Y. Tashiro[1,2*], M. Yoh[3], V. P. Shesterkin[4] , T. Shiraiwa[5], T. Onishi[6], D. Naito[7,8*]

[1]United Graduate School of Agricultural Science, Tokyo University of Agriculture and Technology, Tokyo 1838509, Japan, [2]Institute for Space–Earth Environmental Research, Nagoya University, Aichi 4648601, Japan, [3]Emeritus professor, Institute of Agriculture, Tokyo University of Agriculture and Technology, Tokyo 1838509, Japan, [4]Khabarovsk Federal Research Center of the Far Eastern Branch of the Russian Academy of Sciences, Institute of Water and Ecology Problems (FEB RAS), Khabarovsk 680000, Russia, [5]Institute of Low Temperature Science, Hokkaido University, Hokkaido 0600819, Japan, [6]Faculty of Applied Biological Sciences, Gifu University, Gifu 5011193, Japan, [7]Center for International Forestry Research (CIFOR), Bogor 16115, Indonesia, [8]Faculty/Graduate School of Agriculture, Kyoto University, Kyoto 6068502, Japan

*Current Affiliation

Corresponding author: Yuto Tashiro (`tassy40y@gmail.com`)

Key Points:

- We provide a landcover classification method with high resolution of 30m using Landsat-8 data and machine learning (decision tree analysis).

- Three normalized indices (for vegetation, soil, and water) and slope enabled us to identify the distribution of permafrost wetland.

- Riverine dissolved iron concentrations in the watersheds showed a clear positive correlation with the coverages of permafrost wetland.

Abstract

Dissolved iron (dFe) transported by the Amur River greatly contributes to phytoplankton growth in the Sea of Okhotsk. Nevertheless, there has been little research on the dFe source of rivers, especially in the Amur-Mid Basin which is situated in a sporadic permafrost area. In the Amur-Mid Basin, permafrost generally exists under wetlands in the flat valley, and these permafrost wetlands could be a dFe source of rivers. To asess the importance of the permafrost wetlands as a dFe source, first we made a landcover map with high resolution of 30 m using Landsat-8 data and a machine learning technique (decision tree analysis). As a result of decision tree analysis, three normalized indices (normalized diference vegetation index, normalized difference soil index, and normalized difference water index) and slope enabled us to classify landcovers into three vegetation types: wetland, forest, and grassland. Using this landcover map, we investigated the coverages of the permafrost wetland in the sampled watersheds and examined the correlation with river water chemistry (dFe, dissolved organice carbon: DOC, and electrical conductivity: EC). As a result, dFe and DOC concentrations showed a clear positive correlation (dFe: $r^2 = 0.66$, DOC: $r^2 = 0.46$) with the coverage of permafrost wetlands, while EC showed a negative

correlation with those ($r^2 = 0.45$). These findings are the first to demonstrate the direct evidence about the importance of permafrost wetlands to supply dFe and DOC to rivers in the Amur-Mid Basin.

Plain Language Summary

This study focuses on the source of dissolved iron (dFe) in the Amur River Basin because dFe is known to contribute greatly to phytoplankton growth in the Sea of Okhotsk. Nevertheless, there has been little research on the dFe source of rivers, especially in the Amur-Mid Basin which is situated in a sporadic permafrost area. To gain a better understanding of dFe source in the Amur-Mid Basin, we made a landcover map with high resolution of 30 m using Landsat-8 data and a machine learning technique (decision tree analysis). Then we investigated the coverages of permafrost wetland in the sampled watersheds and found that the larger the coverage of permafrost wetlands, the higher was dFe concentration in the rivers. This is the first study to demonstrate the direct evidence about the importance of permafrost wetlands to supply dFe and DOC to rivers in the Amur-Mid Basin.

1 Introduction

Iron is an essential trace element for the growth of all organisms. It plays an important role in in vivo metabolic processes, including photosynthesis, electron transfer, nitrate reduction, and nitrogen fixation (Crichton, 2001; Sunda, 2012). Martin and Fitzwater (1988) first showed the possibility that iron limits phytoplankton growth in a high-nutrient low-chlorophyll (HNLC) region where the phytoplankton productivity is low despite the abundance of nutrients. Subsequent studies confirmed that dissolved iron (dFe) concentrations in many ocean areas are too low for phytoplankton to fully utilize nutrients (Bruland & Lohan, 2003; Martin et al., 1989, 1990, 1994; Price et al., 1994; Takeda & Obata, 1995). In the past, the main iron source for the ocean was thought to be aeolian dust (Martin & Fitzwater, 1988). However, recent studies have suggested that riverine dFe is also an important source to support primary production in coastal areas and the ocean (Laglera & Vandenberg, 2009; Matsunaga et al., 1998; Moore & Braucher, 2008; Nishioka et al., 2014), stimulating interest in iron dynamics in terrestrial environments.

