (Chen, 2014; Chen et al., 2018; Gao et al., 2018; He et al., 2008; Li et al., 2016; Wu et al., 2015; Yang, 1991). Here we use a slightly

315   larger value of 15 % which provides an upper bound on what is likely and considers uncertainties in prior estimates. For example, prior estimates are based on historic melt rates and cryosphere interactions and do not account for potential nonlinearity under future climate change. Also, most of these studies exclusively examine the Heihe River and neglect smaller tributaries that feed the lower basin and may have different glacial fractions (Li et al., 2014). By using a larger value for the glacial fraction, we can set a lower bound on future water supply in the region.

320   The *Permafrost* scenario models changes in streamflow as a result of permafrost degradation. The baseflow component of streamflow, shown in orange in Fig. 2 is altered in this scenario. Gao et al. (2018) found that winter flow has increased ~ 50 % in the upper basin from 1970 to 2010. This increase in baseflow corresponded to an estimated 8.8 % reduction in permafrost area. We assumed a similar loss of permafrost area by mid-century and chose to increase our baseflow by 50 %

for the *Permafrost* scenario. Gao et al. (2018) chose to assess the impact of permafrost degradation in the freezing season because this is when other contributions to flow are minimal and it's easier to isolate. However, we apply this increase year-round as opposed to only in the freezing or thawing season because subsurface permeability changes and enlargement of the groundwater reservoir due to permafrost degradation could impact baseflow year-round (Jones and Rinehart, 2010; Walvoord and Striegl, 2007). Although there are also likely reductions in peak flows in the thawing season due to permafrost degradation (Carey and Woo, 2001; Ma et al., 2019), it is difficult to generalize these impacts due to other contributions to streamflow such as precipitation. Thus, a reduction in peak flows would be arbitrary. The *Permafrost* scenario serves as an upper bound for future water supply in the basin.

330 To apply the baseflow change, we performed baseflow separation on the observed streamflow using the digital filtering method outlined in Liu, Z. et al. (2019). Digital filtering separates high- from low-frequency signals, in this case runoff from baseflow. Equation (1) solves for surface runoff at the current time step (Q_{dt}) and Equation (2) solves for baseflow (Q_{bt}). β is the filtering parameter and T is the number of passes with the digital filter. The initial parameterization for β and T was taken from estimates for the upper Heihe basin (Liu, Z. et al., 2019; Zhao et al., 2016). After visual inspection, β = 0.90 and T = 3 were selected as the best fit for our data.

335

$$Q_{dt} = \beta Q_{d(t-1)} + \frac{(1 + \beta)}{2}[Q_t - Q_{(t-1)}] \tag{1}$$

340

$$Q_{bt} = Q_t - Q_{dt} \tag{2}$$

The *Combined* scenario represents a system where changes in flow due to glacial melt and permafrost degradation are occurring at the same time. For this case, we apply perturbations to streamflow that are identical in timing and volume to those made in both the *Glacier* and *Permafrost* scenarios (i.e., 15 % reduction in thawing season flow and 50 % increase in baseflow year-round). These changes are made according to the same reasoning as outlined in the descriptions of the *Glacier* and *Permafrost* scenarios above. In many ways, the *Combined* scenario is the most realistic future representation of streamflow, as we do expect glacial reductions and permafrost degradation to both occur. We performed the isolated glacial and permafrost cases in order to quantitatively isolate the different signatures that these changes have and set upper and lower bounds on water supply.

345

350 The *Warming* scenario is designed to evaluate the impact of future warming in the middle basin on the hydrologic system. The *Warming* scenario is identical to the *Combined* scenario except for a global increase of 2 °C in the CMFD temperature forcing data. We selected 2 °C as it is a reasonable mid-century estimate for global temperature increase (IPCC, 2014). In line with previous studies, we decided that simplifying the temperature increase would allow us to better isolate the hydrologic response to warming (Condon et al., 2020).

3 Results

Results are organized into four subsections. Section 3.1 outlines the performance of the *Baseline* model with regards to streamflow and WTD observations. Section 3.2 covers all results related to streamflow specifically. This includes overall time series, anomalies from *Baseline* and seasonal patterns. Similar results are covered in Section 3.3 for subsurface storage. Section 3.4 contains spatial results which allow for the assessment of warming impact between the scenarios.

3.1 Baseline model performance

To assess model performance, we compared model streamflow to observed data at gage HRB2 on the Heihe River. HRB2 is the furthest downstream gage and the closest to the outlet (Fig. 1). The streamflow performance of the *Baseline* scenario is shown in Fig. 3. It is important to note that we are modeling a natural flow state as discussed in Section 2.4. Observational data are subject to operations like pumping and diversion. As we do not include these processes, we expect our model streamflow to be higher than observed. For this reason, our main targets were to match freezing season baseflow (when there is little diversion) and streamflow timing. We don't expect to see an overall perfect match between model and observations.

Observed streamflow was naturalized according to the method outlined in Zhang, A. et al. (2015). The difference between the observed and naturalized streamflow is illustrated in Fig. 3b, c. Figure 3b shows the model (blue) and observed (red) in the 2001 WY with no naturalization applied. The model matches flow well in the freezing season (November to March) as expected due to minimal diversion. Although, the model still tends to overestimate, likely due to permanent differences in groundwater–surface water interactions between a natural and managed catchment (Terrier et al., 2020). It is also likely that there is continued water usage early and late in the freezing season. In warmer months, when water is more heavily diverted and pumped for irrigated agriculture, observed flows can drop close to zero, while the simulated flow remains high (Fig. 3b). Figure 3c shows the same year comparing the naturalized streamflow with simulated. As expected, there is little change in the winter months. However, in the warmer months, we match the magnitude of flows more closely, showing the strong impact of water usage on observed flows.

Figure 3a shows the model comparison to the naturalized streamflow data for the entire period of simulation. We used Spearman's rho as a metric to determine correlation. It tests that the model is increasing or decreasing at the same time as the observed data and places less weight on the difference in magnitude. In this case, when natural flow is so uncertain, it is more helpful in assessing goodness of fit than the more common Nash–Sutcliffe efficiency. The Spearman's rho is 0.72 showing a good positive correlation. However, the model flow still tends to overestimate the peak flows. This is likely due to the uncertainty in our flow naturalization. The naturalization doesn't account for the impacts of groundwater pumping, or water use on other tributaries draining into the Heihe River.
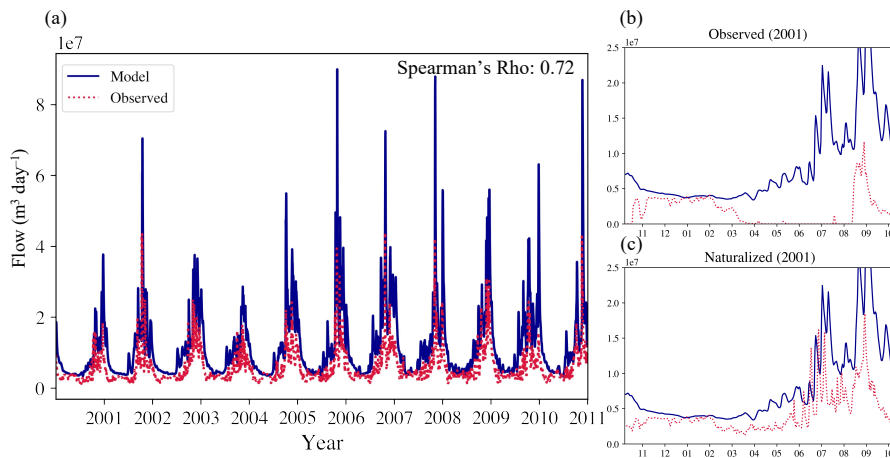
Figure 3. A comparison of observed and modeled flow at the HRB2 gage. (a) shows the model flow in blue for the entire simulation period compared to naturalized flow data in red for HRB2. (b) shows the original observed data in red, which were not subject to naturalization for the 2001 WY. Each number refers to a month, and November 2000 to October 2001 is shown. (c) shows the same year but with the naturalized flows in red. Spearman's rho was calculated for the comparison to naturalized flows.

385   The *Baseline* performance for WTD at 44 observation wells is shown in Fig. 4. The model WTD generally falls within 10 m or less of the observation wells. However, the simulated WTD is significantly shallower where the Heihe River crosses the boundary between the upper and middle basin. This is illustrated by the three dark blue dots (Fig. 4). This is an area of high K ($\sim$ 5.6 m h^{-1}) and as a result, without the pumping and diversion that occur here in the managed system, a much greater volume of water can infiltrate at a rapid rate compared to other parts of the domain and raise the model WTD.

