Earth System Science Data

Supplementary information for

A Gauge-Based Monthly Natural Discharge Reconstruction Dataset for the Eurasian Arctic, 1952–2025

Bingshi Liu[a,b,*], Xiancai Zou[a,b], Jiancheng Li[a,b,c], and Taoyong Jin[a,b]

[a] School of Geodesy and Geomatics, Wuhan University, Wuhan, China

[b] Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan University, Wuhan, China

[c] School of Geosciences and Info-Physics, Central South University, Changsha, China

Corresponding author:

E-mail address: Bingshi Liu (bs.liu@whu.edu.cn).

Contents of this file

Text S1 to S3

Figures S1 to S7

Table S1

Text S1: TWS reconstruction Model Proposed by Liu et al. (2021)

Liu et al. (2021) proposed an state-update empirical equation that links precipitation, temperature, and TWS:

$$TWS_{t_{i+1}} + d_w = \left(TWS_{t_i} + d_w\right) \cdot e^{\tau_{t_{i+1}}} + P_{t_{i+1}} \tag{S1}$$

$$\tau_{t_i} = a + b \cdot P'_{t_i} + c \cdot T_{t_i}^{\mathrm{melt}'} \tag{S2}$$

where TWS_{t_i} represents the TWS anomaly during the month t_i (i = 1, 2, 3,⋯, n); $T_{t_i}^{melt}$ and P_{t_i} are the temperatures above 0 ℃ and precipitation levels during the month t_i, respectively; $*'$ represents the normalized values of these variabilities; $e^{\tau_{t_i}}$ is the decay factor of TWS during the month t_i, related to the precipitation temperature of the month. Specifically, higher precipitation and temperature correspond to a smaller decay factor. d_w is a storage-offset parameter since GRACE only provides relative changes without an absolute reference value. Using GRACE based TWS observations as reference data, they identified the unknown parameters (a, b, c, and d_w) and reconstructed TWS beyond the GRACE era.

The model is based on three core principles. First, precipitation acts as the primary water input to the basin. Second, the antecedent TWS state determines the remaining storage capacity of the basin. Third, precipitation and temperature jointly regulate the release of stored water from the basin.

However, this model has important limitations for Pan-Arctic applications. It represents only rainfall-driven liquid water processes and does not account for the distinct accumulation and melt dynamics of solid water. During snow accumulation periods, runoff and evapotranspiration are limited, but the model does not distinguish between rainfall and snowfall and instead treats all precipitation as acting on the residual storage term. This structure tends to overestimate water loss during snowfall periods. In addition, the model does not adequately represent snowmelt. The response of snowmelt to temperature differs fundamentally from that of ET, and the magnitude of snowmelt depends on solid water storage rather than on total TWS. These limitations likely explain the poor performance of the original model in reconstructing Pan-Arctic TWS.

Text S2: A Runoff-TWS Relationship Developed by Yi et al. (2023)

Yi et al. (2023) developed an empirical relationship to estimate river discharge from TWS for the pan-Arctic region and demonstrated its reliability. This approach assumes that runoff is proportional to liquid TWS in adjacent months.

$$R_{t_{i+1}} = \alpha_1 \cdot (1 - \beta \cdot SC_{t_i}) \cdot \left(TWS_{t_i}^{\text{liquid}} + d_r\right) + \alpha_2 \cdot (1 - \beta \cdot SC_{t_{i+1}}) \cdot \left(TWS_{t_{i+1}}^{\text{liquid}} + d_r\right) \text{(S3)}$$

Where R_{t_i} is the reconstructed runoff for month t_i. Liquid TWS (TWS^{liquid}) was calculated by subtracting solid TWS (TWS^{solid}) from total TWS. SC represents the snow cover fraction, and β is a sensitivity factor ranging from 0 to 1. The term $(1 - \beta \cdot SC)$ describes basin water release efficiency, which was highest during snow-free conditions and lowest during periods of extensive snow cover, while remaining nonzero due to baseflow beneath persistent snow cover. d_r is a storage-offset term added to TWS anomalies to convert them into absolute storage values. The parameters α_1 , α_2 , β , and d_r were estimated using observed discharge data from the corresponding gauging stations.