The Sea of Okhotsk is known as one of the oceans with the richest marine resources in the world. Some studies indicated that abundant dFe derived from forests and wetlands in the Amur River Basin contributes to the creation of high biological productivity in the Sea of Okhotsk (Nishioka et al., 2014; Shiraiwa, 2012; Suzuki et al., 2014). For example, Sanjiang Plain, a large wetland in northeastern China, is an important dFe source for the Amur River, and relatively high dFe concentrations were observed in the Songhua River and the Ussuri River that pass through the Sanjiang Plain (Wang et al., 2012). In addition, wetlands in the Amur-Lower Basin from Khabarovsk to Nikolaevsk-na-Amure also greatly contribute to supply dFe to the Amur River through tributaries (Nagao et al., 2007). Therefore, wide areas of Amur-Lower Basin including

the Sanjiang Plain are believed to be a particularly important dFe source for the Amur River. Some studies reported, however, that relatively high dFe concentration was also observed in the Bureya River and the Zeya River which are the representative large rivers in the Amur-Mid Basin where permafrost is sporadically distributed (Nagao et al., 2007; Shamov et al., 2014) (Figure 1a). Based on the authors' recent field survey on permafrost distribution in the Bureya River Basin, permafrost existence was generally confirmed under wetlands (peat bogs) in flat valleys (Tashiro et al., 2020). Accordingly, these permafrost wetlands can be the most important dFe source for permafrost-affected rivers in the Amur-Mid Basin.

Since permafrost constrains the path of water to the surface of the active layer (a soil layer that thaws during summer) rich in organic matter, permafrost coverage in a watershed is associated with river water chemistry (Olefeldt et al., 2014; Petrone et al., 2006). In the study in Caribou–Poker Creeks of Alaska, Petrone et al. (2006) found that the permafrost-dominated watershed transports more dissolved organic carbon (DOC) and fewer mineral components (Ca^{2+}, Mg^{2+}, K^+, and Na^+) than the permafrost-poor watershed. It is thus important to determine the permafrost coverage to understand biogeochemical cycles from land to river in the arctic regions and to predict the influence of permafrost degradation on river water chemistry under a warming climate.

Examining the permafrost distribution over a wide area, however, is not at all easy. Instead, latitude has often been used to explain the regional difference in chemical compositions of rivers and soil pore waters (Kawahigashi et al., 2004; Pokrovsky et al., 2015, 2016; Raudina et al., 2017, 2018). Pokrovsky et al. (2016), for example, investigated DOC concentration in tributaries of the Ob River and revealed that riverine DOC concentration decreased northward with change in permafrost regime from discontinuous (50–90% coverage) to continuous (>90%) permafrost. However, even within the same regions where the permafrost regime is equal, riverine DOC concentration has a large variation that cannot be explained by permafrost regime or latitude (Olefeldt et al., 2014; Pokrovsky et al., 2015). To better understand the biogeochemical cycles in permafrost regions, grasping the detailed information about permafrost distribution in a given study area is also needed in addition to water sampling.
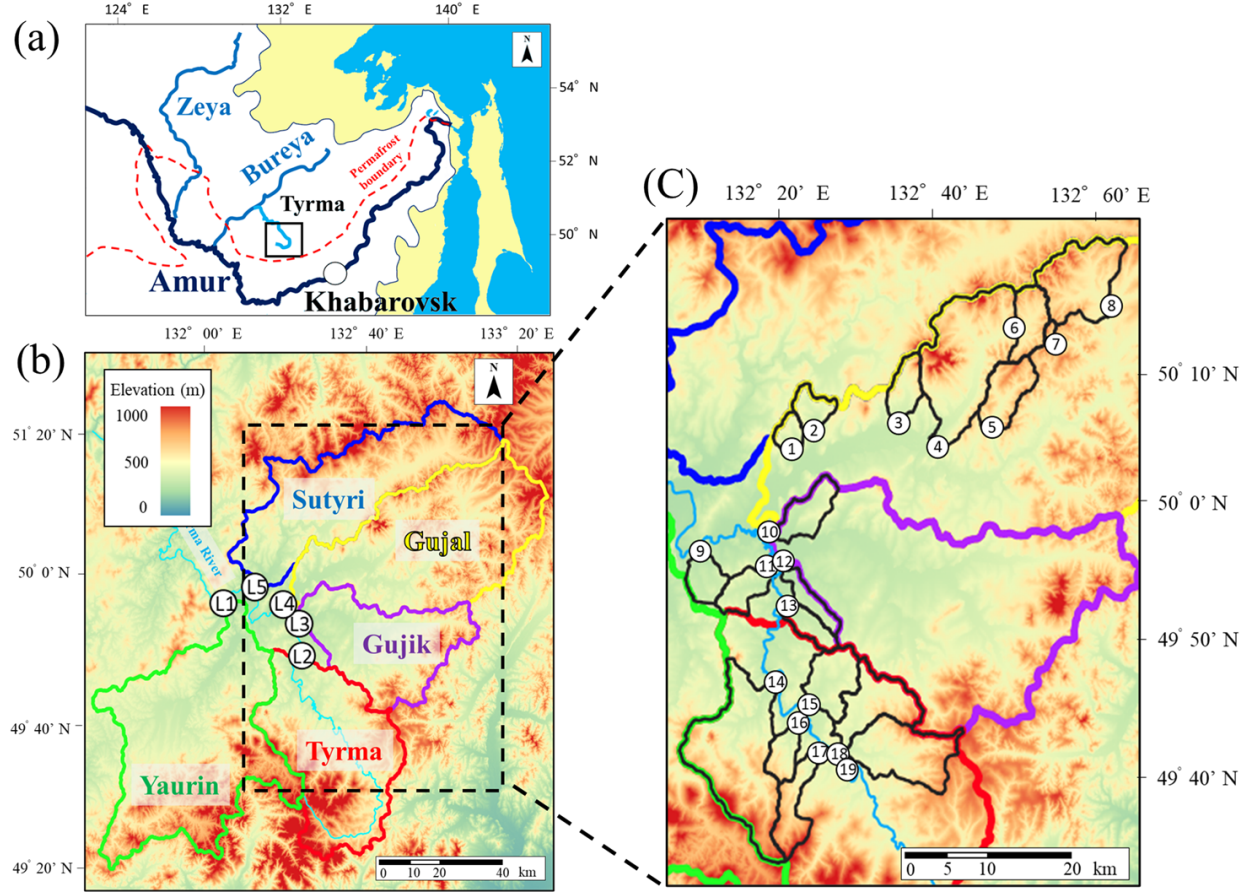
The objectives of this study are: (1) to understand the permafrost distribution on a regional scale utilizing a landcover map based on local survey and remote sensing technique, and (2) to clarify the importance of permafrost wetlands as a dFe source for rivers by investigating the relationship between the coverage of permafrost wetland and dFe concentration. This study provides not only river water quality data in the Amur-Mid Basin, but also shows the usefulness of applying a remote sensing technique to understanding biogeochemical processes in arctic and subarctic regions.

