Supplement of Earth Syst. Sci. Data, 18, 2703–2722, 2026
https://doi.org/10.5194/essd-18-2703-2026-supplement

Open Access Earth System Science Data

Supplement of

A harmonized 2000–2024 dataset of daily river ice concentration and annual phenology for major Arctic rivers

Jiahui Qiu et al.

Correspondence to: Jiahui Qiu (jiahui.qiu@oulu.fi, crisqiu7@gmail.com)

This file includes:

Section S1

Tables S1-S3

Figures S1-S10

S1. Quantifying the potential influence of the fixed-width buffered river mask on gridded RIC in narrow reaches

This section documents (1) how we quantified the prevalence of river reaches narrower than the effective 500 m mask width and (2) a conservative geometric upper bound on the potential positive bias in the reported river ice concentration (RIC) arising from snow-covered non-channel areas included in the buffered mask.

(1) Prevalence of narrower reaches (width < 500 m)

To quantify the proportion of narrow reaches in the analyzed network, we used the GRWL Vector Product (river centerlines) that provides segment-level wetted width information (attribute width_m). For each river, let L_{tot} denote total analyzed centerline length and L_{narrow} denote the subset with width < 500 m. The length-based proportion is:$P_{narrow} = \frac{L_{narrow}}{L_{tot}}$Results are summarized in Table S3. Across all six rivers, the length-weighted proportion of narrow reaches is 12.0%, with the smallest fraction for the Mackenzie River (7.2%).

(2) Conservative geometric upper bound on potential RIC inflation

During winter periods, snow-covered banks included in the buffered mask may be mapped as "ice" as NDSI cannot fully distinguish snow from ice, thereby producing positive bias in the reported RIC for those narrower reaches. For narrower reaches, the mapping uncertainty can be generalized into two conceptual cases depending on the presence of snow on the river banks. When bank snow is absent, mapped ice area is not systematically affected by adjacent bank areas, regardless of ice presence on the watercourse. In this case, the "true" RIC, defined solely by channel ice, is expected to be consistent with the reported RIC. In contrast, when bank snow is present, snow-covered bank areas may be misclassified as ice, leading to a potential overestimation of mapped ice area and RIC. The maximum possible bias can be bounded as:

$$\Delta RIC = RIC_{rep} - RIC_{tru} \leq \frac{A_{bank}}{A_{mask}}$$

where A_{bank} is the bank area inside the buffered mask that is snow-covered, and A_{mask} is the referenced mask area in a 3-km grid-cell.

Under a conservative worst-case assumption—(i) the whole reach is narrower than 500 m, (ii) bank belt is fully snow-covered, and (iii) snow is entirely mapped as ice—the potential "bank belt" area introduced by buffering can be approximated geometrically. With a 250 m buffer applied on both sides, the additional bank-belt width is 0.5 km. For narrower reaches with total centerline length L_{narrow} (km), the maximum added bank area is:

$$A_{bank,max} \approx 0.5 \times L_{narrow}$$

Similarly, within a single 3-km grid cell, if the narrow reach intersects the cell with length l (km), the worst-case bank-belt area inside the cell is:

$$A_{bank,max,cell} \approx 0.5 \times l$$

the total valid river mask (with actual width of w in km) area is:

$$A_{mask,cell} \approx 0.5 \times (w + 0.5) \times l$$

and therefore the maximum absolute per-cell RIC inflation is bounded by:

$$\Delta RIC_{max,cell} \leq \frac{0.5 \times l}{(w + 0.5) \times l} = \frac{0.5}{w + 0.5}$$

In the most pessimistic configuration, i.e., $w \in [0.25,0.5)$ over the entire reach, $\Delta RIC_{max,cell} \in (0.5,0.67]$

(3) Practical considerations

Several factors reduce the expected bias relative to the conservative bounds above:

Partial snow cover and heterogeneity: snow cover on banks is spatially heterogeneous and temporally variable, so only a fraction of the potential bank-belt area is typically snow-covered at a given time.

Imperfect snow-to-ice confusion: although NDSI cannot reliably separate snow and ice, classification errors are not systematic under all conditions.

Post-processing noise suppression: the 3-km gridding and temporal smoothing procedures suppress isolated misclassifications and reduce the influence of localized false positives.

Table S1: Comparison of mean cloud cover over six major rivers based on different satellite strategies. Mean cloud fractions over the six major river basins were compared for the entire hydrological year and the winter season (October 1 to April 30), using both the single-sensor approach (Terra or Aqua) and the final dual-sensor strategy adopted in this study.

	Terra annual	Terra winter	Aqua annual	Aqua winter	Combined annual	Combined winter
Mackenzie	0.624	0.648	0.612	0.621	0.439	0.383
Yukon	0.625	0.624	0.614	0.602	0.465	0.407
Kolyma	0.532	0.515	0.516	0.488	0.350	0.314
Lena	0.588	0.593	0.571	0.559	0.442	0.395
Yenisey	0.703	0.762	0.688	0.743	0.561	0.607
Ob	0.709	0.754	0.698	0.729	0.568	0.589
Average	**0.630**	**0.649**	**0.617**	**0.624**	**0.471**	**0.449**

Table S2: Mean phenology records (freeze-up date, breakup date, and ice duration) of six major rivers over 24-year period.

	Mean freeze-up date	Mean breakup date	Mean ice duration (day)
Mackenzie	October 20	April 24	187.45
Yukon	October 6	May 6	212.71
Kolyma	October 1	May 23	234.36
Lena	October 13	April 28	198.09
Yenisey	November 3	April 8	156.97
Ob	October 27	April 10	166.72
Average	**October 20**	**April 22**	**184.97**

Table S3: Summary of the river centerline length for the six major Arctic rivers in this study.