Text S3: Apportionment Entropy for seasonal variation

We employed Apportionment Entropy (AE), an information-theory metric, to quantify the seasonality of monthly river discharge. Unlike parametric indices such as the variation and standard deviation, the AE metric is non-parametric and can capture higher-order moments (Konapala et al., 2020). AE measures the disorder or dispersion of an apportionment, making it well-suited for assessing changes in the seasonal distribution of hydrological fluxes. It has been widely applied in hydrology and climatology to assess the seasonality of rainfall (Feng et al., 2013), precipitation, evapotranspiration (Konapala et al., 2020), and runoff (Wang et al., 2024).

To estimate the AE for a given hydrological year *i* (defined as September–August), the total annual discharge AT_i is first calculated by summing the monthly discharge values $M_{m,i}$.

$$AT_i = \sum_{m=1}^{12} M_{m,i} \tag{S4}$$

The AE for year *i* is then given by:

$$AE_i = -\sum_{m=1}^{12} \frac{M_{m,i}}{AT_i} \log_2\left(\frac{M_{m,i}}{AT_i}\right) \tag{S5}$$

Where, a high AE value corresponds to low seasonality. The maximum value, $\mathrm{AE} = \log_2 12$, indicates a uniform monthly distribution, where the discharge is equal in every month. Conversely, a low AE value indicates high seasonality. The minimum value, $\mathrm{AE} = 0$, occurs when the entire annual discharge is concentrated in a single month.