2 Materials and Methods

2.1 Study site

A field survey was conducted in the Tyrma region, which is approximately 270 km northwest of Khabarovsk in the Russian Far East (Figure 1a). Mean annual air temperature is $-1.96℃$ and annual precipitation is 654.6 mm. The Tyrma region is located just north of the permafrost boundary and is situated in a sporadic permafrost area (Obu et al., 2019; Shamov et al., 2014). The sporadic permafrost area mostly covers the Bureya River Basin and the Zeya River Basin which are the large tributaries of the Amur River (Figure 1a). In this paper, Amur-Mid Basin indicates the Bureya River Basin and the Zeya River Basin. In the Tyrma region, four large rivers, namely, the Yaurin River, the Gujik River, the Gujal River, and the Sutyri River join the Tyrma River and eventually join the Bureya River (Figure 1b).

Vegetation in the Tyrma region is roughly divided into two types: wetlands in the flat valleys are characterized by shrubs, such as bog blueberry (*Vaccinium uliginosum*), cowberry (*Vaccinium vitis-idaea L.*), and ledum (*Ledum decumbens*) and scattered larches (*Larix gmelinii var. gmelinii*) (Figure 2a & Figure S1); and the forest on the ridges and hillslopes are characterized by spruces (*Picea ajanensis*) and white birches (*Betula platyphylla*) (Figure 2b). Topsoil layers in the wetlands and the forests are composed of peat soils. In particular, thick peat soil layers are formed in the wetlands due to long-term accumulation of sphagnum biomass. This type of wetland, called "*Mari*" is a typical landscape in the flat valleys of the Tyrma region. Even more important is the fact that permafrost generally exists underneath a thick peat soil layer in wetlands (Tashiro et al., 2020). In this paper, the word "wetland" or "permafrost wetland" indicates *Mari.* In addition to the forests and wetlands, grasslands are often found along large rivers (Figure 3c). Grasslands cover flat areas of land (floodplain) near large rivers and are characterized by herbaceous plants spreading over sediment deposits.

Figure 1. Map of the study region. (a) Location of the Tyrma region. Permafrost boundary was drawn with reference to Shamov et al. (2014). (b) Elevation map of the Tyrma region. Areas enclosed by colored lines and numbers L1–L5 shows the catchment areas and the sampling sites of large river waters. (c) Detailed elevation map for the sampled small watersheds. Areas enclosed by black lines and numbers 1–19 show the catchment areas and the sampling sites of small river waters. Maps (b) and (c) were created by the authors using Quantum GIS based on a 30 m resolution digital elevation model (DEM) provided by Japan Aerospace Exploration Agency (JAXA).

Figure 2. Typical vegetation type in the Tyrma region. (a) wetland (*Mari*), (b) forest, (c) grassland. Other photos about the vegetation of wetland (*Mari*) are shown in Figure S1.