	Analyzed river segments length (km)	Narrower reaches length (km)	Ratio
Mackenzie	11,890.13	858.79	7.2%
Yukon	20,919.17	2,326.00	11.1%
Kolyma	19,152.67	2,081.93	10.9%

Lena	62,037.11	8,377.70	13.5%
Yenisey	57,836.47	6,377.58	11.0%
Ob	35,586.66	4,805.59	13.5%

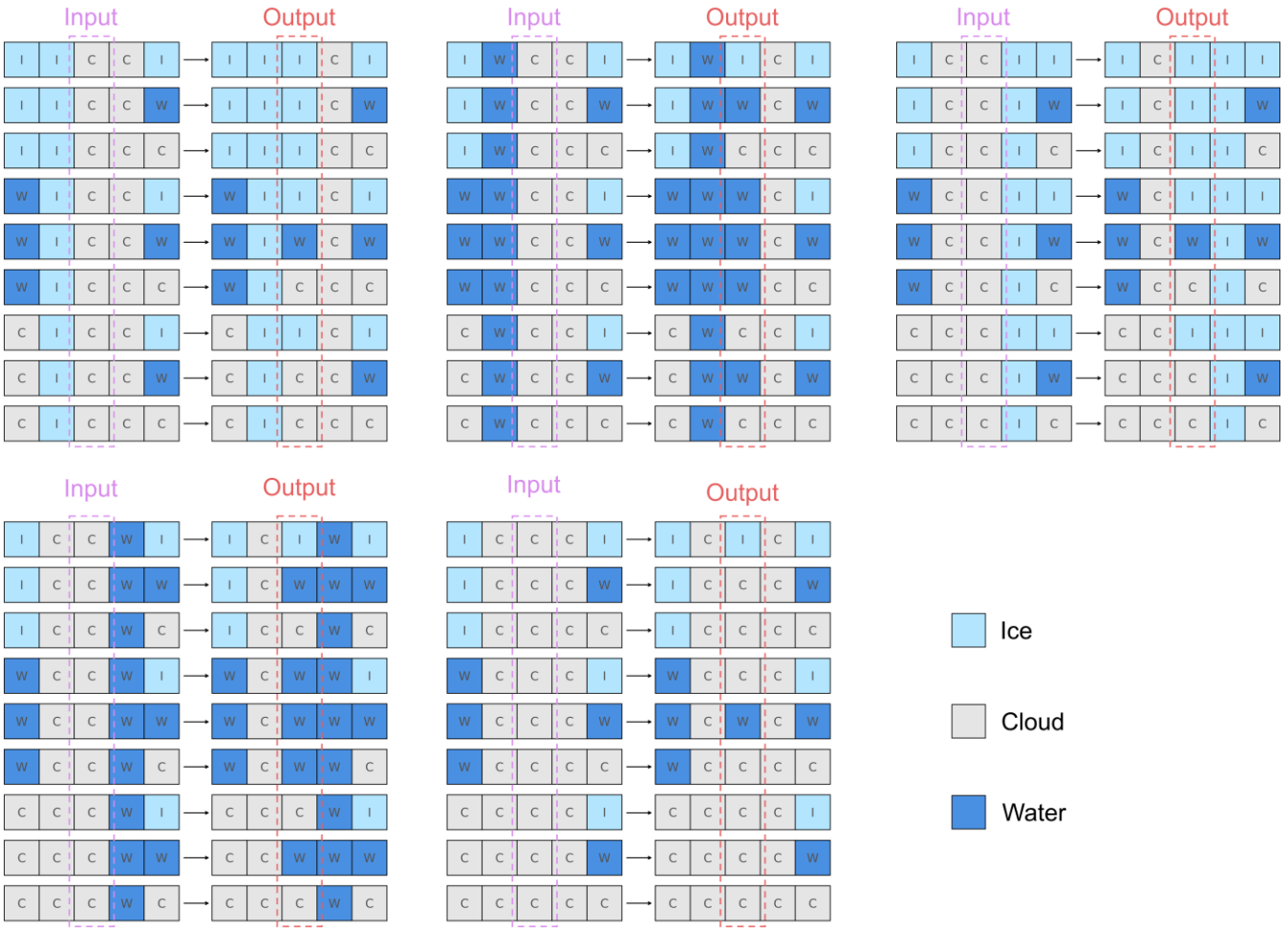

Figure S1: Logic flow for reclassifying central-day cloud pixels using a 5-day window. Of the 45 possible combinations, 13 lead to reclassification as river ice, 13 as water, and 19 remain as cloud.