390   There are three additional outlier points where the model WTD is much deeper than expected (orange). The discrepancy at these points is likely to do with our spatial resolution and uncertainty of the actual well locations in the modeling domain. These wells are located where there is a sharp gradient in WTD due to elevation changes between lowlands and a mountain range in the north. Small differences in well location could result in very different predictions of WTD. Overall, based on the results of both streamflow and WTD, we concluded that the model performs satisfactorily when the natural flow

395   state is considered.
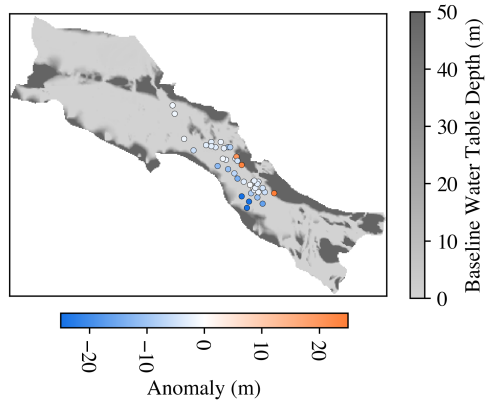
Figure 4. The points refer to the locations of the 44 observation wells in the middle basin. They are colored by the difference of the average observed and model WTD for the *Baseline* simulation across the simulation period (2001- 2011 WY). A negative value (blue) means the model has a shallower water table, while a positive value (orange) means the model has a deeper water table than observed. A value of zero (white) means there is no difference. The background (grey) shows the mean *Baseline* WTD for the simulation.

3.2 Streamflow

A streamflow time series at the outlet of the Heihe River for each of the five scenarios is shown in Fig. 5. Flows are low in the
400 colder months, consisting almost entirely of baseflow. Most of the streamflow occurs in the late summer and early fall, with high and peaky flows. There is strong overlap between the scenarios, showing that all our climate warming cases still have consistent overall behavior during simulation. When scenarios do diverge from *Baseline*, it is typically in the expected order. The *Permafrost* scenario (green), has the highest net increase in flow, followed by the *Combined* (yellow) and *Warming* (orange) scenarios. The *Glacier* scenario (blue) is the only one with flow below the *Baseline*, except for very occasionally,
405 *Warming*.
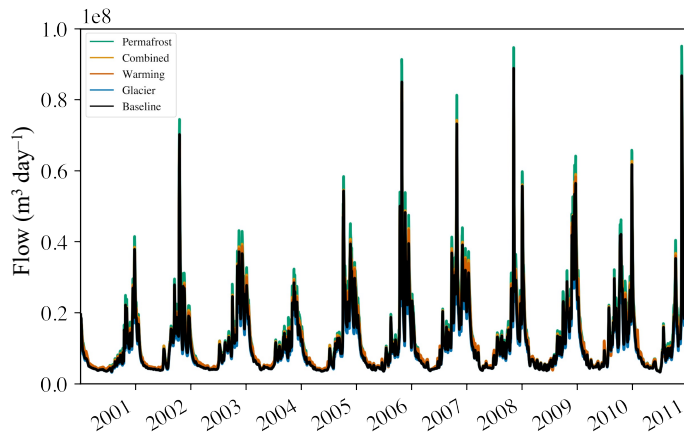
Figure 5. Daily streamflow at the Heihe River outlet. Summed from hourly model outputs for the five scenarios:
Baseline, _Glacier_, _Permafrost_, _Combined_ and _Warming_. The scenarios were run with forcing from 2001–2011 WY.

410   To isolate scenario impact from the baseline dynamics in the model, we primarily discuss our results in terms of how
they differ from the _Baseline_ scenario. Inflow perturbation refers to the difference in streamflow input between the scenarios
and the _Baseline_. The outlet anomaly refers to the difference in flow at the river outlet. Storage anomaly refers to the difference
in storage at a given time. The anomaly fraction (for outlet or storage) refers to the outlet or storage anomaly divided by the
inflow perturbation.

415   Figures 6a and 6b show the inflow perturbation and outlet anomaly plotted together for the Heihe and Beide Rivers.
First, in Fig. 6a for the Heihe River, we see that the outlet anomaly is always smaller in magnitude than the inflow perturbation.
This means that as water moves from the inlet to the outlet, the inflow perturbation signal is dampened. That is, a negative
inflow perturbation (reduction in water from _Baseline_) such as in the _Glacier_ scenario will become less negative. A positive
inflow perturbation (increase in water input from _Baseline_), like the other three scenarios, will become less positive. The only
420   exception is 2011, where the outlet anomaly is slightly less than the inflow perturbation for the _Permafrost_ scenario.

In Fig. 6b for the Beide River, we see this same dampening signal from 2001 to 2004 for the _Permafrost_ and _Combined_
scenarios, and until 2005 for the _Glacier_ scenario. After this year, the inflow perturbation starts to be amplified as opposed to
dampened. This is represented in Fig. 6b by the outlet anomaly plotting below the inflow perturbation in the _Glacier_ scenario,
and above in the _Permafrost_ and _Combined_ scenarios. The _Warming_ scenario however continues to exhibit dampening of the
425   inflow perturbation throughout the simulation.
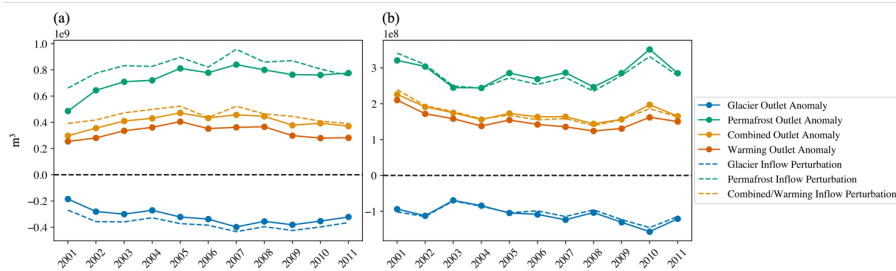
Figure 6. Inflow perturbations are shown in dashed lines, representing the magnitude of flow difference between the four scenarios and the *Baseline*. Solid lines show outlet anomalies, or differences between scenarios and *Baseline* at the outlet. (a) shows these two metrics for the Heihe River outlet while (b) shows the same for the Beide River outlet. The Heihe River is the drainage area for 12 tributaries, while the Beide River is the drainage area for two.

The anomaly fraction refers to the fraction of the inflow perturbation that is still present at the outlet. (Fig. 7). A fraction of one means that the anomaly at the outlet is equal to the inflow perturbation. A number less than one indicates that the outlet anomaly is less than the inflow perturbation and that the signal was dampened. A number greater than one indicates the outlet anomaly is more than the inflow perturbation or that the signal was amplified as it moved downstream. If a fraction is negative, it means that the outlet anomaly is the opposite sign of the inflow perturbation, that is a reduction in flow at the inlet becomes an increase in flow at the outlet or vice versa.

First in Fig. 7a, we see that the fraction is always less than one in all years, and all scenarios for the Heihe River, except in 2011 for *Permafrost*. This corroborates what we see in Fig. 6a which is that the inflow perturbation is almost always dampened in the Heihe River drainage. The first year of simulation shows a large increase in fraction for all scenarios, with a smaller increase for the *Warming* scenario. After this point, there is an increasing trend in fraction for all scenarios from 2001 to 2006. *Combined* and *Glacier* both exhibit a small decrease in fraction from 2003 to 2004. After 2006, trends become much more variable, except for in the *Glacier* scenario which shows smaller fractional changes from year to year. The *Warming* and *Combined* scenarios show very similar patterns because they have the same inflow perturbation. However, the *Warming* scenario is shifted downwards and maintains a smaller fraction throughout. *Warming* trends differ from *Combined* in 2010 and 2011. The *Permafrost* scenario shows slightly less variability compared to the *Combined* and *Warming* scenarios. It also exhibits a switch to amplifying behavior in 2011.

The Beide River shows a large increase in anomaly fraction in the first year in all scenarios except *Warming*, which only has a small increase. This is similar to what is seen in the Heihe River. Likewise, there is a general increasing trend for *Glacier, Permafrost* and *Combined* scenarios until 2006. However, after this year, patterns between the two drainages differ. There is a large jump in fraction for the *Glacier* scenario showing a switch to the amplification of the negative inflow signal.

Permafrost and *Combined* also exhibit a switch to amplifying behavior and the fractions are not as variable as for the Heihe
River. The differences between the *Warming* and *Combined* scenarios are also more apparent in the Beide than Heihe. The
Warming scenario is clearly more variable than *Combined* in the Beide, however they are both similarly variable in the Heihe.

It should be noted that the range of the fractional changes are smaller for the Beide than Heihe River. The fractions
for the Beide range between 0.83 (0.92 without *Warming*) and 1.1 while for the Heihe they range between 0.65 (0.68 without
Warming) and 1.1 across all years and scenarios. Thus, the overall change in anomaly fraction from the start to end of the
simulation period is greater in the Heihe River.
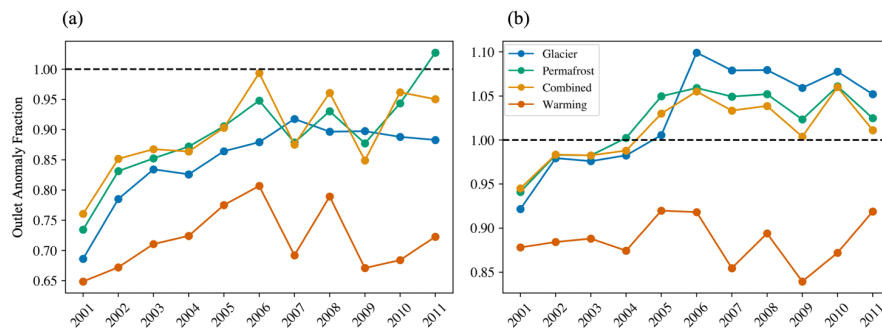
**Figure 7. (a) shows the annual average outlet anomaly fraction (outlet anomaly divided by inflow perturbation) for
the Heihe River and (b) for the Beide River for four scenarios (*Glacier*, *Permafrost*, *Combined* and *Warming*).**

The anomaly fraction is also quantified on a monthly time scale to assess seasonal impacts (Fig. 8). The anomaly
fraction cannot be assessed for the *Glacier* scenario in the freezing season because the inflow perturbation is zero from
November to March. The anomaly fraction generally increases for all scenarios across the thawing season (April to October)
as total flows are increasing. However, the range of months over which the increasing trend persists varies. For *Glacier* and
Warming, it begins in April. *Permafrost* has a few months delay, with the increase not beginning until June. The *Combined*
scenario has the shortest and least consistent increasing window, from July to December. *Warming* and *Permafrost* also
increase until December, whereas *Glacier* only increases until October, after which it cannot be assessed.