2.2 Calculation of normalized difference indices

A flowchart of the landcover classification process is shown in Figure 3. Landsat-8 data (Level2, Collection2, Tier1) with a spatial resolution of 30 m was used for landcover classification. Importing and processing of these data were conducted in Google Earth Engine which is a platform for scientific analysis and visualization of geospatial datasets (Gorelick et al., 2017). First, Landsat-8 data from 2013 to 2021 were extracted. Then, the data in the summer months (June, July, and August: JJA) were extracted. Finally, medians of these JJA data were extracted to calculate normalized difference indices as described in the next paragraph.
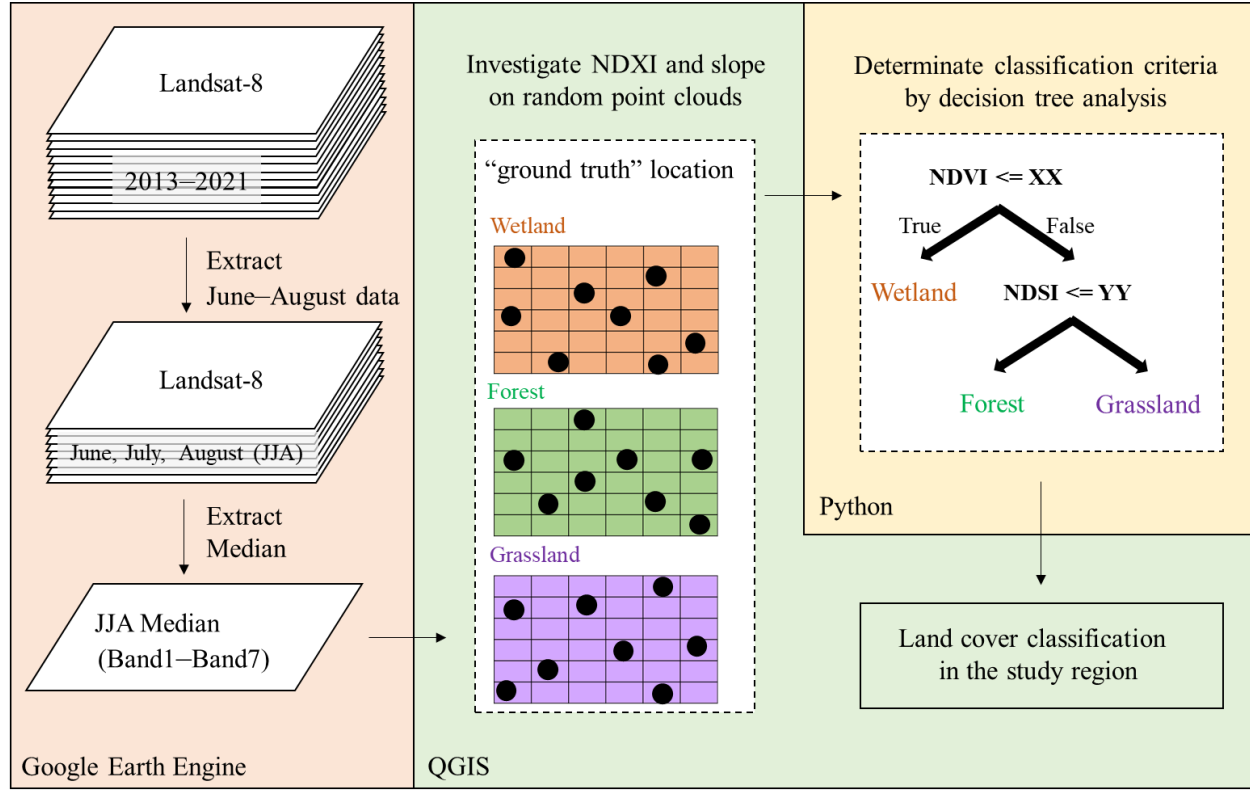
For classifying landcover by satellite image analysis, it is important to presume several surface conditions such as vegetation, soil, and water. Thus, we calculated three indices using JJA-median Landsat-8 data: normalized difference vegetation index (NDVI), normalized difference soil index (NDSI), and normalized difference water index (NDWI). The formulas to calculate the three indices are as follows:

$$\text{NDVI} = (\text{Band5}_{\text{NIR}} - \text{Band4}_{\text{Red}}) / (\text{Band5}_{\text{NIR}} + \text{Band4}_{\text{Red}}) \qquad (1)$$

$$\text{NDSI} = (\text{Band6}_{\text{SWIR1}} - \text{Band5}_{\text{NIR}}) / (\text{Band6}_{\text{SWIR1}} + \text{Band5}_{\text{NIR}}) \qquad (2)$$

$$\text{NDWI} = (\text{Band5}_{\text{NIR}} - \text{Band6}_{\text{SWIR1}}) / (\text{Band5}_{\text{NIR}} + \text{Band6}_{\text{SWIR1}}). \qquad (3)$$

Here, each band of data was JJA-median value. Band names and wavelength of Landsat-8 are shown on the following website (https://www.usgs.gov/faqs/what-are-band-designations-landsat-satellites). NDVI, NDSI, and NDWI show different ranges depending on vegetation activity, soil types, and water content of leaves (Gao, 1996; Takeuchi and Yasuoka, 2004; Tucker et al., 1985). In this paper, NDVI, NDSI, and NDWI are collectively described as NDXI. The code for extracting JJA-median Landsat-8 data and calculating NDXI in Google Earth Engine is available in the Supporting Information (Figure S2).

Figure 3. Schematic flowchart of several processes for landcover classification. The detailed explanations of the processes in Google Earth Engine, QGIS, and Python are described in subsections 2.2, 2.3, and 2.4, respectively.

2.3 Investigation of NDXI and slope across different vegetation types

Investigation of NDXI in different vegetation types was conducted by GIS software Quantum GIS (QGIS version 3.20.0). In this study, we attempted to classify landcover in the Tyrma region into three vegetation types: wetland, forest, and grassland (Figure 2). First, 30 random point clouds were generated in a area where these vegetations are dominated. The random point cloud is a tool to sample a certain number of data in a given area, and we can specify the minimum distance between points (Figure S3). Here, we specified 30 m as a minimum distance to avoid generating points on the same grid of Landsat-8 data with 30 m resolution. Most importantly, the areas where the random point clouds were generated are locations that we confirmed the actual vegetation type (ground truth) at the study site. The random point clouds were generated in three ground truth areas for each vegetation type, that is, a total of 270 points (3 vegetation types × 3 ground truth areas × 30 point clouds). Then, the values of NDXI calculated using JJA-median Landsat-8 data were investigated for all 270 points. In addition to NDXI, the degree of slope at all 270 points was also

investigated. The slope data were created using a 30 m resolution digital elevation model (DEM) provided by Japan Aerospace Exploration Agency (JAXA). The coordinates, the values of NDXI, and the degree of slope for all 270 points are summarized in Table S1.