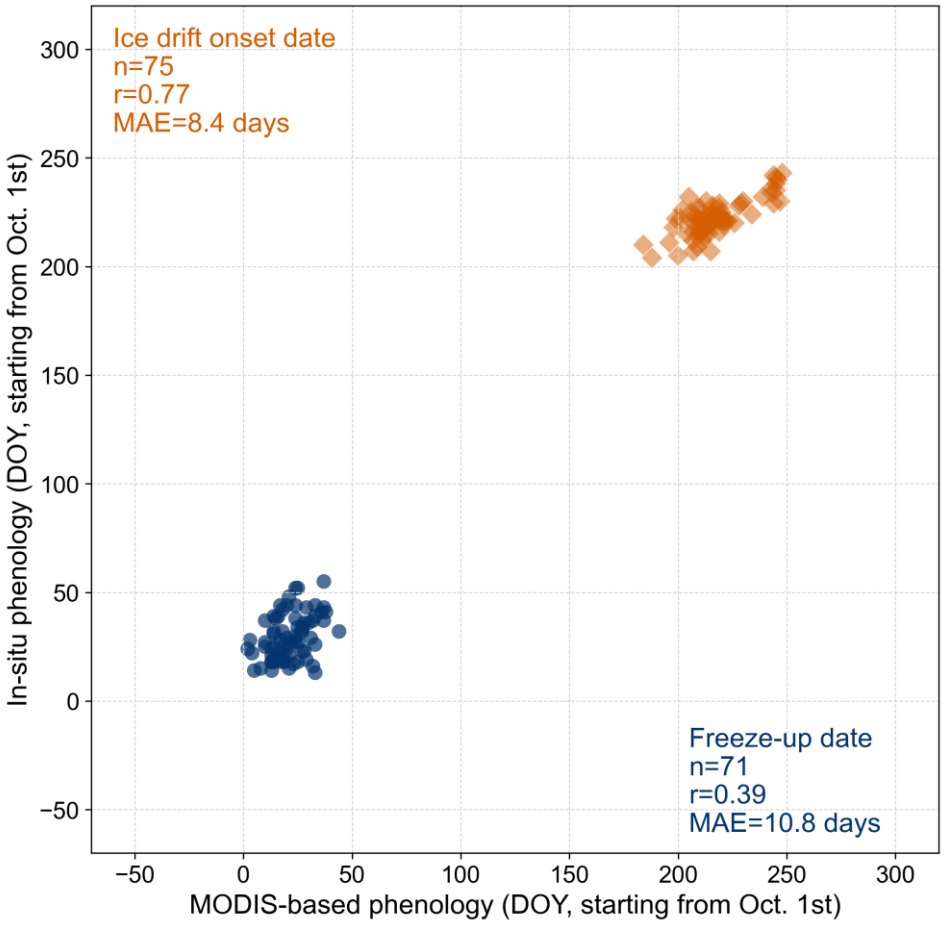

Figure S2: Comparison of MODIS-derived river ice phenology (freeze-up and breakup dates) with long-term gauge records. Satellite-based breakup events were compared with the in situ ice drift onset date.

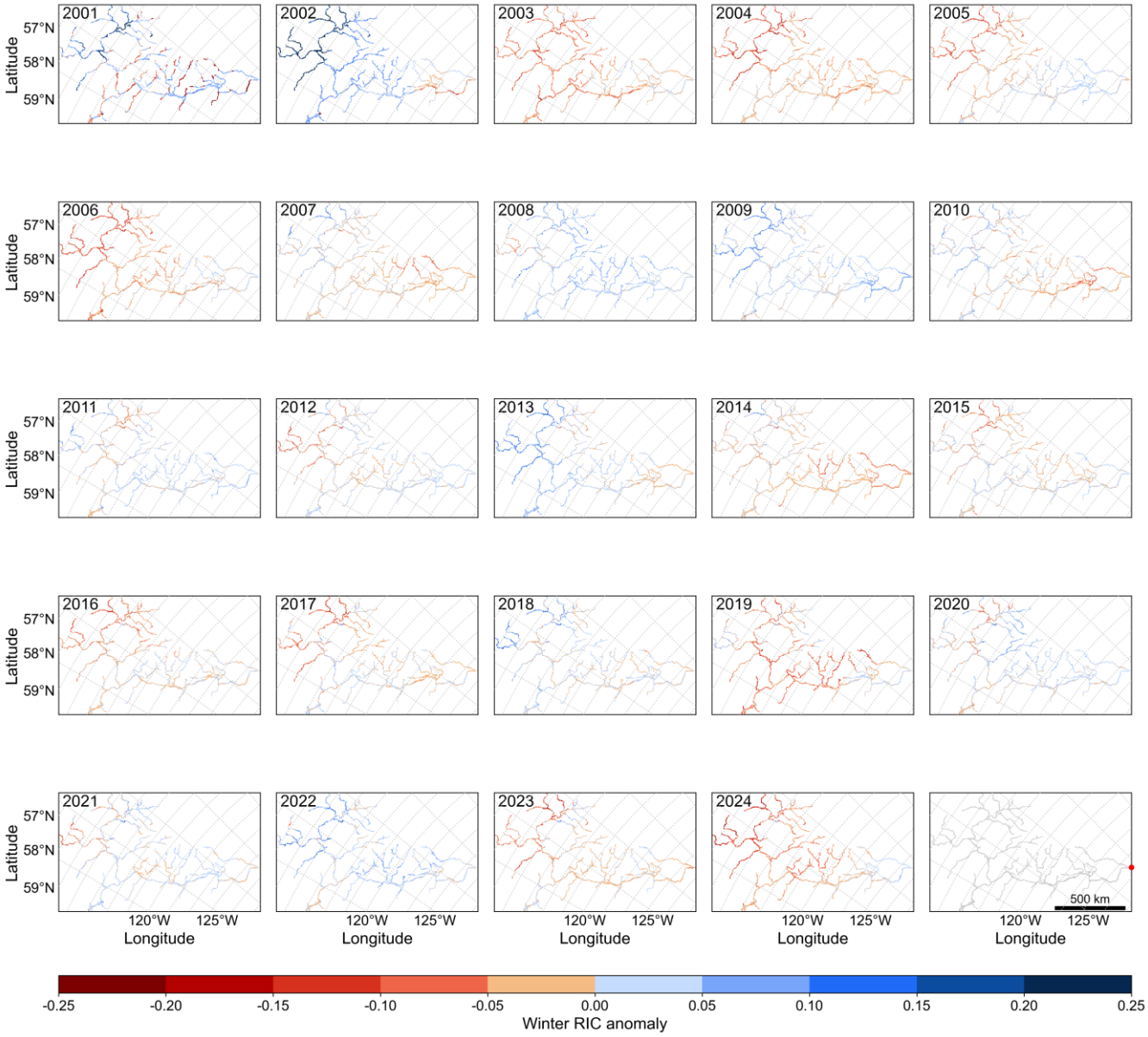

Figure S3: Maps show hydrological year winter (October–April) anomalies in MODIS-derived RIC across the Mackenzie River basin. Anomaly is defined as the winter-mean RIC for a given hydrological year minus the 24-year climatological winter mean (2001–2024) computed at each pixel. All maps are presented in the EPSG:3995 coordinate reference system.