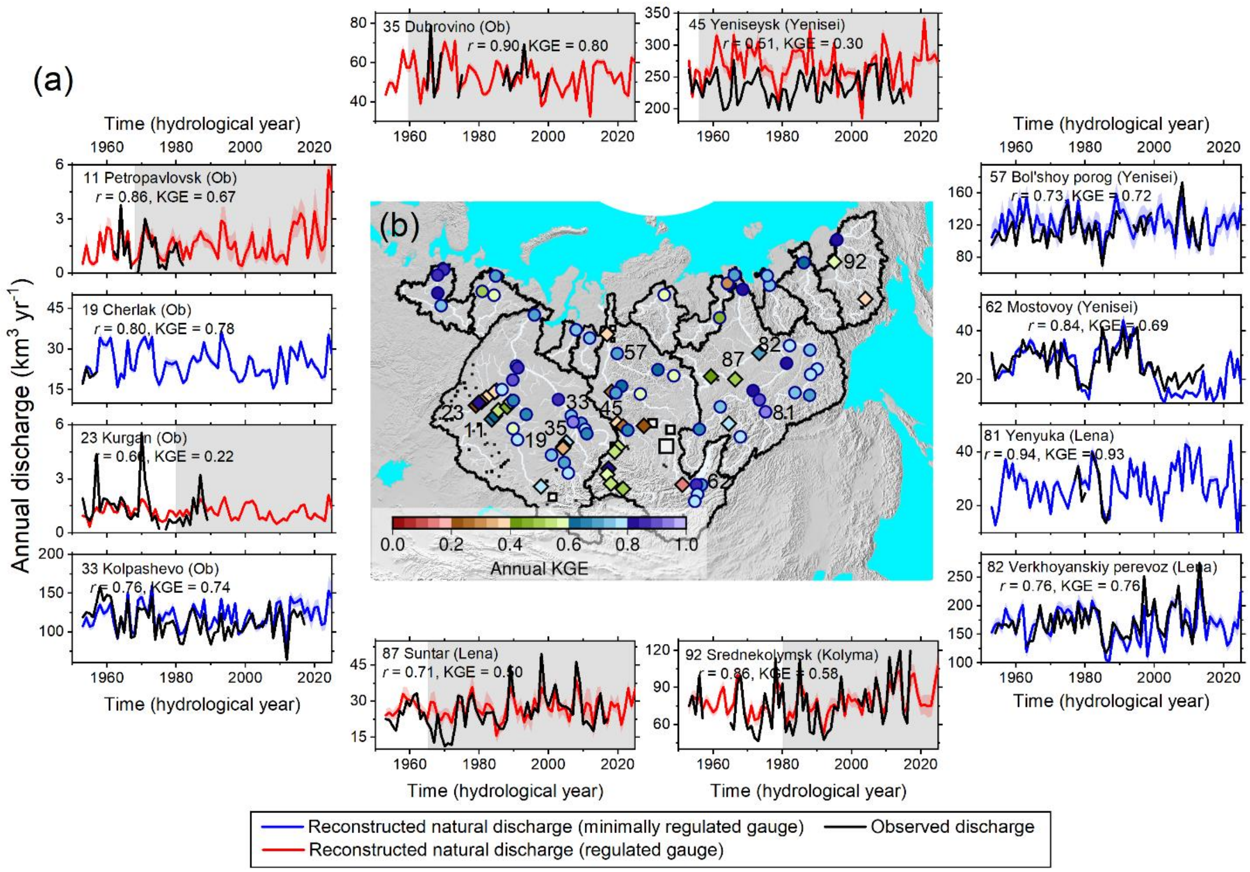

Figure S1. (a) Observed and reconstructed natural annual discharge at selected gauges. Annual values were aggregated by hydrological year from September to the following August. Shaded intervals indicate post-regulation periods. (b) Spatial distribution of annual Kling–Gupta efficiency (KGE) values between reconstructed and observed discharge.