2.4 Determination of classification criteria by decision tree algorithm

To classify landcovers based on NDXI and slope, the criteria were determined by supervised machine learning. Here, we utilized decision tree analysis in Python (version 3.8.5). A decision tree is an algorithm that classifies data gradually based on generated rules and outputs a tree-like graph. Of the data in all 270 points, 30% were used as test data, and the other 70% were used as learning data. This data classification was done in three stages. The obtained criteria as a result of the decision tree analysis were extrapolated to the whole Tyrma region, and landcover was classified into wetland (*Mari*), forest, and grassland. The code for the decision tree analysis in Python is available in Figure S4.

2.5 Sampling of river waters

Water samples were collected in July 2019. Sampling of the Tyrma Main River was conducted just before the confluence with the Gujik River, and the sampling of other large rivers (the Yaurin River, the Gujik River, the Gujal River, and the Sutyri River) was conducted just before the confluence with the Tyrma River (Figure 1b). In addition, water samples were also collected in 19 small rivers: 8 rivers in the Gujal River system and 19 rivers in the Tyrma River system (Figure 1c). Two hundred milliliters of water was sampled using a disposable syringe (TERUMO, SS-50ESZ) and immediately filtered through 0.45 µm disposable filters made of cellulose acetate (ADVANTEC, DISMIC 25CS045AS). One hundred milliliters of the filtered water was preserved in an acid-washed propylene bottle for dFe measurement and the other 100 mL was preserved in a propylene bottle for DOC measurement. Both samples were kept in a refrigerator until analysis. Also, electrical conductivity (EC) was measured using a portable EC meter (ES-71, HORIBA) at the time of water sampling.

2.6 Chemical analyses and statistics

dFe concentration was determined by the 1,10-phenanthroline method (Russian international technical standards 52.24.358-2006: https://files.stroyinf.ru/Index2/1/4293837/4293837319.htm). Here, we describe this method briefly. First, 1 mL of 10% hydroxylammonium chloride was added to 50 mL of the sample. Second, this was boiled for 15–20 min until the volume reached 25 mL to separate organic iron complexes into organic compounds and Fe ( ). Third, after cooling, ammonium hydroxide was added until ~pH4. Fourth, 3 mL of ammonium acetate buffer and 1 mL of 1,10-phenanthroline were added, and ultrapure water was added until the volume reached 50 mL. Finally, 20 min after the color development, the absorbance at a wavelength of 510 nm was measured with an ultraviolet–visible spectrophotometer (SHIMADZU UV mini-1240). In this paper, we define dFe as Fe that was determined by this process. The detection limit for dFe by the 1,10-phenanthroline method was 0.02 mg L^{-1}. DOC concentration was deter-
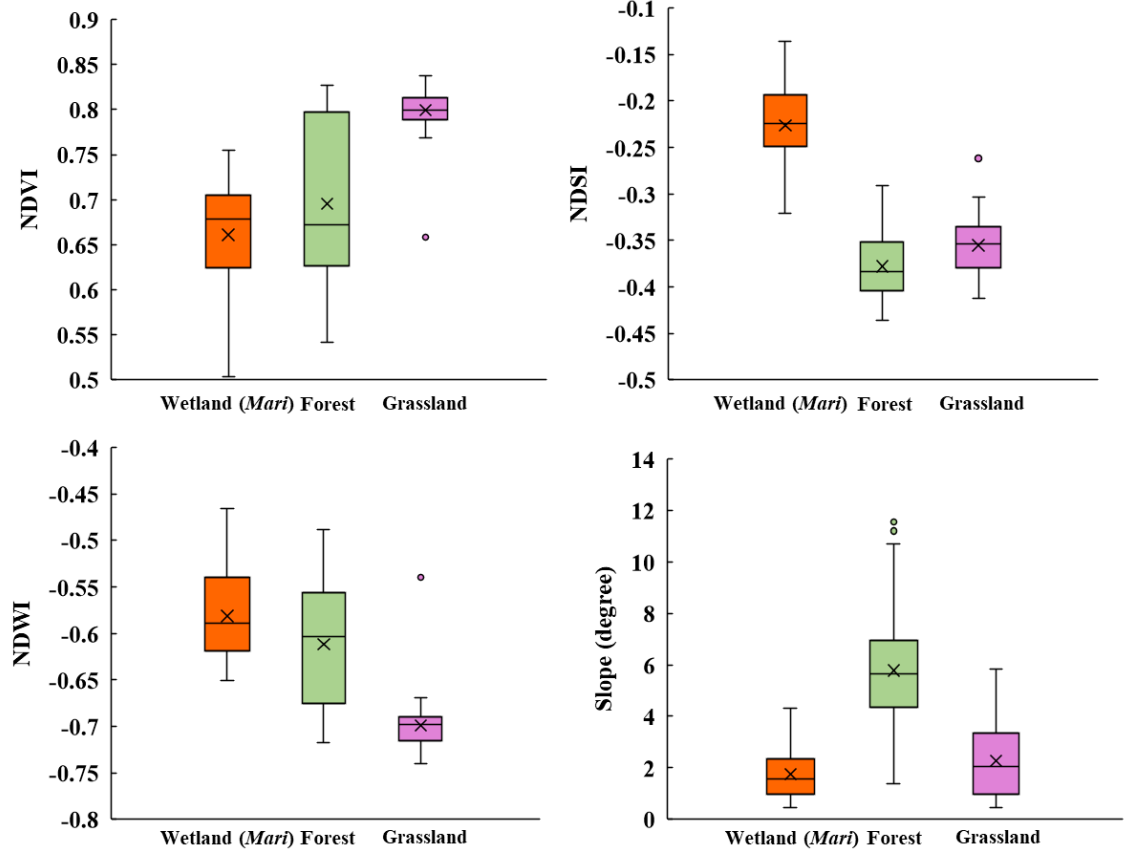
mined with a total organic carbon (TOC) analyzer (SHIMADZU TOC-LCSH) using the catalytic combustion oxidation method. The detection limit for TOC by the TOC analyzer was 0.1 mg L^{-1}, and standard solutions for DOC analysis were prepared using Potassium Hydrogen Phthalate (C6H4(COOK)(COOH)) (Nacalai tesque).