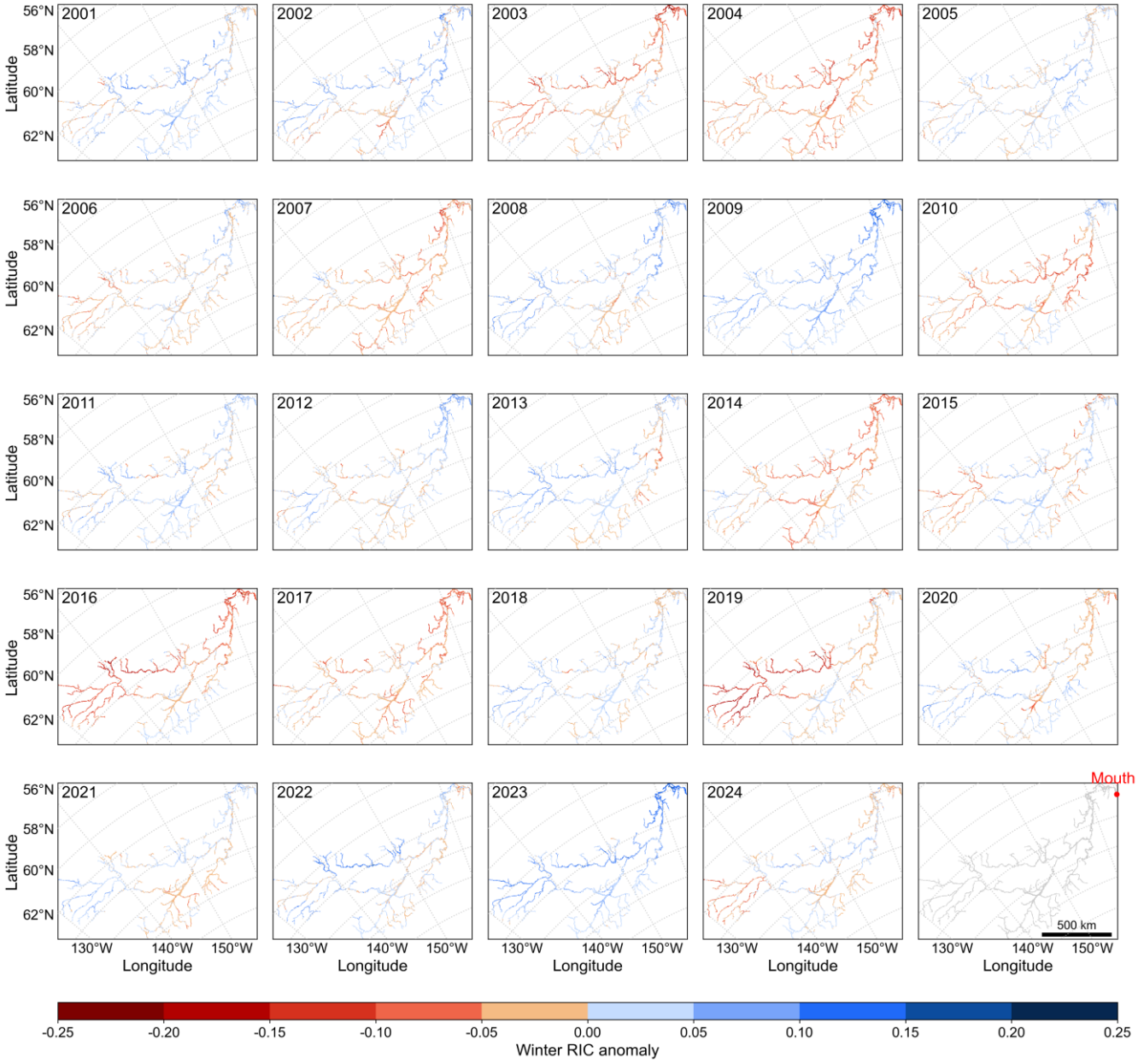

Figure S4: Maps show hydrological year winter (October–April) anomalies in MODIS-derived RIC across the Yukon River basin. Anomaly is defined as the winter-mean RIC for a given hydrological year minus the 24-year climatological winter mean (2001–2024) computed at each pixel. All maps are presented in the EPSG:3995 coordinate reference system.