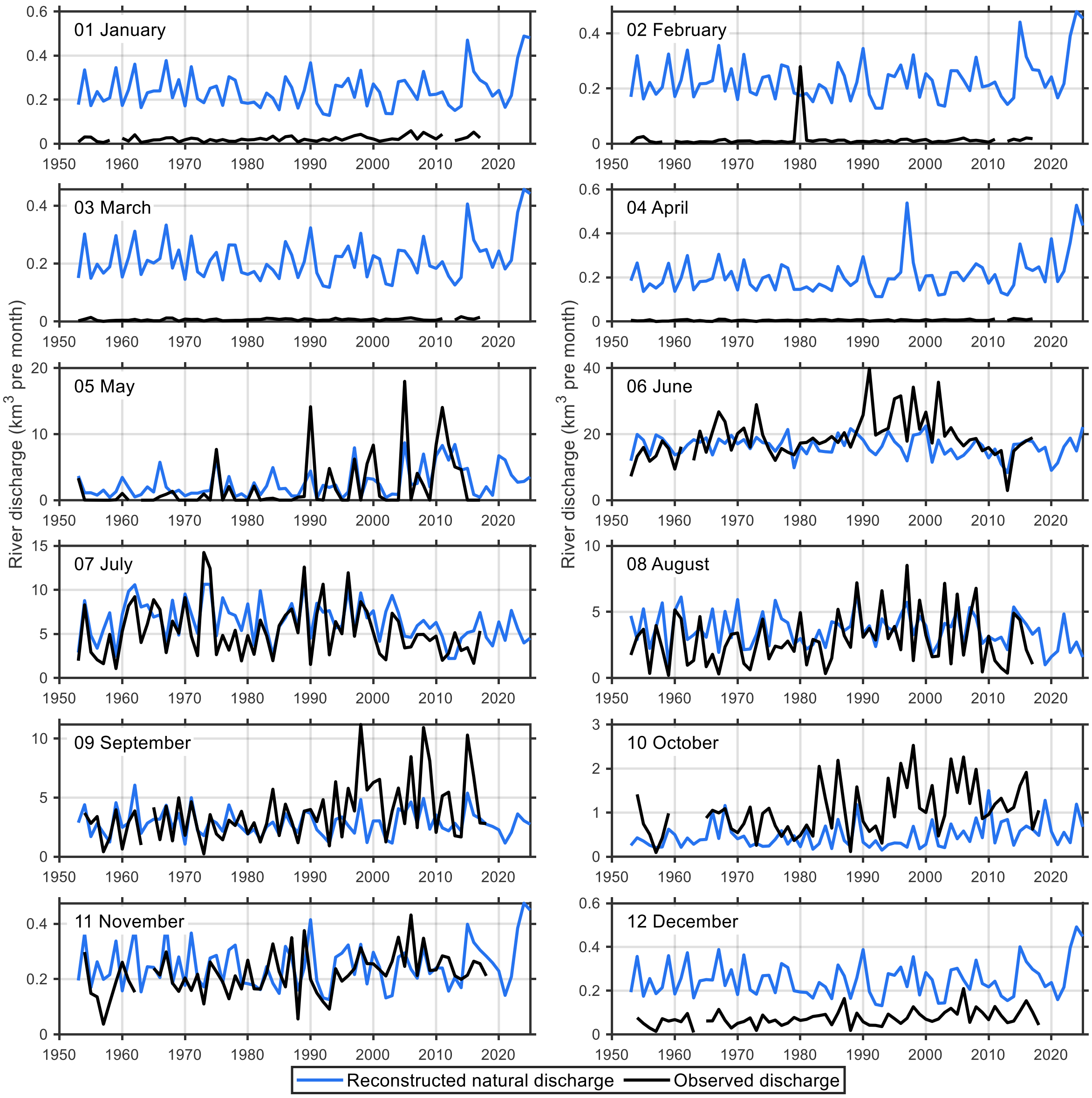

Figure S2. Monthly comparison between reconstructed natural discharge and observed discharge at gauge 68, 7.5km d/s of mouth of river pur in the Olenek basin. The main discrepancy occurred in June during 1990–2000, when observed discharge showed substantially larger magnitude and variability than the reconstruction. In some years, observed June discharge exceeded 20 km^3 $month^{-1}$, accounting for nearly 60% of the mean annual discharge at this gauge. Systematic differences were also found during the cold season from December to April, when reconstructed discharge was generally higher than observations by approximately 0.2 km^3 $month^{-1}$. These differences may reflect combined uncertainties in discharge observations, meteorological forcing, snow-process representation, and the reconstruction framework.

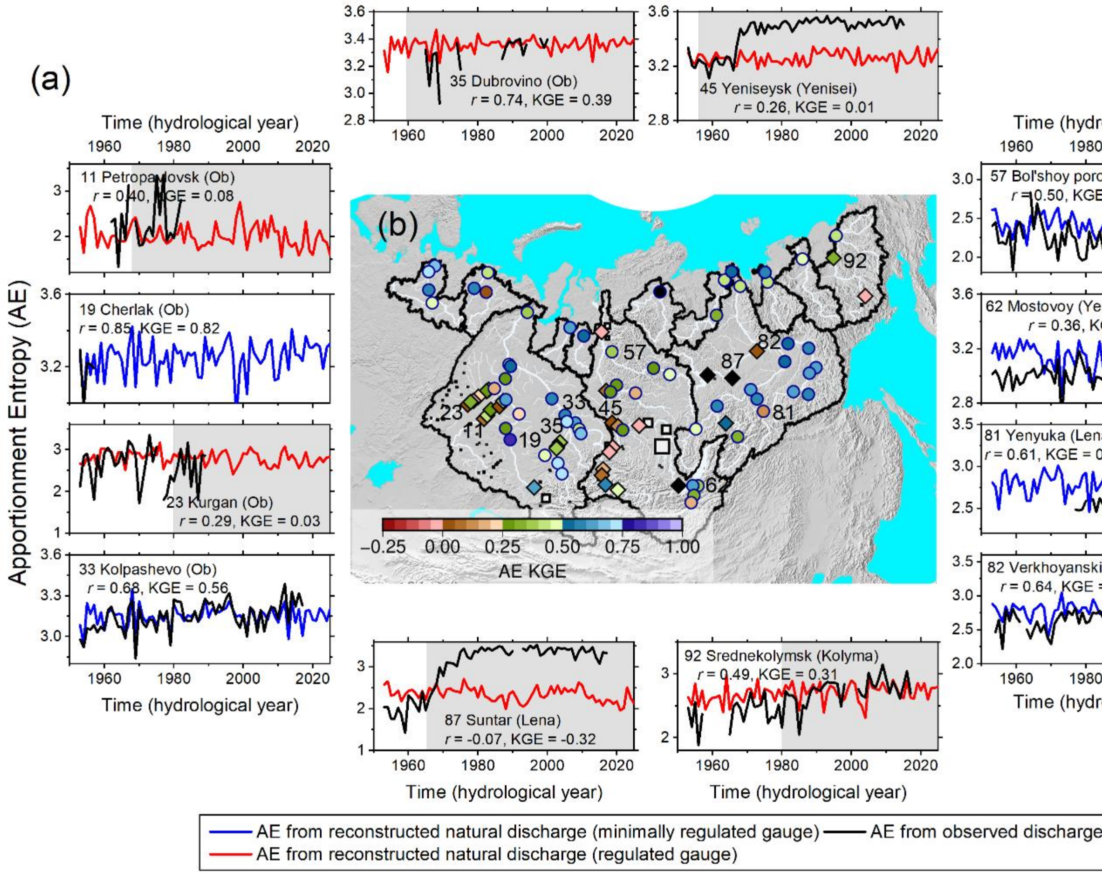

Figure S3. Evaluation of seasonal discharge allocation using Apportionment Entropy (AE). (a) AE time series derived from reconstructed monthly natural discharge and observations at selected gauges. Shaded intervals indicate post-regulation periods. (b) Spatial distribution of Kling–Gupta efficiency (KGE) values between reconstructed and observed AE.