Based on the produced landcover map (subsections 2.2–2.4), the coverage of wetland (*Mari*) was investigated for each catchment area of 5 large rivers and 19 small rivers. The correlation of water chemistry (dFe, DOC, and EC) and the coverage of wetland (*Mari*) was assessed using liner regression analysis and non-liner regression analysis. For non-liner regression analysis, three common functions (power, exponential, and logarithmic) were investigated to create a approximation curve. Both liner and non-liner regression calculations were performed by least-square method with Microsoft Excel Solver (version 2021). The approximation line or curve with the highest coefficient of determination (r^2) was selected as the most suitable regression equation to represent the coefficient between water chemistry and the wetland coverage. Note that coefficient of determination (r^2) and Pearson's correlation coefficient (r) for each approximation curve were calculated by fitting liner regression model for the log-transformed data.

3 Results

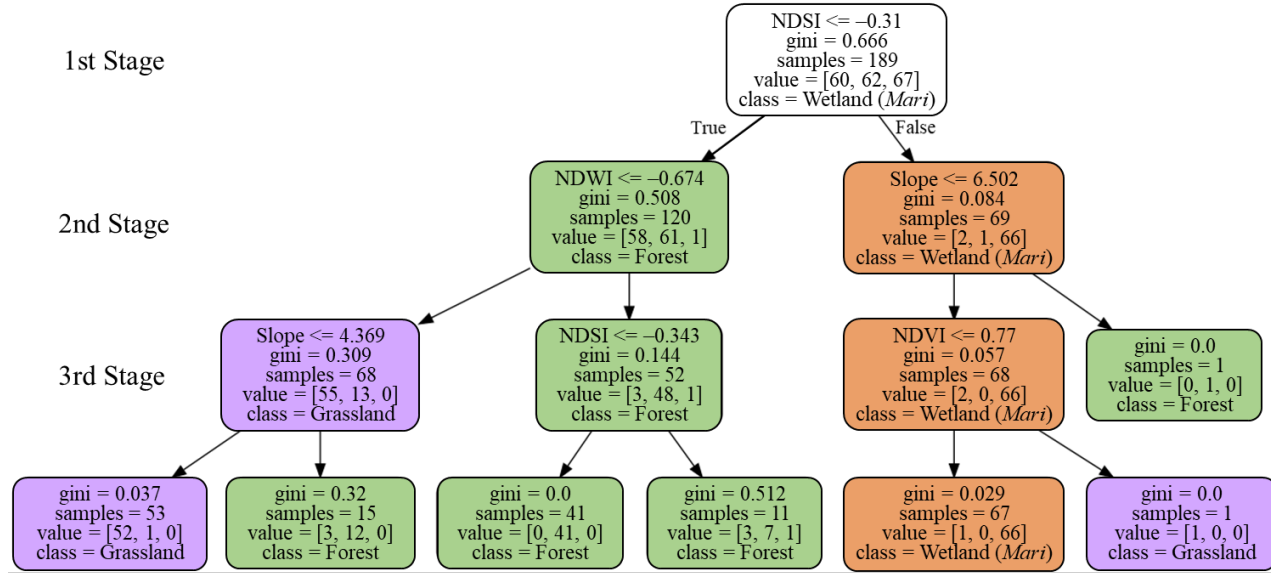3.1 Landcover classification by decision tree analysis based on NDXI and slope

Ranges of NDXI and slope on the point clouds in wetland (*Mari*), forest, and grassland are shown in Figure 4. Here we focus attention on the differences of the NDXI and slope in the wetland compared with those in the forest and the grassland. In the wetland, NDVI was in the range of 0.50–0.75, NDSI was –0.32 to –0.14, and NDWI was –0.65 to –0.47. NDVI and NDSI in the wetland largely overlapped with those of the forest, indicating that NDVI and NDWI were not useful in distinguishing between wetland and forest. On the other hand, NDSI in the wetland was clearly higher than that in the forest and the grassland. The range of slope in the wetland was quite low at 0.43–4.31 degrees. Compared with this, the forest clearly showed a higher range, but the grassland showed almost the same range as the wetland; accordingly, wetland cannot be distinguished from grassland just by the slope. From these findings, NDSI seems to be the most effective index for identifying the distribution of wetland (*Mari*).