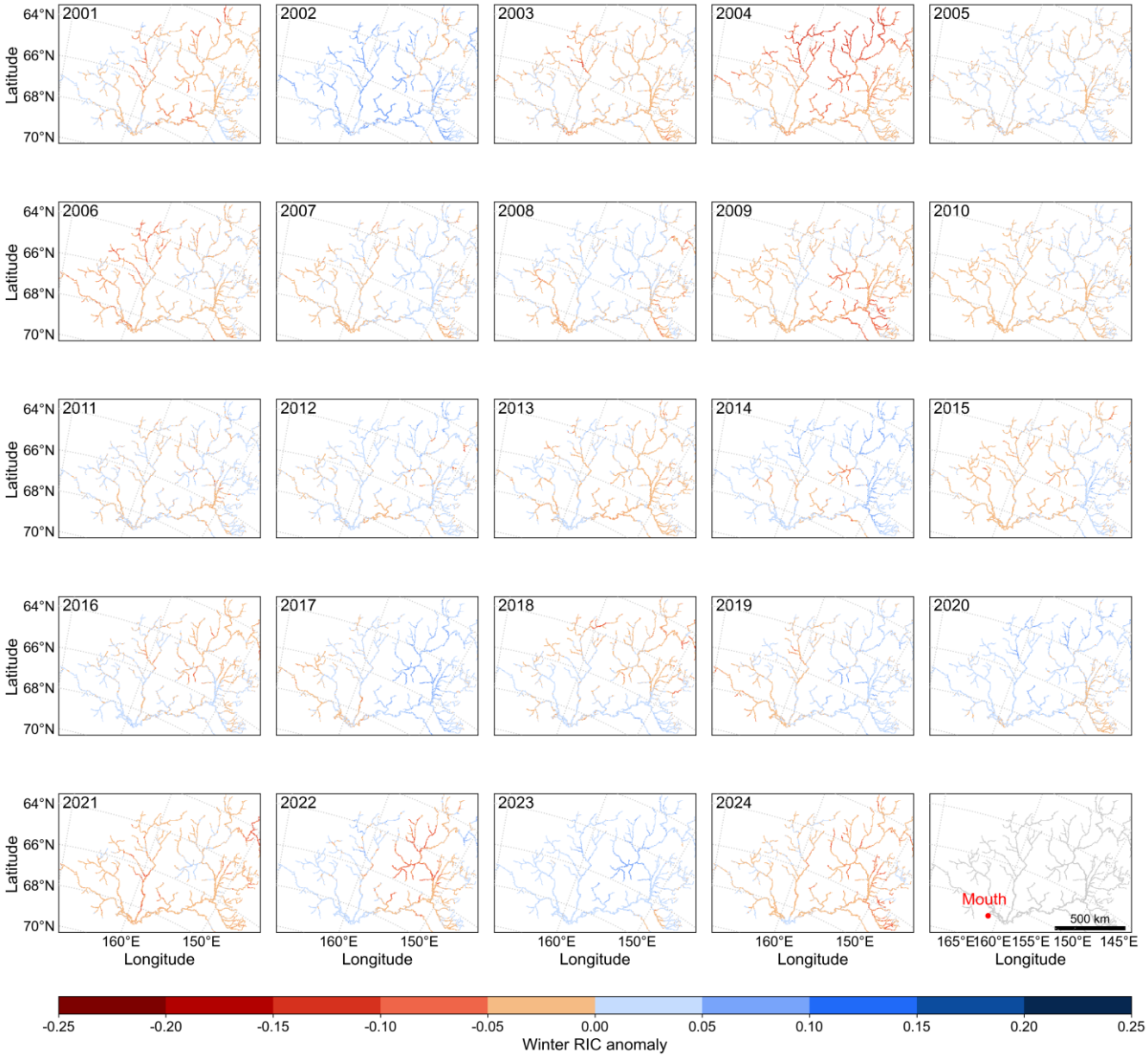

Figure S5: Maps show hydrological year winter (October–April) anomalies in MODIS-derived RIC across the Kolyma River basin. Anomaly is defined as the winter-mean RIC for a given hydrological year minus the 24-year climatological winter mean (2001–2024) computed at each pixel. All maps are presented in the EPSG:3995 coordinate reference system.