Variability tends to be consistent between most months across all scenarios. However, there are some months which
stand out. These typically correspond to the periods with the lowest flows. For example, the inflow perturbation is smallest
(closest to zero) in April for *Glacier*, *Combined* and *Warming*. This month is clearly more variable for *Combined* and *Warming*,
but not for *Glacier*. The *Permafrost* scenario has its smallest inflow perturbation in January, which also corresponds to the
month of greatest variability. However, the entire period of December to April has an inflow perturbation of similar magnitude,
but notably different variability between months.

The anomaly fraction is always positive for *Glacier*, *Permafrost* and *Combined* in every month, however they do differ in the timing of dampening versus amplifying behavior (a fraction less than, as opposed to greater than, one). Looking at the mean anomaly fraction, the *Glacier* scenario only shows amplifying behavior in October. The *Permafrost* scenario has a mean anomaly fraction above one in November and December but shows significant amplifying behavior from September to January. The *Combined* scenario has strong amplifying behavior in April, and the only other month where the mean fraction goes above one is December, with the rest of the year showing dampening. The *Warming* scenario is the only scenario to have a negative fraction at any time of year, with values often going negative in April and May. In these months, even though we added more flow than in the *Baseline* at the inlet, by the time the signal reaches the outlet, the flow is less than in the *Baseline*. Other than these months, *Warming* shows dampening behavior except for in November, December and March.
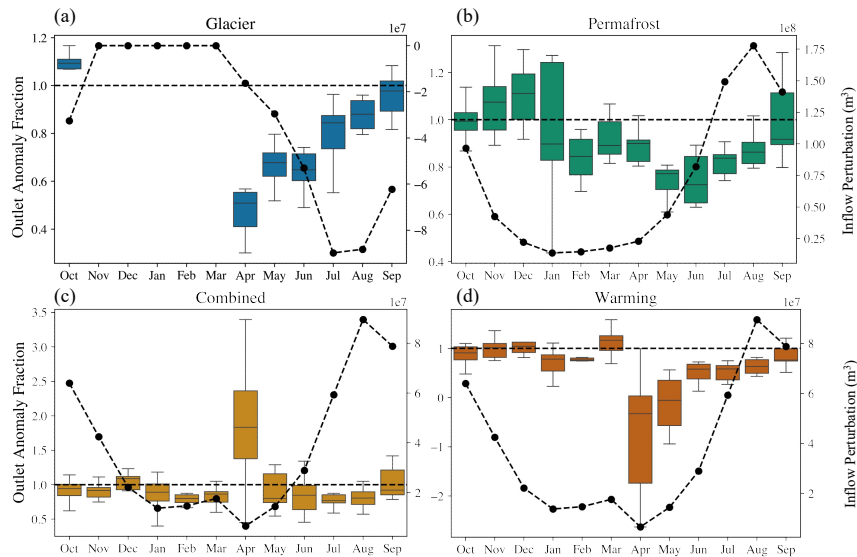
Figure 8. A boxplot of the monthly outlet anomaly fraction for the Heihe River outlet for each scenario is shown. (a) is the *Glacier* scenario, (b) *Permafrost*, (c) *Combined* and (d) *Warming*. The colored boxes represent the interquartile range (IQR) and the line in the center is the mean outlet anomaly in that month across all years of simulation. The whiskers extend from +/– 1.5IQR and cover 99.3 % of the distribution. The outliers are not shown. Each point on the black dashed line is the average inflow perturbation in that month across the entire simulation period.

3.3 Subsurface storage

Subsurface storage has a positive trend in all scenarios, including *Baseline* (Fig. 9a). This increase in storage in all scenarios is possibly attributed to increasing precipitation in the region. However, major findings are taken with reference to *Baseline*. This allows us to isolate the processes we are interested in and remove the influence from variables in the climate forcing that are the same across scenarios, such as precipitation. Relative to *Baseline*, the largest increase in storage is *Permafrost* followed by *Combined* and *Warming*, with *Glacier* losing storage relative to *Baseline*. *Permafrost* and *Combined* have clear positive storage trends over the course of the simulation, except for 2011 for *Permafrost*. *Glacier* has a negative trend. The *Warming* scenario has a positive trend until 2005, after which it becomes variable, with no clear trend.
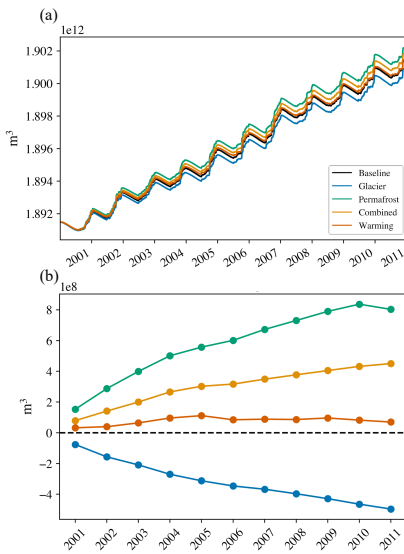
Figure 9. (a) timeseries of total daily subsurface storage for all scenarios and (b) the anomaly of subsurface storage, calculated by subtracting *Baseline* storage from each of the scenarios.

The anomaly fraction for subsurface storage is shown in Fig. 10 for (a) total storage (b) deep (below depth of 10 m) and (c) shallow (above depth of 10 m). The *Glacier* scenario has exclusively negative fractions which indicates a net loss in storage relative to *Baseline*. The opposite is true for *Permafrost* and *Combined* which both have net gains in storage. In 2011, the negative fraction in the *Permafrost* scenario means that the *Permafrost* scenario did not gain as much storage as the *Baseline* in this year. The *Warming* scenario is variable from year to year, sometimes losing and sometimes gaining storage relative to *Baseline*.

All anomaly fractions in all simulations in the total (Fig. 10a) and deep subsurface (Fig. 10b) tend to approach zero. This decreasing trend is not as strong in the total subsurface (Fig. 10a), which has a flatter trend. The shallow subsurface

500 anomaly fraction (Fig. 10c) is much more variable, where the fractions tend to approach zero until 2006, and then diverge again. The combination of deep and shallow subsurface trends likely causes the flattening we see in the total subsurface. A computed average of the anomaly fraction in the total subsurface after 2005 results in 0.06 for *Glacier*, 0.04 for *Permafrost* and 0.03 for *Combined* and –0.02 for *Warming*.
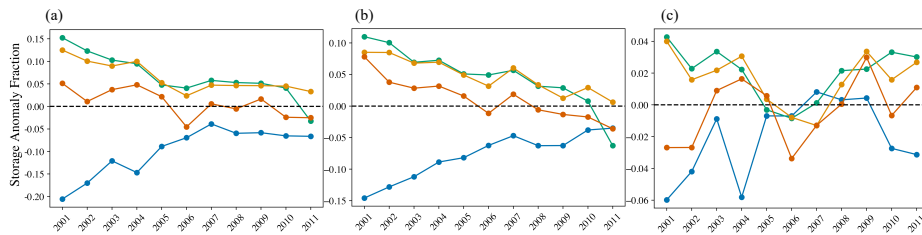
Figure 10. The subsurface storage anomaly fraction (storage anomaly divided by inflow perturbation). (a) shows the time series for total subsurface storage while (b) is for deep subsurface storage, that is depths below 10 m. (c) is shallow subsurface storage which corresponds to depths above 10 m.

505 The mean anomaly fraction almost always shows dampening behavior (less than one) in Fig. 11. This indicates that a quantity of water less than the inflow perturbation applied is added or lost from groundwater storage. The only exception is the *Warming* scenario (Fig. 11d) with positive amplifying behavior (fraction greater than one) in April and negative amplifying behavior (fraction less than negative one) in May. In April, this is interpreted as more water than what was added at the inflow relative to *Baseline* being added to groundwater storage. In May, more water is lost from groundwater storage than the inflow

510 perturbation.