Figure 4. Box plot showing the comparison of NDXI and slope between wetland (*Mari*), forest, and grassland. Lower end of the box denotes 25th percentile, center line denotes 50th percentile (exclusive median), and upper end denotes 75th percentile. Cross mark in the box denotes the average. Bar denotes the largest and smallest values within 1.5 times interquartile range. Outside values denotes >1.5 times the interquartile range. n = 90 for each box plot. The underlying data for this figure are available in Table S1.
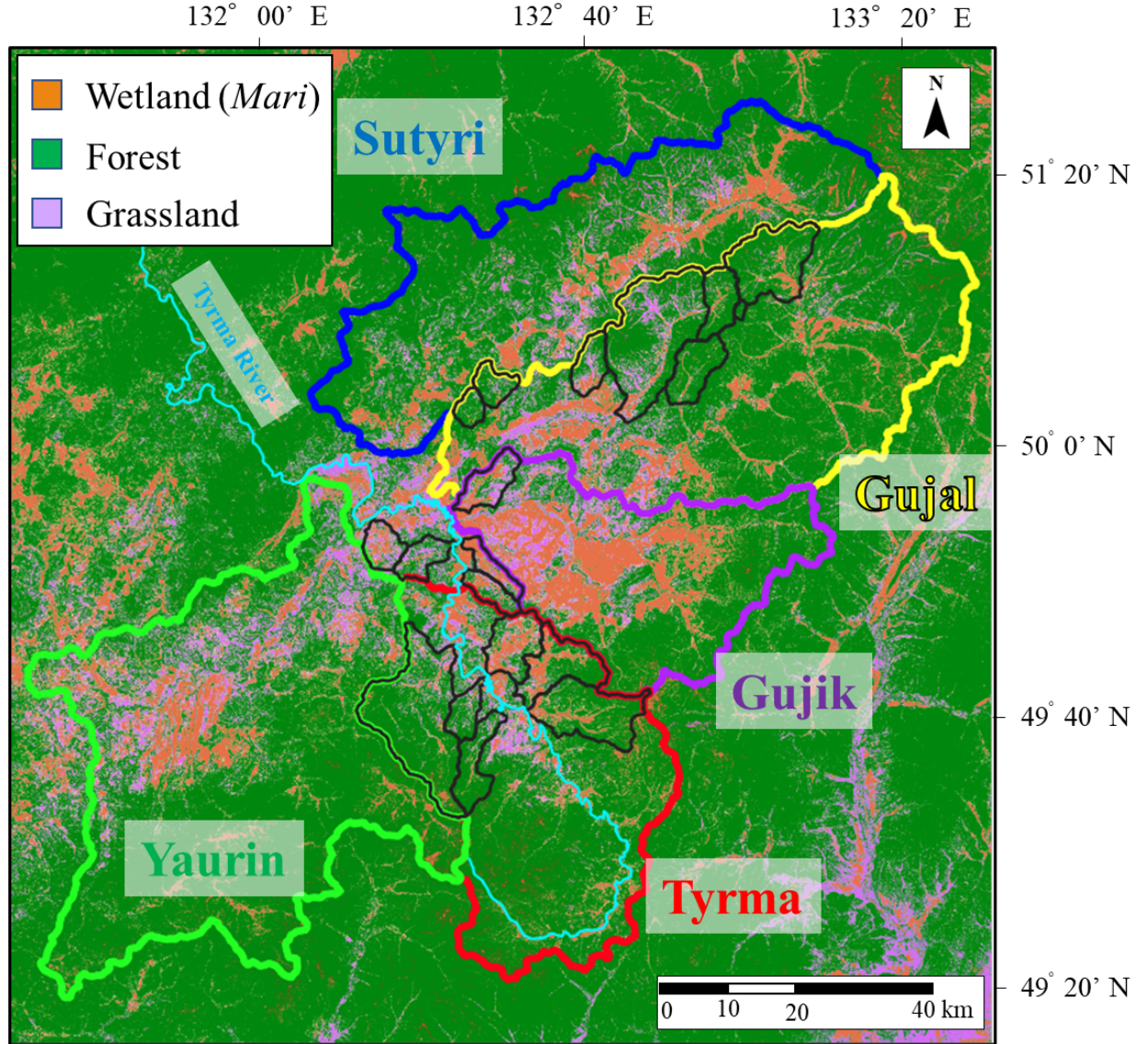
A graph of the decision tree analysis is shown in Figure 5. The accuracy rate of this model for test data was 92.6% (Figure S4). Accuracy rate is generally used to evaluate a prediction precision for unknown data and is calculated by dividing the number of correct classification for test data by total number of test data (30% of 270 samples: 81). As shown in Figure 4, NDSI was selected as a criterion at the first stage of decision tree analysis. Based on NDSI –0.31, most samples of the forest and grassland were classified into the true direction, while almost all samples of the wetland (*Mari*) were classified into the false direction. This result indicates that wetland can be distinguished from forest and grassland, but only by NDSI. On the other hand, many samples of forest

and grassland remained in the box of true direction after the first stage. Then, NDWI –0.674 was used as a criterion to distinguish forest and grassland: most samples of grassland were classified into the true direction, while most samples of forest were classified into the false direction. Finally, at the third stage, a small number of samples were classified by slope, NDSI, and NDVI. From these results, it follows that three-stage classification by the decision tree analysis based on NDXI and slope was enough to distinguish landcover as wetland (*Mari*), forest, and grassland.

5. Output graph of the decision tree analysis. The contents in each box are: top part is the classification criterion; gini denotes Gini coefficient; samples denote the total number of samples for classification 189 in the box at the 1st stage means 70% of the total samples (270), which is used as learning data; value denotes the number of samples, from left to right, Grassland, Forest, and Wetland (*Mari*); class denotes the greatest share of classified samples. Box colors (orange, green, and purple) corresponds with the class (wetland, forest, and grassland).