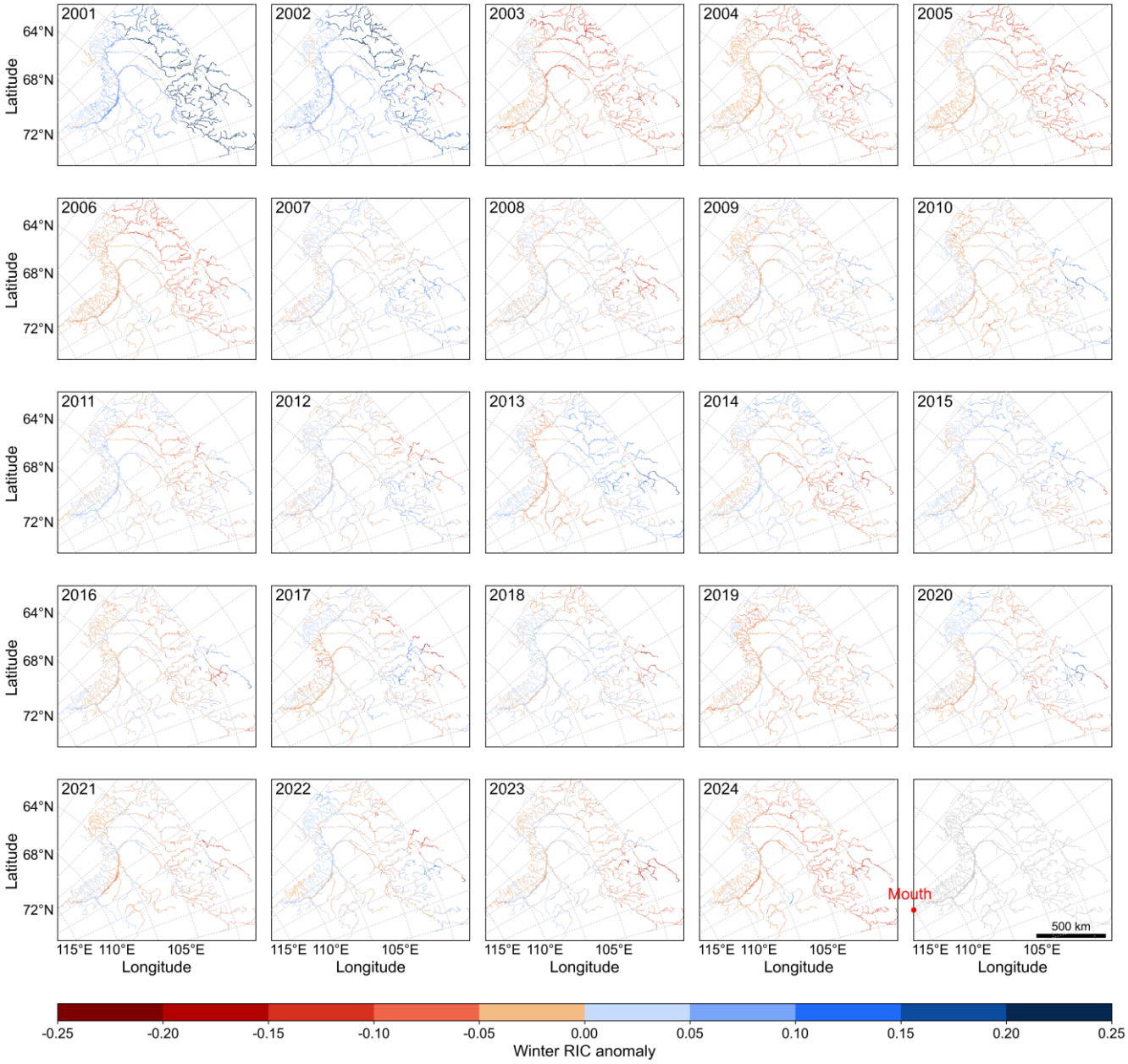

Figure S6: Maps show hydrological year winter (October–April) anomalies in MODIS-derived RIC across the Lena River basin. Anomaly is defined as the winter-mean RIC for a given hydrological year minus the 24-year climatological winter mean (2001–2024) computed at each pixel. All maps are presented in the EPSG:3995 coordinate reference system.

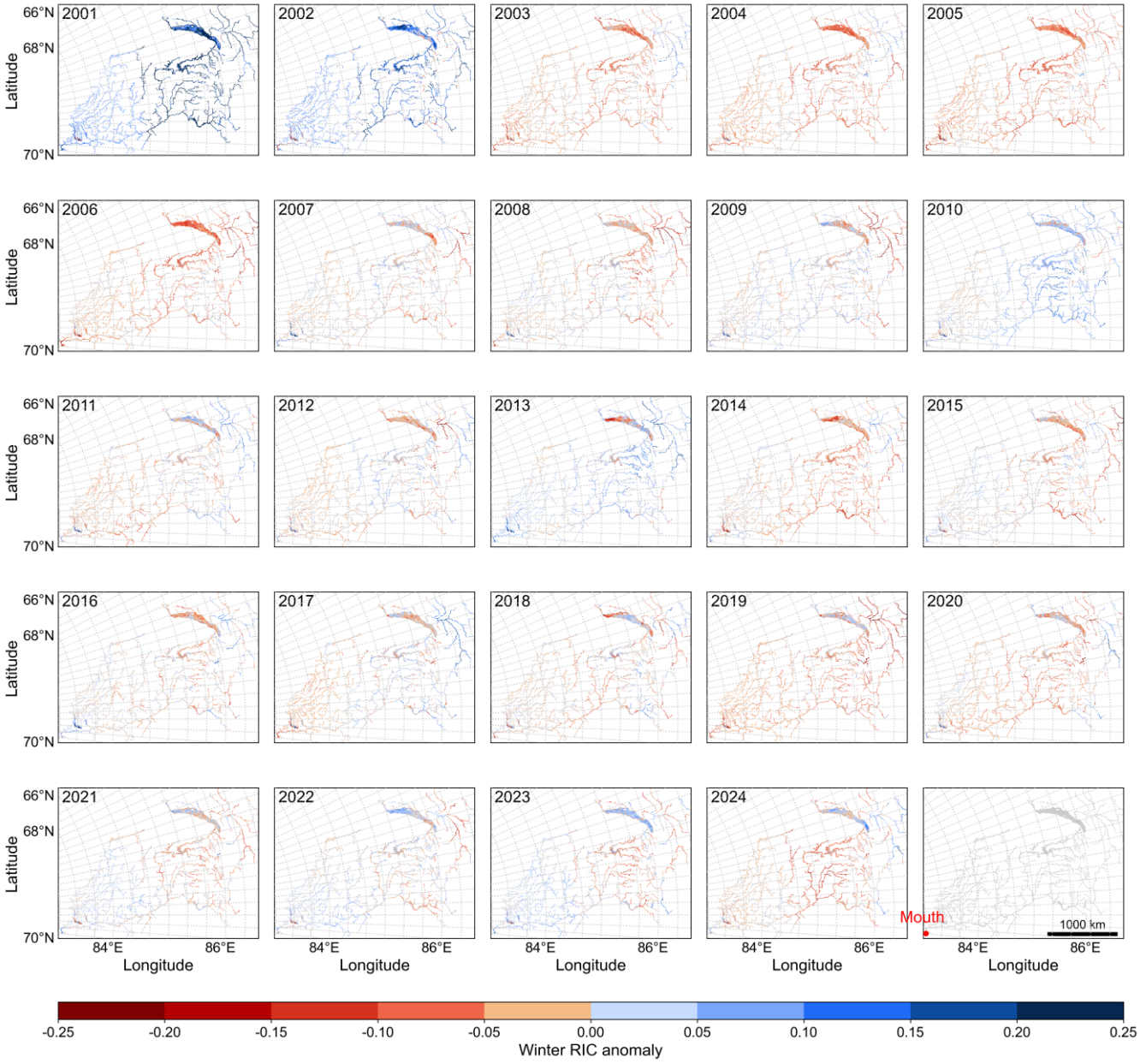

Figure S7: Maps show hydrological year winter (October–April) anomalies in MODIS-derived RIC across the Yenisey River basin. Anomaly is defined as the winter-mean RIC for a given hydrological year minus the 24-year climatological winter mean (2001–2024) computed at each pixel. All maps are presented in the EPSG:3995 coordinate reference system.