When the fraction is opposite the sign of the inflow perturbation, this represents changes in groundwater storage in the opposite direction as the inflow perturbation. In the *Glacier* scenario, this only occurs in April (Fig. 11a). This means that in April, even though there has been a flow decrease, the monthly storage increase is greater than the *Baseline*. As for *Permafrost* (Fig. 11b), while the mean fraction is never negative, there are months across the simulation period that are

515 negative. This means that even though there is a positive inflow perturbation, there has been a smaller increase in groundwater storage relative to *Baseline* that month. The *Combined* (Fig. 11c) and *Warming* (Fig. 11d) scenario both have months where the mean anomaly fraction is negative, May and June for *Combined* and May through August for *Warming*.

Where the inflow perturbation is positive, the anomaly fraction increases for all scenarios across most of the thawing season (April to October) as flows are increasing. Where it is negative, as for *Glacier* the trend decreases in a similar time

520   frame. The *Warming* scenario on the other hand continues to increase in anomaly fraction until November and exhibits a sharp drop from April to May. In the freezing season (November to March), the *Combined* and *Permafrost* scenarios increase at first in November and December, but then begin a decreasing trend until March and April respectively. The *Glacier* scenario cannot be assessed in the freezing season because the inflow perturbation is zero.

       Variability trends are largely consistent between *Glacier*, *Permafrost* and *Combined*. The variability tends to be

525   higher in the late thawing season (July to October). The variability tends to be smaller in the freezing season when flows are lower. The *Warming* scenario is extremely variable in April and May.
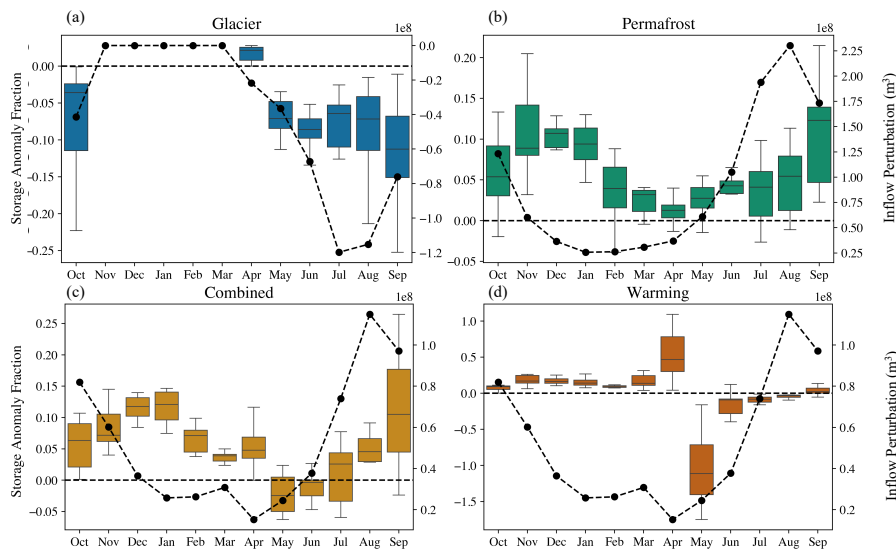
Figure 11. A boxplot of the monthly subsurface storage anomaly fraction for each scenario is shown. (a) is the *Glacier* scenario, (b) *Permafrost*, (c) *Combined* and (d) *Warming*. The colored boxes represent the IQR and the line in the center is the mean storage anomaly in that month across all years of simulation. The whiskers extend from +/– 1.5 IQR and cover 99.3 % of the distribution. The outliers are not shown. Each point on the black dashed line is the average inflow perturbation in that month across the simulation period.

3.4 The impact of warming

530 The *Combined* and *Warming* scenarios have the same inflow perturbations at the inlets (Fig. 12). However, the *Warming* scenario has 2 °C of warming relative to the temperatures in *Combined* applied across the entire simulation domain. When looking at ET, in the *Combined* scenario in January (Fig. 12a), only small regions of ET are greater relative to the *Baseline*. These regions are only in areas near the inlets. They are also more pronounced in areas which have higher flow, like the main stem of the Heihe.
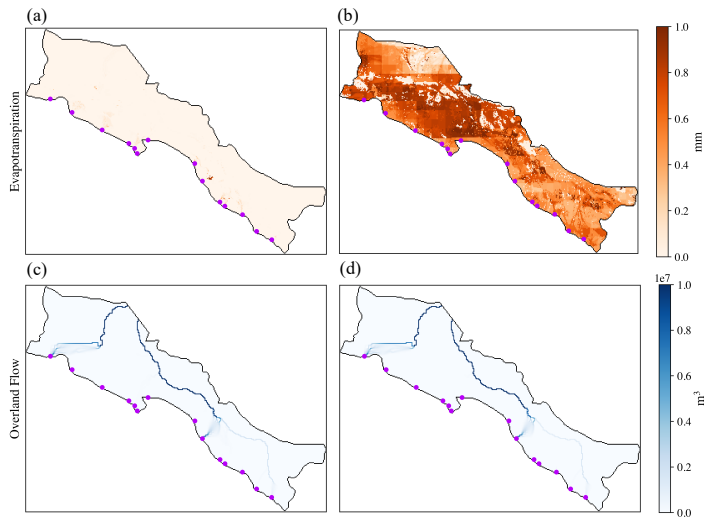
Figure 12. The difference of the average monthly sum of ET in January from *Baseline* for the *Combined* scenario (a) and *Warming* scenario (b). The difference in the sum of overland flow in January from *Baseline* is shown in (c) for *Combined* and (d) for *Warming*.

535

The *Warming* scenario on the other hand shows large differences in ET across the domain relative to *Baseline*. These differences become more dramatic in July (Fig. 13b) where the maximum ET difference almost quadruples. The differences for the *Combined* scenario are also more pronounced and widespread in July compared to January. It is now possible to see ET differences along the river channel into the domain, and not only near the inlets.
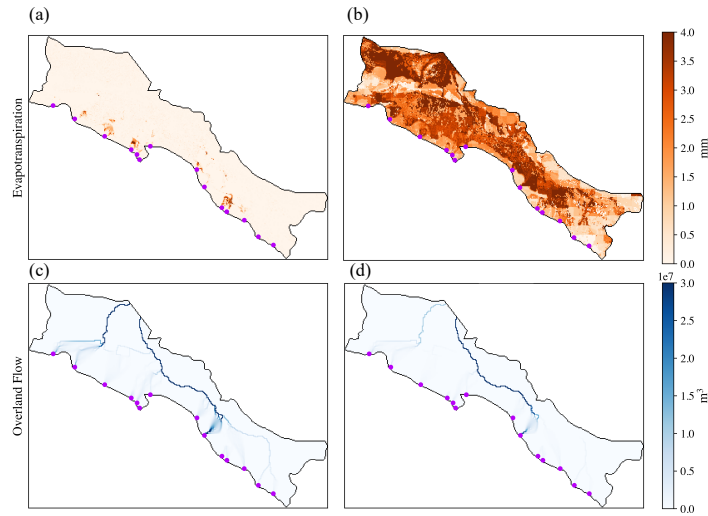
Figure 13. The difference of the average monthly sum of ET in July from *Baseline* for the *Combined* scenario (a) and *Warming* scenario (b). The difference in the sum of overland flow in July from *Baseline* is shown in (c) for *Combined* and (d) for *Warming*.

The river network relative to *Baseline* also differs in January versus July. There is overall greater connectivity of tributaries and higher flows, especially in the *Combined* scenario. Additionally, while the river network between *Combined* and *Warming* looks very similar in January (Fig. 12c, d), the river network as it compares to *Baseline* is noticeably different between them in July (Fig. 13c, d). There is much less flow arriving at both the Heihe and Beide River outlets when comparing *Warming* and *Combined* as well as significantly less connection of inlets to the main river network.

The *Glacier* scenario (Fig. 14a) has an increase in WTD around the river inlets, although not as significant in magnitude as the rising water table in the *Permafrost* scenario (Fig. 14b). The *Combined* (Fig. 14c) and *Warming* (Fig. 14d) scenarios show a similar rise in water table near the river inlets. Several inlets have a rise in water table greater than 20 m near the boundary between the upper and middle Heihe relative to the *Baseline*. The areas of greatest increase are not necessarily directly adjacent to this boundary. The *Warming* scenario has broad areas across the domain where the water table has fallen several meters. This is not seen in any other scenario and so is almost certainly the result of increasing the temperature.
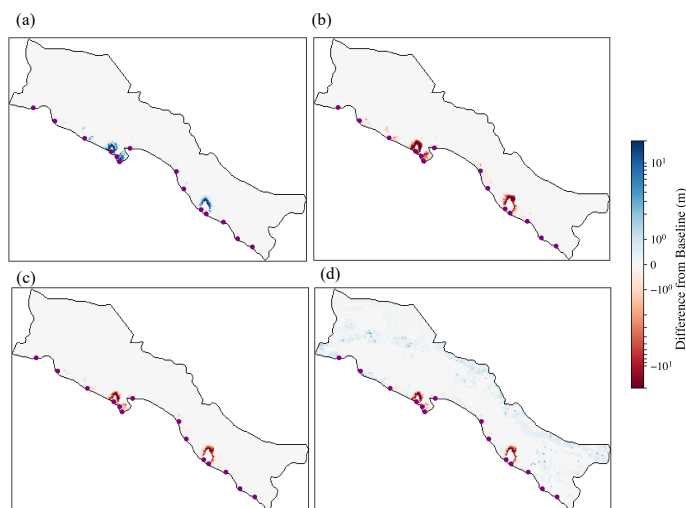
Figure 14. The difference in WTD from *Baseline* after 11 years of simulation. (a) shows the *Glacier* scenario, (b) is the *Permafrost* scenario, (c) is the *Combined* scenario and (d) is the *Warming* scenario. The scale is in meters on a log scale, with red being decreases in WTD (rising water table) and blue being increases in WTD (falling water table).