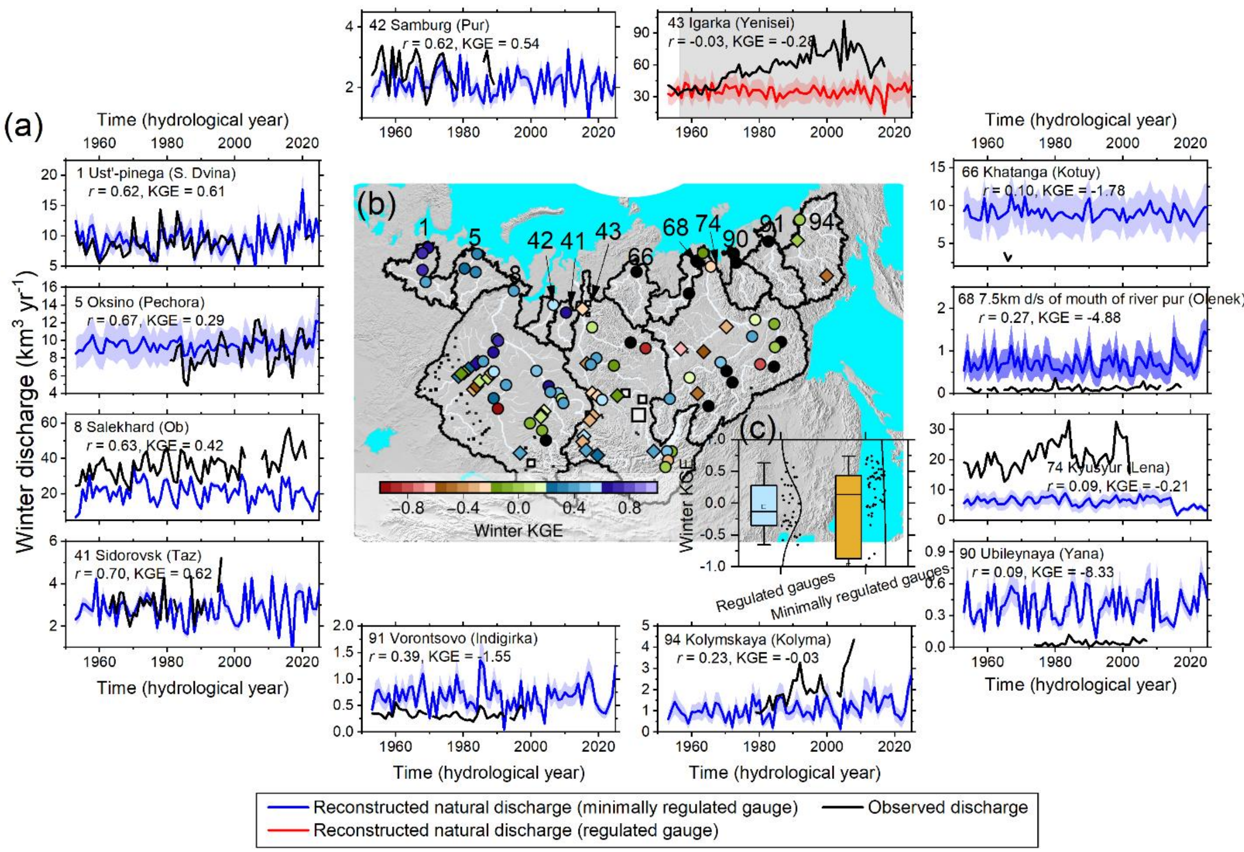

Figure S4. (a) Observed and reconstructed natural winter (DJF) discharge at selected gauges. Shaded interval indicates post-regulation periods. (b) Spatial distribution of winter Kling–Gupta efficiency (KGE) values between reconstructed and observed discharge. (c) Boxplots of winter KGE values at 33 regulated and 61 minimally regulated gauges.