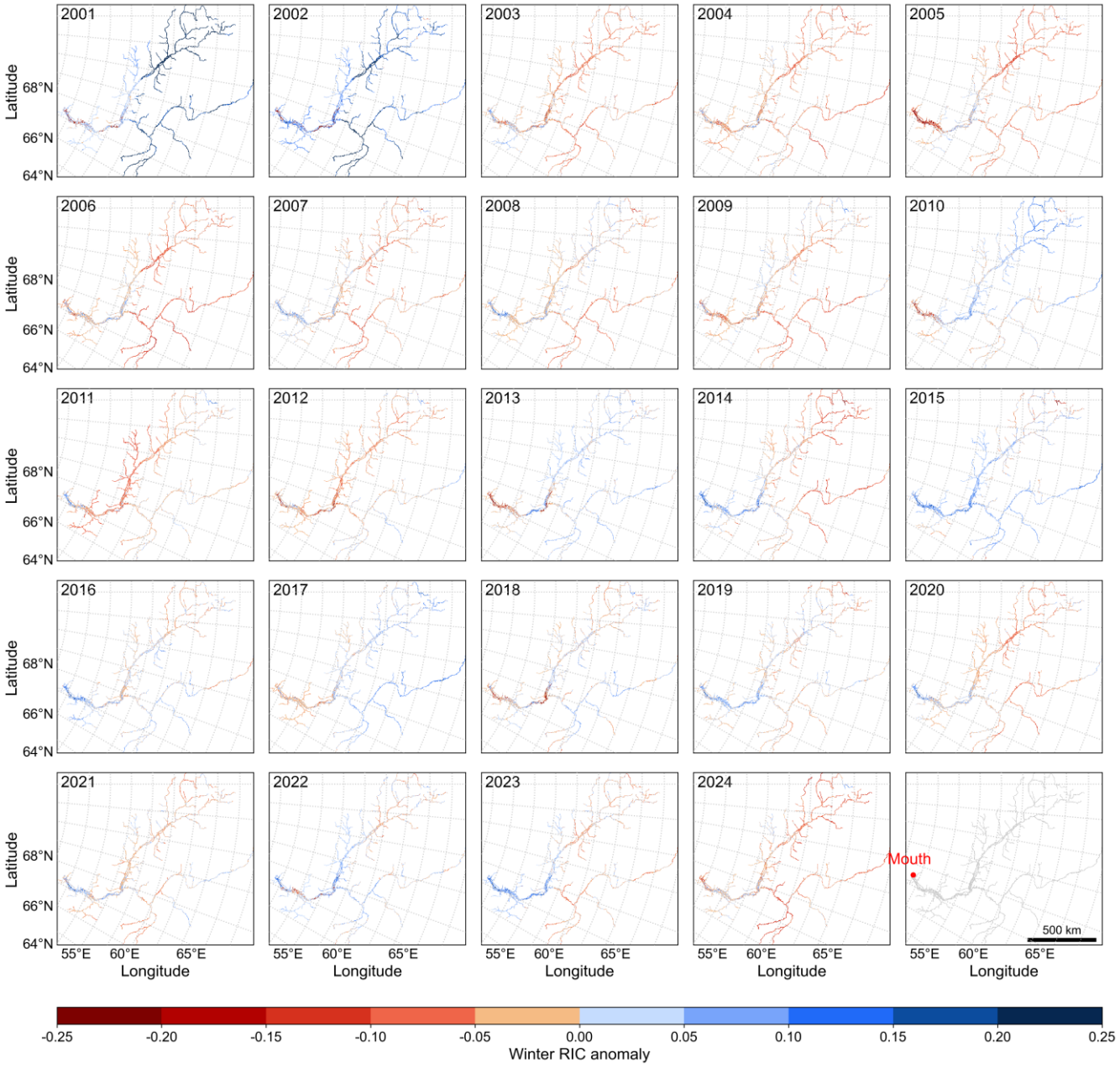

Figure S8: Maps show hydrological year winter (October–April) anomalies in MODIS-derived RIC across the Ob River basin. Anomaly is defined as the winter-mean RIC for a given hydrological year minus the 24-year climatological winter mean (2001–2024) computed at each pixel. All maps are presented in the EPSG:3995 coordinate reference system.