4 Discussion

4.1 Mediation of cryosphere-based streamflow changes by the middle basin

First, we will explore the impacts from changes in the upper basin cryosphere (i.e., the *Glacier*, *Permafrost* and *Combined* scenarios) on the middle-basin. The *Glacier* scenario has an overall decrease in streamflow relative to *Baseline*, while the *Permafrost* scenario has an overall increase. Throughout all scenarios across all years (apart from 2011 for *Permafrost*), only a fraction of whatever change is applied at the inlet is still present at the outlet (Fig. 7). When streamflow is decreased, stream height falls in the river channel resulting in increased baseflow from groundwater storage releasing to the stream (in locations where the groundwater is shallow and connected). The net result is that streamflow losses are compensated for by groundwater discharge resulting in a negative inflow perturbation at the inlet becoming less negative as we move downstream to the outlet (Fig. 6). When we increase flow, the opposite occurs (i.e., stream level increases can induce increased

groundwater recharge or decrease groundwater discharge). The net effect is a buffering of the streamflow perturbation, in this case the positive inlet perturbation decreasing as you move downstream to the outlet (Fig. 6).

The buffering effect is largest in the first few years when the water table is at its greatest distance from the *Baseline* scenario. For the *Glacier* scenario, this means the most water leaving storage to augment streamflow and for the *Permafrost* scenario, this means the greatest amount of water added to storage. As the water table and stream height equilibrate, the gradient decreases and we expect smaller volumes of water to be exchanged. However, applied differences in streamflow with regards to the *Baseline* result in permanent differences in behavior between the scenarios as a new equilibrium is reached. After 2002, there is little shift in the rolling average of the anomaly fraction. This indicates that the system has reached a new equilibrium.

On average after 2002, 88 %, 92 % and 91 % of the applied streamflow perturbation propagates downstream for *Glacier*, *Permafrost* and *Combined* respectively (Fig. 7), while 8 %, 5 % and 5 % of the perturbation is compensated for by changes in groundwater storage (Fig. 10). This leaves 4 %, 3 % and 4 % of the streamflow perturbation which may be compensated for by a shifting relationship with ET. Figure 13b shows elevated ET for the *Combined* scenario relative to *Baseline*, particularly near inlets and major river branches which supports this assessment.

Multiplying the storage anomaly fraction by the average annual inflow perturbation for each scenario allows us to estimate the total annual change in groundwater storage (note that the magnitude of the inflow perturbation varies across scenarios). We obtain a 39.5 million m^3 decrease of water each year for *Glacier*, an increase of 56.8 million m^3 for *Permafrost*, and an increase of 33.5 million m^3 for *Combined*. This means that the increase in baseflow from permafrost degradation will more than offset the reduction in flow from glacial loss as we see in the *Combined* scenario. It should also be noted that we selected an upper bound scenario for glacial flow reduction. So, it is highly likely that this is a conservative estimate, and that subsurface storage will increase in the future in the middle basin as a result of these two process changes.

Estimates of groundwater pumping in the middle basin can range anywhere from 220 million m^3 (Zeng et al., 2012) to 858.6 million m^3 per year (Tian et al., 2018), after adjusting for groundwater being a presumed 30 % of total water usage (Tian et al., 2018) and the middle basin accounting for 90 % of water usage (Liu et al., 2009; Deng and Zhao, 2015). This results in a range of 3.9–15.0 % of annual usage added to groundwater storage in the *Combined* scenario depending on the pumping estimate used. This is likely to be highly impactful to human systems because the perturbation is applied to rivers, around which most human settlements and agriculture are located. However, this large uncertainty in groundwater usage estimates makes it difficult to assess the impact of these upstream inflow changes.

Watershed response to streamflow changes varies with the initial state of the water table. In the *Baseline* scenario, groundwater is already shallow across a large portion of the domain, especially near the river network (Fig. 5). The areas of the domain with a deep water table at the start of simulation tend to be close to the boundary with the upper basin where there is a larger elevation gradient (Fig. 1). This means that increased flows have a limited area within which to infiltrate and add significantly to storage compared to *Baseline*. This is seen in Fig. 14a, which shows limited regions with a significant rise in water table. Saturated subsurface conditions combined with a high precipitation year result in the behavior illustrated in 2011

Deleted: 15.0–3.9

for *Permafrost*. Here, despite increased flows, less water is added to storage than the *Baseline* because a higher fraction of water is running off (Figs. 9b and 10a). It is worth noting then that over time, permanent increases to subsurface storage may not be as significant as the increase in streamflow would suggest. However, if we consider the system with human impacts of pumping and diversion, it is possible we would continue to see streamflow supplementing subsurface storage in a more stable way year after year.

4.2 Differences between the Heihe and Beide Rivers

The Heihe River drains most of the gaged tributaries (12) coming into the basin, while the Beide River in the western part of the domain only drains two gaged tributaries before crossing into the lower basin (Fig. 1). The result is a significantly smaller stream network draining to the Beide. Additionally, the distance from the inlet to outlet is on average shorter for branches of the Beide than Heihe River (Fig. 1). Finally, there is a greater variation in WTD around the tributaries of the Heihe than Beide (Fig. 3). These factors combined result in a different groundwater response to streamflow perturbation.

From 2003–2005 the inflow perturbation is always less than the outlet anomaly (dampening) for the Beide River (Fig. 6b) and after 2004 it switches to amplifying behavior (fraction greater than one) of the inflow perturbation signal (Fig. 7b). The Heihe River, apart from 2011 *Permafrost*, always shows dampening behavior (fraction less than one). In Fig. 14a, the two leftmost gages (purple markers), which drain to the Beide River, have little change in WTD relative to *Baseline*. This implies that there is little ability for the streamflow signal to be buffered by interactions with groundwater storage. This would explain the lack of significant dampening behavior in the Beide, while for the Heihe we see dampening that is present but diminishing over time. There are several possible explanations for cases of amplifying behavior. First, in scenarios like *Permafrost* and *Combined* which have a net increase in flow, reduced infiltration of that signal over time as shallow groundwater storage fills results in increased runoff. In the *Glacier* scenario, where there is a reduction in flow, an amplification in this negative signal in the Beide may be due to falling water tables (Fig. 14a) which induce further infiltration and streamflow losses (Fig. 7b).

4.3 Seasonal differences

The difference in the intra-annual patterns between the scenarios largely depends on if we are in a baseflow, or runoff dominated month. In general, the applied streamflow perturbations are most dampened in the early thawing season and summer (i.e., the outlet anomaly fractions are smallest). For example, April is the month with the lowest outlet anomaly fraction for the *Glacier* scenario. In this month, on average ~ 50 % of the streamflow reduction is buffered either by the release of groundwater from storage and reduced ET (Fig. 8a). This is compared to September, where essentially none of the reduction is buffered, with a mean outlet anomaly fraction close to one (Fig. 8a). This means that more flow than expected based on the magnitude of the loss of the glacial fraction will arrive at the outlet during the spring and early summer. Conversely, in scenarios where flow is increased relative to *Baseline*, such as *Permafrost* and *Combined*, the month with the smallest mean anomaly fractions occurs a little later (June for *Permafrost* and July for *Combined)*. Increased infiltration into the subsurface and greater losses to ET result in a larger fraction of the streamflow increase being lost before the basin outlet. In these cases,

despite increased flows at the inlets mediated by permafrost changes, there will be less surface water available in the stream network than expected from spring to mid-summer.

To determine if these differences in anomaly fraction are due to changing relationships with groundwater storage, we can look at Fig. 11. First, for the *Glacier* scenario the only month with a positive fraction is April. In April, less storage is lost relative to *Baseline*. This is counterintuitive when considering the small outlet anomaly fraction (Fig. 8a). However, the *Glacier* scenario has the same inflow as the *Baseline* during the freezing season, allowing for large increases in groundwater storage. The switch from a positive to negative fraction from April to May signifies that any surplus storage gained in the freezing season is lost by May. This accounts for the degree of dampening seen in the outlet anomaly fraction in April. The rest of the thawing season, the *Glacier* scenario gains less storage than the *Baseline*.

In the *Permafrost* and *Combined* scenarios, the storage anomaly tends to increase throughout the thawing and into the early freezing season (Fig. 11). While the magnitude of the inflow perturbation is increasing, there is also an increase in the variability and range of the storage anomaly fraction. The variability tends to decrease in the freezing season, while the storage anomaly fraction remains high. Looking at July as an example (Fig. 11b, c), depending on the year up to 10 % of the inflow perturbation could be added to subsurface storage. However, in other years that fraction can be negative. That is, despite elevated flow over *Baseline*, less storage was added for the scenario that year. This is likely related to increasing flow across the entire thawing season. If subsurface storage near the stream network is fully saturated, then more of the inflow perturbation will pass through to the outlet and not infiltrate. This is reflected in outlet anomaly fractions approaching one across the thawing season (Fig. 8). Regardless of if there is a decrease or increase in flow, at the end of the thawing season there is sufficient flow to saturate the subsurface adjacent to the stream network.