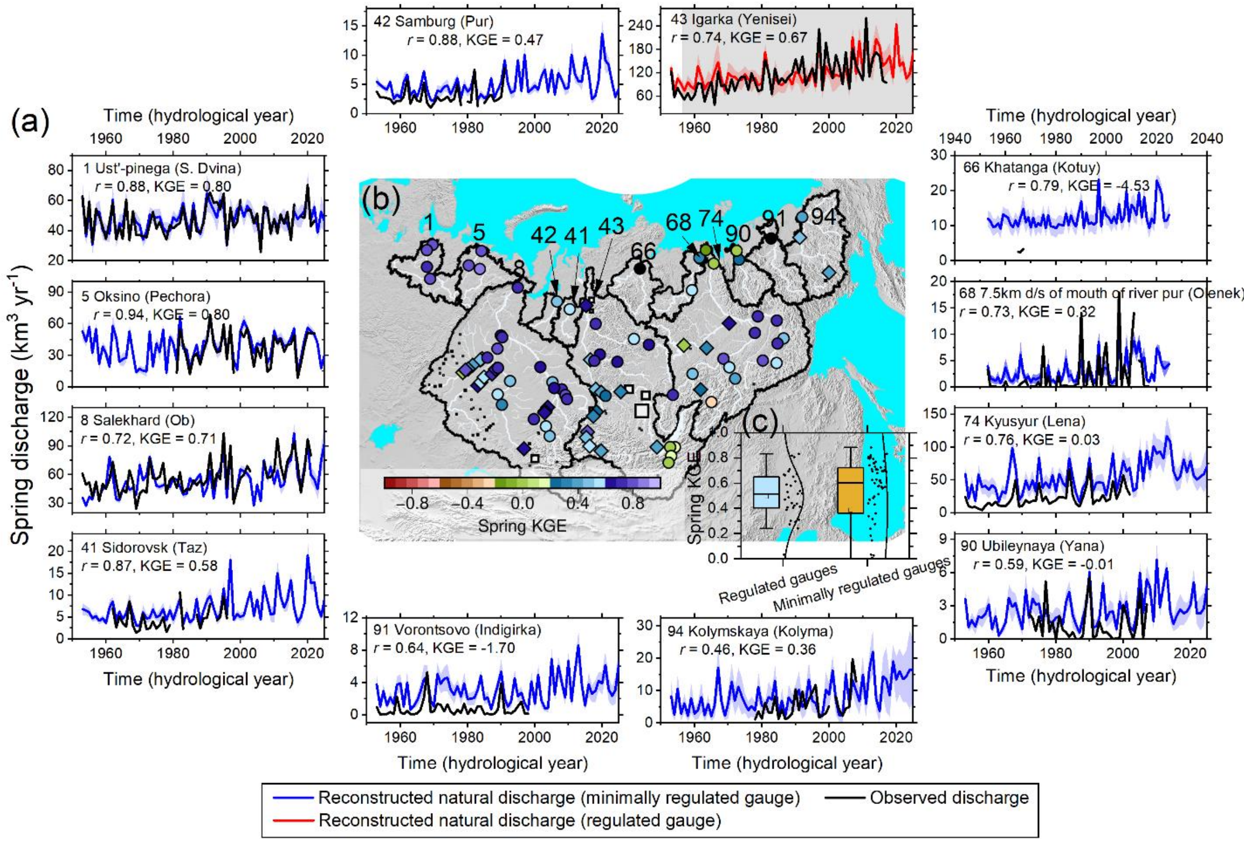

Figure S5. (a) Observed and reconstructed natural spring (MAM) discharge at selected gauges. Shaded interval indicates post-regulation periods. (b) Spatial distribution of spring Kling–Gupta efficiency (KGE) values between reconstructed and observed discharge. (c) Boxplots of spring KGE values at 33 regulated and 61 minimally regulated gauges.