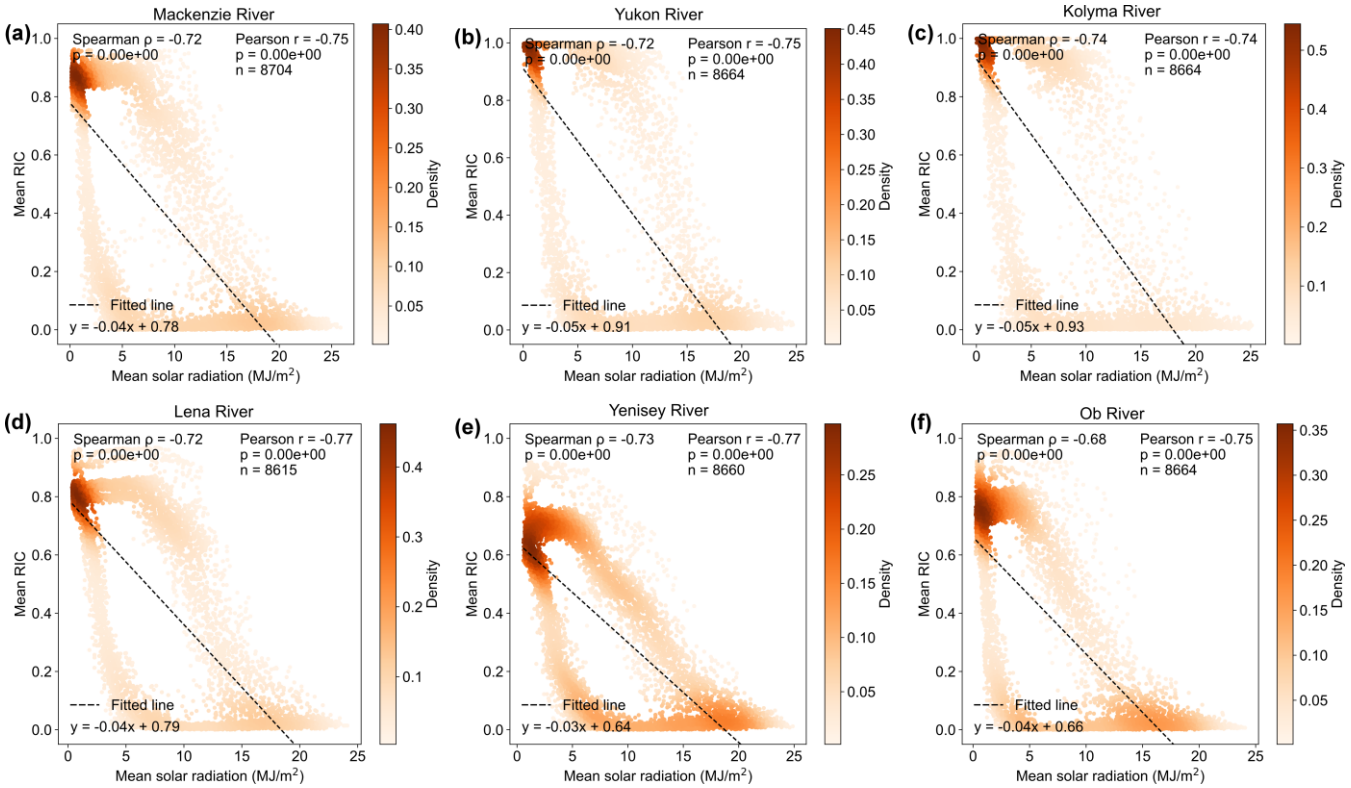

Figure S9: Correlation between MODIS-derived winter RIC and net solar radiation across six major Arctic rivers. Each point denotes the basin-mean RIC and contemporaneous net solar radiation for a given date.

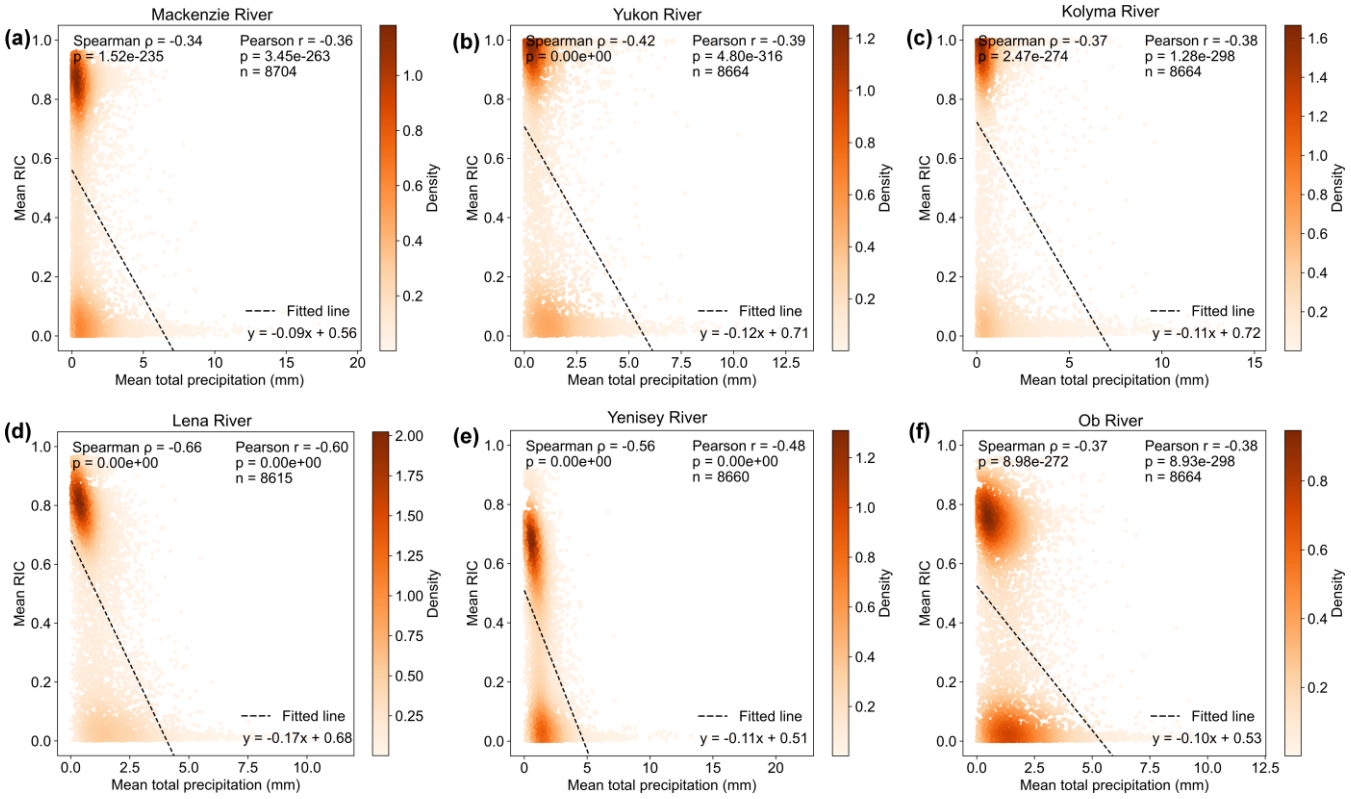

Figure S10: Correlation between MODIS-derived winter RIC and total precipitation across six major Arctic rivers. Each point denotes the basin-mean RIC and contemporaneous total precipitation for a given date.