Storage anomaly fractions in the freezing season tend to be above zero and less variable (Fig. 11b, c). Lower winter flows result in a smaller likelihood of oversaturating the subsurface near the river network. Second, lower connectivity of tributaries to the main stem during this low flow period increases the amount of streamflow that infiltrates before arriving at the outlet. This second point can be visualized with differences in river network connectivity between January and July (Figs. 12c and 13c). This means that changes to subsurface storage are more consistent in the freezing season.

To summarize, the results above suggest that with glacial loss, surface water supply in spring and summer will be more stable than expected due to supplementation by groundwater. However, this means that in the late summer and fall, groundwater supplies will be more depleted than in the past. As for *Permafrost* and *Combined*, a larger fraction of the streamflow perturbation infiltrating in spring and summer means that flows will not be as high in the stream network as would be expected from the magnitude of the flow increase. However, streamflow should be at anticipated levels in the late summer and fall. As for groundwater, the supply should be more robust overall, but the permafrost mediated changes to baseflow will have a more predictable impact on groundwater storage in the freezing than thawing season.

4.4 The influence of warming temperature

Increasing the temperature in the middle basin changes many of the overall impacts of the *Combined* scenario discussed in previous sections. Changes to streamflow impact WTD and ET in a limited area of the domain (Fig. 14). *Warming* the domain on the other hand, will impact the entire middle basin. In Figs. 12b and 13b, ET is elevated across the domain compared to the *Baseline*. Increases in ET across the domain in the *Warming* scenario result in lower flows on the main river stems, and loss of connection of smaller tributaries in the *Warming* scenario compared to *Combined,* even though they have identical inflows (Fig. 13c, d). Impacts are less pronounced in January (Fig. 12c, d) when ET is lower.

Increased ET reduces shallow groundwater storage and decreases the chance of oversaturating the subsurface during high flow summer periods. These two factors combined cause dampening to persist throughout the simulation period for the *Warming* scenario. For example, in Fig. 6a, b both rivers have consistent dampening behavior in each year of simulation. This is not the case for the *Combined* scenario which switches to amplifying behavior in the Beide and is variable in the Heihe. Likewise for the outlet anomaly fraction, more than 80 % of the flow increase is never present at the outlet for the Heihe River (Fig. 7a). Less runoff reaching the stream network due to increased ET also contributes to this result.

The *Warming* scenario also has a small net loss in subsurface storage (equating to ~ 0.24 % of annual use) (Figs. 10a and 14d). This is due to diffuse, small drops in WTD throughout the domain. The *Warming* scenario does have a similar rise in water table near the river inlets as in the *Combined* scenario (Fig. 14c, d) which does not fully counteract the losses in subsurface storage. The *Warming* scenario also has more variable and negative fractions compared to *Combined* in the shallow subsurface (Fig. 10c). Additionally, less water available to infiltrate results in a steady declining trend in the deep subsurface overtime (Fig. 10b). If we assume that rising WTD and increasing groundwater storage near the inlets slows down overtime as indicated, reductions in groundwater storage due to warming may be more significant relative to *Baseline* than they appear initially.

The *Warming* scenario also shows markedly different behavior than the *Combined* scenario in the spring. In April and May when streamflow increase at the inlet are small, increases in ET are larger than the streamflow perturbations (Fig. 8d). As a result, the net impact is a streamflow decrease at the outlet relative to *Baseline.* The *Warming* scenario also gains less storage compared to *Baseline* throughout the summer, rarely showing increases in subsurface storage relative to *Baseline* until October (Fig. 11d). This differs from the *Combined* scenario which shows relative increases to subsurface storage starting in July (Fig. 11c). The behavior between the two scenarios for both the outlet and storage anomalies are similar in the freezing season where the impact of increased ET, even with 2 °C warming, is minimal. Ultimately, the benefits of a higher flow regime will not be as strong in the middle basin in conjunction with the impacts of warming.

4.5 Caveats

We have made several simplifying assumptions throughout this research in order to design a well constrained experiment. However, these assumptions may also influence our findings. We briefly discuss three principal assumptions in our research

and how they may impact our results. First, we model a natural flow state even though the middle basin is subject to intensive surface and groundwater usage. Our results are valuable to understanding the physical processes and progression of upstream flow changes on the middle basin. However, these impacts will change when modeled with water usage. Next, we only looked at perturbations to streamflow related to temperature changes in the upper basin. There are other processes that may occur under future climate change which we did not address such as precipitation changes. However, precipitation trends are less predictable and difficult to disentangle from the impact of permafrost degradation. Last, it would be valuable to run the model for a longer period. This would allow for a better analysis of the long-term response of groundwater storage to changing streamflow. However, we do remain constrained by data availability.

There are two main ways we would like to expand upon this work. First, adding water management operations in the middle basin would give a more realistic view of how these changes will impact the modern basin. While the physical processes are unlikely to change, the magnitude of the impact will shift. Second, it would be ideal to link the middle basin domain to a model of the upper basin. That way, glacial melt and permafrost degradation would not be simplified and could be linked directly to processes modeled in the middle basin. This would allow for a more physically based change in flow timing and magnitude.

5 Conclusion

Climate warming in the upper basin cryosphere is essentially inevitable. The disappearance of glaciers will decrease overall streamflow, while permafrost degradation will increase baseflow. Examining the downstream impacts on an ecologically and economically important region, such as the middle basin of the Heihe, is of critical importance. Through targeted changes to upper basin discharge and middle basin temperature, this study provides valuable insight into the future of water resources in the middle basin. Overall, our results indicate that there will likely be an increase in streamflow and groundwater storage from combined changes to discharge coming out of the cryosphere. Additionally, even when reductions are severe such as in the *Glacier* scenario, impacts to middle basin water supply are not as extreme as expected. Groundwater exchanges can mediate some of the short-term impacts and dampen the overall shift in streamflow volume.

These cryosphere mediated shifts vary seasonally in their influence on the stability and volume of surface and groundwater storage. This is important to understanding the impact the use of surface or groundwater may have at different times of year for activities such as irrigation. Groundwater supplementation of the glacial fraction makes the stream network more resilient than expected in the spring, but at the cost of reduced groundwater storage in the summer and fall. As for the permafrost mediated changes, they tend to promote increased infiltration in the spring and early summer. This means that while there may be lower flows throughout the stream network than anticipated by the increases observed at the inlet, the result is larger and more reliable groundwater stores in the middle basin.

However, the impacts of warming on the middle basin may be more dramatic than those resulting from streamflow changes. This is because streamflow impacts tend to be in a limited area around the stream network. We find that widespread

warming can overwhelm the signal from streamflow shifts occurring in the upper basin through increasing ET which thereby reduces streamflow and groundwater storage in the middle basin. Our findings are relevant to other semi-arid basins with mountainous water sources that are facing uncertainty and water stress under climate warming.

Code and data availability

All data and programming scripts used to produce this work can be found on Cyverse under the following DOI: https://doi.org/10.25739/kmk7-b046. Any data or scripts not included in this repository can be made available upon request. Finally, any source data that is not specifically attributed in the manuscript can be accessed from the Heihe Program Data Management Center (http://heihedata.org).

Author contribution

AT under the advisement of LC built the model as well as designing and running the scenarios. AT also analyzed all data and created all figures with feedback and guidance from LC. AT wrote the manuscript with revising and editing provided by LC.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

The authors would like to thank the National Science Foundation (Award number 1855912) for funding all aspects of this project. Additionally, our gratitude to our collaborators at Southern University of Science and Technology, particularly Dr. Yong Tian, who was integral in providing source data and information about the Heihe River Basin. Finally, for use of supercomputing resources to run simulations, the National Center for Atmospheric Research.

References

Ashby, S. F., and Falgout, R. D.: A Parallel Multigrid Preconditioned Conjugate Gradient Algorithm for Groundwater Flow Simulations, Nuclear Science and Engineering, 124, 145-159, 10.13182/NSE96-A24230, 1996.

Carey, S. K., and Woo, M.-K.: Slope runoff processes and flow generation in a subarctic, subalpine catchment, Journal of Hydrology, 253, 110-129, 2001.

Chen, R., Wang, G., Yang, Y., Liu, J., Han, C., Song, Y., Liu, Z., and Kang, E.: Effects of Cryospheric Change on Alpine Hydrology: Combining a Model With Observations in the Upper Reaches of the Hei River, China, Journal of Geophysical Research: Atmospheres, 123, 3414-3442, 10.1002/2017jd027876, 2018.

Chen, Y., Zhang, D., Sun, Y., Liu, X., Wang, N., and Savenije, H. H. G.: Water demand management: A case study of the Heihe River Basin in China, Physics and Chemistry of the Earth, Parts A/B/C, 30, 408-419, 10.1016/j.pce.2005.06.019, 2005.

Chen, Y.: Water Resources Research in Northwest China, 1-444 pp., 2014.