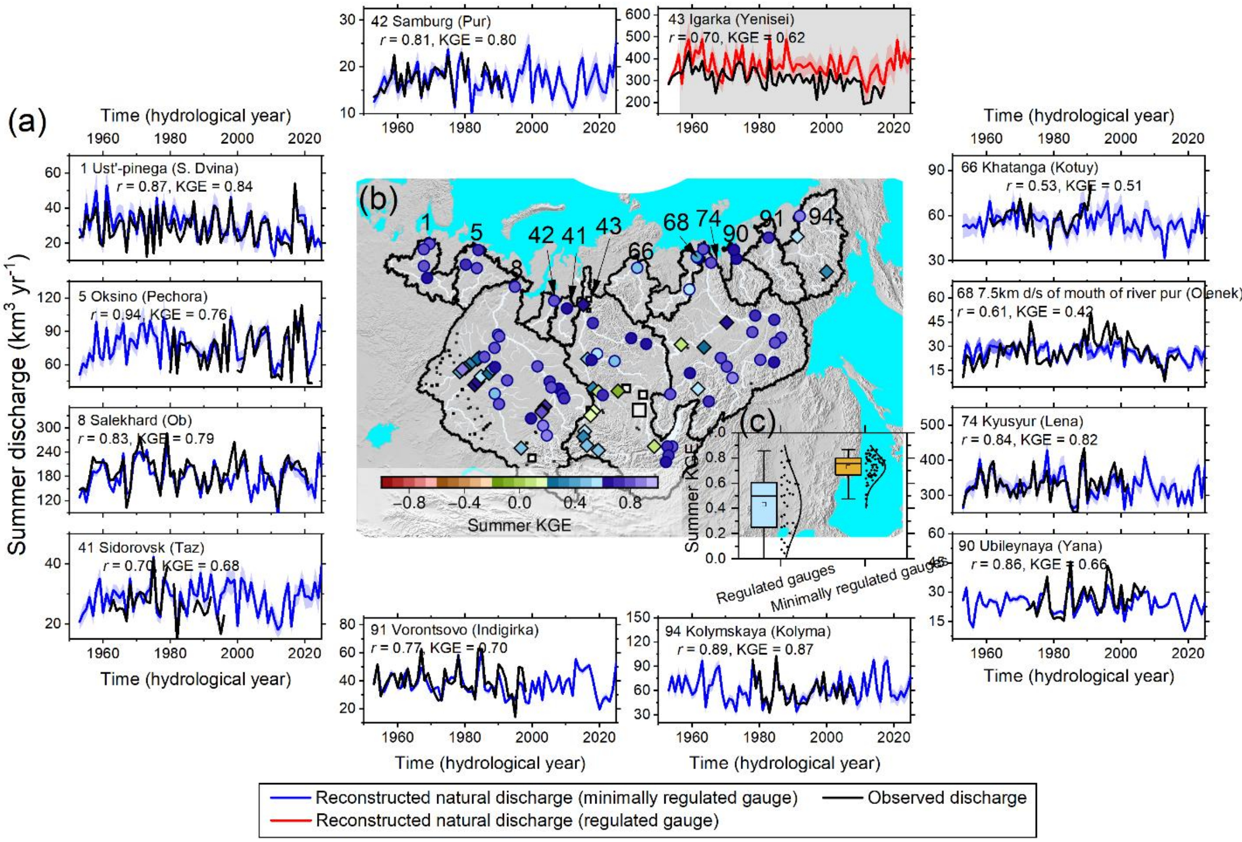

Figure S6. (a) Observed and reconstructed natural summer (JJA) discharge at selected gauges. Shaded interval indicates post-regulation periods. (b) Spatial distribution of summer Kling–Gupta efficiency (KGE) values between reconstructed and observed discharge. (c) Boxplots of summer KGE values at 33 regulated and 61 minimally regulated gauges.