Chen, Y., Li, Z., Fan, Y., Wang, H., and Deng, H.: Progress and prospects of climate change impacts on hydrology in the arid region of northwest China, Environ Res, 139, 11-19, 10.1016/j.envres.2014.12.029, 2015.

780 Condon, L. E., and Maxwell, R. M.: Feedbacks between managed irrigation and water availability: Diagnosing temporal and spatial patterns using an integrated hydrologic model, Water Resources Research, 50, 2600-2616, 10.1002/2013wr014868, 2014.

Condon, L. E., and Maxwell, R. M.: Modified priority flood and global slope enforcement algorithm for topographic processing in physically based hydrologic modeling applications, Computers & Geosciences, 126, 73-83, 10.1016/j.cageo.2019.01.020,
785 2019.

Condon, L. E., Atchley, A. L., and Maxwell, R. M.: Evapotranspiration depletes groundwater under warming over the contiguous United States, Nat Commun, 11, 873, 10.1038/s41467-020-14688-0, 2020.

Deng, X., and Zhao, C.: Identification of Water Scarcity and Providing Solutions for Adapting to Climate Changes in the Heihe River Basin of China, Advances in Meteorology, 2015, 1-13, 10.1155/2015/279173, 2015.

790 Ferguson, I. M., and Maxwell, R. M.: Role of groundwater in watershed response and land surface feedbacks under climate change, Water Resources Research, 46, 10.1029/2009wr008616, 2010.

Foster, L. M., and Maxwell, R. M.: Sensitivity analysis of hydraulic conductivity and Manning's n parameters lead to new method to scale effective hydraulic conductivity across model resolutions, Hydrological Processes, 33, 332-349, 10.1002/hyp.13327, 2018.

795 Gao, B., Yang, D., Qin, Y., Wang, Y., Li, H., Zhang, Y., and Zhang, T.: Change in frozen soils and its effect on regional hydrology, upper Heihe basin, northeastern Qinghai–Tibetan Plateau, The Cryosphere, 12, 657-673, 10.5194/tc-12-657-2018, 2018.

Gilbert, J. M., and Maxwell, R. M.: Examining regional groundwater–surface water dynamics using an integrated hydrologic model of the San Joaquin River basin, Hydrology and Earth System Sciences, 21, 923-947, 10.5194/hess-21-923-2017, 2017.

800 Gilbert, J. M., and Maxwell, R. M.: Contrasting warming and drought in snowmelt-dominated agricultural basins: revealing the role of elevation gradients in regional response to temperature change, Environmental Research Letters, 13, 10.1088/1748-9326/aacb38, 2018.

Gleeson, T., Moosdorf, N., Hartmann, J., and van Beek, L. P. H.: A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity, Geophysical Research Letters, 41, 3891-3898, 10.1002/2014gl059856,
805 2014.

Gochis, D. J., Yu, W., and Yates, D. N.: The WRF-Hydro model technical description and user's guide, version 3.0. NCAR Technical Document. , 120 pages, Available online at: http://www.ral.ucar.edu/projects/wrf_hydro/, 2015.

Han, F., Zheng, Y., Tian, Y., Li, X., Zheng, C., and Li, X.: Accounting for field-scale heterogeneity in the ecohydrological
810 modeling of large arid river basins: Strategies and relevance, Journal of Hydrology, 595, 10.1016/j.jhydrol.2021.126045, 2021.

He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., and Li, X.: The first high-resolution meteorological forcing dataset for land process studies over China, Scientific Data, 7, 25, 10.1038/s41597-020-0369-y, 2020.

He, J. Q., Song, G. J., Jiang, X., Li, Q. L., and Wu, X. B.: Relation between glacial meltwater runoff and mountainous runoff in 2006 in four typical river basins of Heihe river water system, Journal of Desert Research, 28, 1186-1190 (In Chinese), 2008.

815 Hein, A., Condon, L. E., and Maxwell, R. M.: Evaluating the relative importance of precipitation, temperature and land-cover change in the hydrologic response to extreme meteorological drought conditions over the North American High Plains, Hydrology and Earth System Sciences, 23, 1931-1950, 10.5194/hess-23-1931-2019, 2019.

Hu, X., Lu, L., Li, X., Wang, J., and Guo, M.: Land Use/Cover Change in the Middle Reaches of the Heihe River Basin over 2000-2011 and Its Implications for Sustainable Water Resource Management, PLoS One, 10, e0128960,
820 10.1371/journal.pone.0128960, 2015.

Huntington, J. L., and Niswonger, R. G.: Role of surface-water and groundwater interactions on projected summertime streamflow in snow dominated regions: An integrated modeling approach, Water Resources Research, 48, 10.1029/2012wr012319, 2012.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II, and III to the Fifth Assessment Report
825 of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)], IPCC, Geneva, Switzerland, 151 pp., 2014.

Jones, J. B., and Rinehart, A. J.: The long-term response of stream flow to climatic warming in headwater streams of interior Alaska, Canadian Journal of Forest Research, 40, 1210-1218, 10.1139/x10-047, 2010.

830 Jones, J. E., and Woodward, C. S.: Newton-Krylov-multigrid solvers for large-scale, highly heterogenous, variably saturated flow problems, Advances in Water Resources, 24, 763-774, https://doi.org/10.1016/S0309-1708(00)00075-0, 2001.

Kang, E., Cheng, G., Lan, Y., and Jin, H.: A model for simulating the response of runoff from the mountainous watersheds of inland river basins in the arid area of northwest China to climatic changes, Science in China Series D: Earth Sciences, 42, 52-63, 1999.

Kollet, S. J., and Maxwell, R. M.: Integrated surface–groundwater flow modeling: A free-surface overland flow boundary
835 condition in a parallel groundwater flow model, Advances in Water Resources, 29, 945-958, 10.1016/j.advwatres.2005.08.006, 2006.

Kollet, S. J., and Maxwell, R. M.: Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model, Water Resources Research, 44, 10.1029/2007wr006004, 2008.

Li, X., Cheng, G., Ge, Y., Li, H., Han, F., Hu, X., Tian, W., Tian, Y., Pan, X., Nian, Y., Zhang, Y., Ran, Y., Zheng, Y., Gao,
840 B., Yang, D., Zheng, C., Wang, X., Liu, S., and Cai, X.: Hydrological Cycle in the Heihe River Basin and Its Implication for Water Resource Management in Endorheic Basins, Journal of Geophysical Research: Atmospheres, 123, 890-914, 10.1002/2017jd027889, 2018.

Li, Z., Feng, Q., Liu, W., Wang, T., Cheng, A., Gao, Y., Guo, X., Pan, Y., Li, J., Guo, R., and Jia, B.: Study on the contribution of cryosphere to runoff in the cold alpine basin: A case study of Hulugou River Basin in the Qilian Mountains, Global and
845 Planetary Change, 122, 345-361, 10.1016/j.gloplacha.2014.10.001, 2014.

Li, Z., Feng, Q., Wang, Q. J., Song, Y., Li, H., and Li, Y.: The influence from the shrinking cryosphere and strengthening evopotranspiration on hydrologic process in a cold basin, Qilian Mountains, Global and Planetary Change, 144, 119-128, 10.1016/j.gloplacha.2016.06.017, 2016.

Liu, B., Zhao, W., Chang, X., Li, S., Zhang, Z., and Du, M.: Water requirements and stability of oasis ecosystem in arid region,
850 China, Environmental Earth Sciences, 59, 1235-1244, 10.1007/s12665-009-0112-7, 2009.

Liu, Y., Hao, L., Zhou, D., Pan, C., Liu, P., Xiong, Z., and Sun, G.: Identifying a transition climate zone in an arid river basin using the evaporative stress index, Natural Hazards and Earth System Sciences, 19, 2281-2294, 10.5194/nhess-19-2281-2019, 2019.

Liu, Z., Liu, S., Ye, J., Sheng, F., You, K., Xiong, X., and Lai, G.: Application of a Digital Filter Method to Separate Baseflow
855 in the Small Watershed of Pengchongjian in Southern China, Forests, 10, 10.3390/f10121065, 2019.

Lu, Z., Wei, Y., Xiao, H., Zou, S., Xie, J., Ren, J., and Western, A.: Evolution of the human–water relationships in the Heihe River basin in the past 2000 years, Hydrology and Earth System Sciences, 19, 2261-2273, 10.5194/hess-19-2261-2015, 2015.

Ma, Q., Jin, H.-J., Bense, V. F., Luo, D.-L., Marchenko, S. S., Harris, S. A., and Lan, Y.-C.: Impacts of degrading permafrost on streamflow in the source area of Yellow River on the Qinghai-Tibet Plateau, China, Advances in Climate Change Research,
860 10, 225-239, 10.1016/j.accre.2020.02.001, 2019.

Markstrom, S. L., Niswonger, R. G., Regan, S. R., Prudic, D. E., and Barlow, P. M.: GSFLOW - coupled ground-water and surface-water flow model based on the integration of the precipitation-runoff modeling system (PRMS) and the modular ground-water flow model (MODFLOW-2005), U.S. Geoological Survey Techniques and Methods 6-D1, Chapter 1, Section D, Book 6, 2008.

865 Maxwell, R. M., and Miller, N. L.: Development of a Coupled Land Surface and Groundwater Model, Journal of Hydrometeorology, 6, 233-247, https://doi.org/10.1175/JHM422.1, 2005.

Maxwell, R. M.: A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling, Advances in Water Resources, 53, 109-117, 10.1016/j.advwatres.2012.10.001, 2013.

Maxwell, R. M., Condon, L. E., and Kollet, S. J.: A high-resolution simulation of groundwater and surface water over most of
870 the continental US with the integrated hydrologic model ParFlow v3, Geoscientific Model Development, 8, 923-937, 10.5194/gmd-8-923-2015, 2015.

Niu, L., Ye, B., Li, J., and Sheng, Y.: Effect of permafrost degradation on hydrological processes in typical basins with various permafrost coverage in Western China, Science China Earth Sciences, 54, 615-624, 10.1007/s11430-010-4073-1, 2010.

Qin, D., Zhou, B., and Xiao, C.: Progress in studies of cryospheric changes and their impacts on climate of China, Journal of
875 Meteorological Research, 28, 732-746, 10.1007/s13351-014-4029-z, 2014.

Qin, J., Ding, Y., and Yang, G.: The hydrological linkage of mountains and plains in the arid region of northwest China, Chinese Science Bulletin, 58, 3140-3147, 10.1007/s11434-013-5768-4, 2013.

Shi, Y., Shen, Y., Kang, E., Li, D., Ding, Y., Zhang, G., and Hu, R.: Recent and Future Climate Change in Northwest China, Climatic Change, 80, 379-393, 10.1007/s10584-006-9121-7, 2006.

Song, C., Wang, G., Mao, T., Dai, J., and Yang, D.: Linkage between permafrost distribution and river runoff changes across the Arctic and the Tibetan Plateau, Science China Earth Sciences, 63, 292-302, 10.1007/s11430-018-9383-6, 2019.

Sun, T., Wang, J., Huang, Q., and Li, Y.: Assessment of Water Rights and Irrigation Pricing Reforms in Heihe River Basin in China, Water, 8, 10.3390/w8080333, 2016.

Sun, Z., Zheng, Y., Li, X., Tian, Y., Han, F., Zhong, Y., Liu, J., and Zheng, C.: The Nexus of Water, Ecosystems, and Agriculture in Endorheic River Basins: A System Analysis Based on Integrated Ecohydrological Modeling, Water Resources Research, 54, 7534-7556, 10.1029/2018wr023364, 2018.

Terrier, M., Perrin, C., de Lavenne, A., Andréassian, V., Lerat, J., and Vaze, J.: Streamflow naturalization methods: a review, Hydrological Sciences Journal, 66, 12-36, 10.1080/02626667.2020.1839080, 2020.

Thatch, L. M., Gilbert, J. M., and Maxwell, R. M.: Integrated Hydrologic Modeling to Untangle the Impacts of Water Management During Drought, Ground Water, 58, 377-391, 10.1111/gwat.12995, 2020.

Tian, Y., Zheng, Y., Wu, B., Wu, X., Liu, J., and Zheng, C.: Modeling surface water-groundwater interaction in arid and semi-arid regions with intensive agriculture, Environmental Modelling & Software, 63, 170-184, 10.1016/j.envsoft.2014.10.011, 2015a.

Tian, Y., Zheng, Y., Zheng, C., Xiao, H., Fan, W., Zou, S., Wu, B., Yao, Y., Zhang, A., and Liu, J.: Exploring scale-dependent ecohydrological responses in a large endorheic river basin through integrated surface water-groundwater modeling, Water Resources Research, 51, 4065-4085, 10.1002/2015wr016881, 2015b.

Tian, Y., Xiong, J., He, X., Pi, X., Jiang, S., Han, F., and Zheng, Y.: Joint Operation of Surface Water and Groundwater Reservoirs to Address Water Conflicts in Arid Regions: An Integrated Modeling Study, Water, 10, 10.3390/w10081105, 2018.

Viviroli, D., and Weingartner, R.: The hydrological significance of mountains: from regional to global scale, Hydrology and Earth System Sciences, 8, 1016-1029, 2004.

Viviroli, D., Archer, D. R., Buytaert, W., Fowler, H. J., Greenwood, G. B., Hamlet, A. F., Huang, Y., Koboltschnig, G., Litaor, M. I., López-Moreno, J. I., Lorentz, S., Schädler, B., Schreier, H., Schwaiger, K., Vuille, M., and Woods, R.: Climate change and mountain water resources: overview and recommendations for research, management and policy, Hydrology and Earth System Sciences, 15, 471-504, 10.5194/hess-15-471-2011, 2011.

Walvoord, M. A., and Striegl, R. G.: Increased groundwater to stream discharge from permafrost thawing in the Yukon River basin: Potential impacts on lateral export of carbon and nitrogen, Geophysical Research Letters, 34, 10.1029/2007gl030216, 2007.

Wang, G. X., and Cheng, G. D.: The characteristics of water resources and the changes of the hydrological process and environment in the arid zone of northwest China, Environmental Geology, 39, 783-790, 2000.

Wu, F., Zhan, J., Wang, Z., and Zhang, Q.: Streamflow variation due to glacier melting and climate change in upstream Heihe River Basin, Northwest China, Physics and Chemistry of the Earth, Parts A/B/C, 79-82, 11-19, 10.1016/j.pce.2014.08.002, 2015.

Xiong, Z., and Yan, X.: Building a high-resolution regional climate model for the Heihe River Basin and simulating precipitation over this region, Chinese Science Bulletin, 58, 4670-4678, https://doi.org/10.1007/s11434-013-5971-3, 2013.

Xu, M., Kang, S., and Li, J.: Evaluation of Water Storage Change of Inland Cryosphere in Northwestern China, Advances in Meteorology, 2015, 1-12, 10.1155/2015/681634, 2015.

Yang, Z. N.: Glacier Water Resources in China, Gansu Science and Technology Press, 1991.

Yao, Y., Zheng, C., Liu, J., Cao, G., Xiao, H., Li, H., and Li, W.: Conceptual and numerical models for groundwater flow in an arid inland river basin, Hydrological Processes, 29, 1480-1492, 10.1002/hyp.10276, 2015a.

Yao, Y., Zheng, C., Tian, Y., Liu, J., and Zheng, Y.: Numerical modeling of regional groundwater flow in the Heihe River Basin, China: Advances and new insights, Science China Earth Sciences, 58, 3-15, 10.1007/s11430-014-5033-y, 2015b.

Yao, Y., Tian, Y., Andrews, C., Li, X., Zheng, Y., and Zheng, C.: Role of Groundwater in the Dryland Ecohydrological System: A Case Study of the Heihe River Basin, Journal of Geophysical Research: Atmospheres, 123, 6760-6776, 10.1029/2018jd028432, 2018.

925 Ye, B., Yang, D., Zhang, Z., and Kane, D. L.: Variation of hydrological regime with permafrost coverage over Lena Basin in Siberia, Journal of Geophysical Research, 114, 10.1029/2008jd010537, 2009.

You, Q., Wu, T., Shen, L., Pepin, N., Zhang, L., Jiang, Z., Wu, Z., Kang, S., and AghaKouchak, A.: Review of snow cover variation over the Tibetan Plateau and its influence on the broad climate system, Earth-Science Reviews, 201, 10.1016/j.earscirev.2019.103043, 2020.

930 Zeng, Z., Liu, J., Koeneman, P. H., Zarate, E., and Hoekstra, A. Y.: Assessing water footprint at river basin level: a case study for the Heihe River Basin in northwest China, Hydrology and Earth System Sciences, 16, 2771-2781, 10.5194/hess-16-2771-2012, 2012.

Zhang, A., Zheng, C., Wang, S., and Yao, Y.: Analysis of streamflow variations in the Heihe River Basin, northwest China: Trends, abrupt changes, driving factors and ecological influences, Journal of Hydrology: Regional Studies, 3, 106-124, 935 10.1016/j.ejrh.2014.10.005, 2015.

Zhang, A., Liu, W., Yin, Z., Fu, G., and Zheng, C.: How Will Climate Change Affect the Water Availability in the Heihe River Basin, Northwest China?, Journal of Hydrometeorology, 17, 1517-1542, 10.1175/jhm-d-15-0058.1, 2016.

Zhang, L., Nan, Z., Yu, W., and Ge, Y.: Modeling Land-Use and Land-Cover Change and Hydrological Responses under Consistent Climate Change Scenarios in the Heihe River Basin, China, Water Resources Management, 29, 4701-4717, 940 10.1007/s11269-015-1085-9, 2015.

Zhang, Y., Xu, C.-Y., Hao, Z., Zhang, L., Ju, Q., and Lai, X.: Variation of Melt Water and Rainfall Runoff and Their Impacts on Streamflow Changes during Recent Decades in Two Tibetan Plateau Basins, Water, 12, 10.3390/w12113112, 2020.

Zhao, W., Li, Z. L., and Wang, Y. H.: Separation and characteristics of base flow in the upper reaches of Heihe River Basin, South North Water Transf. Water Sci. Tech., 14, 26-31 (In Chinese), 2016.

945