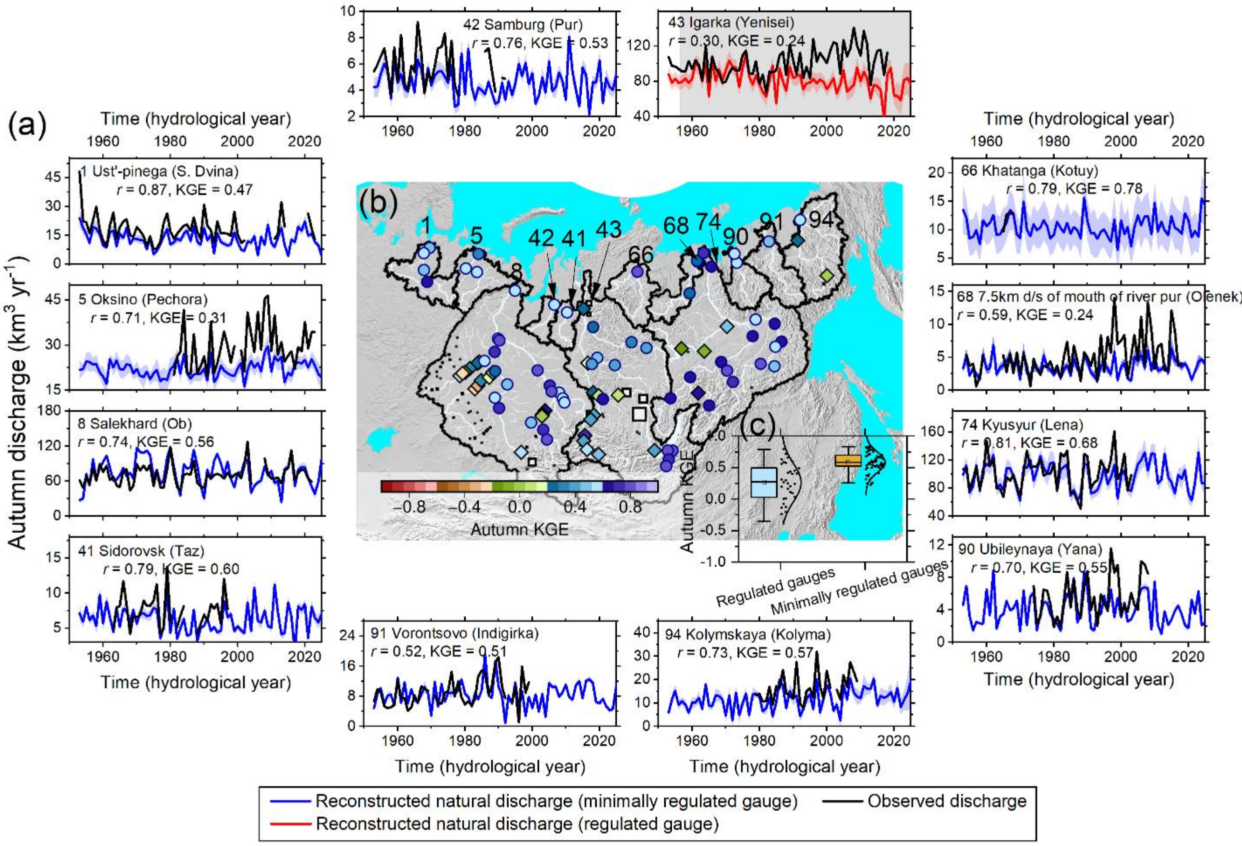

Figure S7. (a) Observed and reconstructed natural autumn (SON) discharge at selected gauges. Shaded interval indicates post-regulation periods. (b) Spatial distribution of autumn Kling–Gupta efficiency (KGE) values between reconstructed and observed discharge. (c) Boxplots of autumn KGE values at 33 regulated and 61 minimally regulated gauges.

Table S1. Acronyms that involved in this Paper

BEST	Berkeley Earth Surface Temperatures
CRU	Climatic Research Unit Time-series data version 4.09
Crocus-ERA5	Crocus-ERA5 snow product
CPC	Climate Prediction Center Global Unified Daily Precipitation/Temperature
CSR	Center for Space Research at the University of Texas
ERA5−Land	ERA5-Land monthly averaged data
ERA5 single	ERA5 monthly averaged data on single levels
GLDAS NOAH	Global Land Data Assimilation System Noah Land Surface Model v2.0
GPCC	Global Precipitation Climatology Centre data Monthly Version 2022
GRACE	Gravity Recovery and Climate Experiment
GRACE–FO	GRACE Follow–On
GRDC	Global Runoff Data Centre
GSFC	Goddard Space Flight Center
JRA-3Q	Japanese Reanalysis for Three Quarters of a Century
JPL	Jet Propulsion Laboratory
KGE	Kling–Gupta efficiency
MCMC	Markov Chain Monte Carlo
SWE	Snow Water Equivalent
TWS	Terrestrial Water Storage
U. Delaware	University of Delaware Air Temperature & Precipitation
WaterGAP	Water – Global Assessment and Prognosis