The decision-tree's classification model was extrapolated for the Tyrma region, and the produced landcover map is shown in Figure 6. Comparing this map with the terrain map (Figure 1b), we can see that wetland (*Mari*) covers in the valley areas widely. Across the five large river basins, wetland is particularly distributed in the Gujik River Basin. Grassland also seems to cover some of the valley areas, but the area of grassland is apparently smaller than that of wetland. Mountain slopes, the areas from valley edge to ridge, were mostly covered with forest.

Figure 6. Landcover map in the Tyrma region as a result of extrapolating the decision-tree's classification model. Orange areas denote the places classified into wetland (*Mari*). Similarly, green areas denote forest, and purple areas denote grassland. Areas enclosed by color and black lines show the catchment areas of the sampled rivers.

3.2 River water chemistry (dFe, DOC, EC) and their relationship with the coverage percentage of wetland (*Mari*) in the watersheds.

Using the produced landcover map (Figure 6), the coverage percentage of wet-

land (*Mari*) in the sampled river watersheds was calculated. This result is shown in Table 1 with the results of dFe, DOC, and EC. The range of dFe concentration in the five large rivers was 0.12–0.38 mg L^{-1}, with dFe being especially high in the Yaurin and Gujik Rivers. The range of DOC concentration was 13.4–19.0 mg L^{-1}, and similar to dFe, DOC concentration was higher in the Yaurin and Gujik rivers than in other large rivers. EC in the large rivers was low on average and its range was 3.90–5.63 mS m^{-1}. The range of wetland coverage was relatively high in the Gujik Basin at 34.8%, but there was no significant difference among other large rivers of 14.7–16.5%. Alternatively, the small rivers showed more variety of water chemistry compared with the large rivers: they showed quite wide ranges for dFe concentration of 0.02–0.54 mg L^{-1} and DOC concentration of 7.4–29.5 mg L^{-1}. EC in the small rivers was also in a wide range of 2.89–14.00 mS m^{-1}. The coverage of wetland in the small river watersheds was 1.5–55.9%, especially higher in the rivers of the Tyrma River system (Rivers 9–19 in Table1) than those of the Gujal River system (Rivers 1–8 in Table1).

Table 1

List of sampled rivers, watershed area, wetland coverage, and water chemistry

Number	Large river	Watershed (km²)	Wetland (*Mari*) (%)	dFe (mg L^{-1})	DOC (mg L^{-1})	EC (mS m^{-1})
L1	Yaurin					
L2	Tyrma					
L3	Gujik					
L4	Gujal					
L5	Sutyri					
Number	Small river	Watershed (km²)	Wetland (*Mari*) (%)	dFe (mg L^{-1})	DOC (mg L^{-1})	EC (mS m^{-1})
	Sofron					
	Yakagulin					
	No					
	Name					
	Kevity					

Table 1

List of sampled rivers, watershed area, wetland coverage, and water chemistry

	Jangsovo	
	Kevity (upstream)	
	Kevity-makit	<0.02
	Sula	
	No Name	
	Allan	
	Talkanji	
	Kuvitikan	
	Kuvitikan	
	Verkhny Kovun	
	Right Talanja	
	Talanja	
	Tokchka-birakan	
	Small Nigba	
	Big Nigba	

The relationships between the river water chemistry and the wetland coverage are shown in Figure 7. Riverine dFe and DOC concentrations clearly increased with the wetland coverage (dFe: $r^2 = 0.66$, DOC: $r^2 = 0.46$). In contrast, riverine EC exponentially decreased with wetland coverage (EC: $r^2 = 0.45$). In addition to these findings, we found that riverine dFe concentration had a strong positive correlation with DOC concentration ($r^2 = 0.63$), and a negative correlation with EC ($r^2 = 0.74$). Since EC is an index showing how much dissolved ion components are present in the water, these results strongly suggest that wetland (*Mari*) supplies dFe- and DOC-rich and mineral-poor groundwater to rivers (see Discussion 4.2).

Figure 7. Variations in (a) dFe concentration, (b) DOC concentration, and (c) EC with an increase in the coverage of wetland (*Mari*). Figure panels on the right-hand side shows variations in dFe concentration with increases in (d) EC and (e) DOC concentration. As written in the subsection 2.6, the approximation line or curve on the each graph is the most suitable regression equation among the calculated liner function and non-liner functions (power, exponential, and logarithmic).

4 Discussion

16

4.1 Permafrost distribution in the Tyrma region and comparison of dFe and DOC concentrations in other similar permafrost-regime regions.

Since the produced landcover map in this study (Figure 6) was based on NDXI and slope in the ground truth areas, the areas determined to be wetland (*Mari*) most likely have the same vegetation as wetland (*Mari*). Moreover, according to the authors' previous work and field experience, permafrost existence underneath the wetland (*Mari*) was confirmed in many areas of the Tyrma region (Tashiro et al., 2020). These facts allow us to expect that permafrost is distributed with a fairly high possibility under the wetland-covered areas shown in the landcover map (Figure 6). Given this perspective and the calculated wetland coverage in the large river basins (Table 1), the permafrost coverage percentage in the Tyrma region can be estimated at around 15-35%. This coverage corresponds with the estimated permafrost regime (sporadic permafrost distribution: 10–50%) by Obu et al. (2019) using a hemispheric-scale model.

In the Ob River, flows north and west across the western Siberian lowland, there have been some studies about river water chemistry in the sporadic permafrost areas (around N60°–N62°). Pokrovsky et al. (2016) showed that riverine dFe concentrations (<0.45 µm like this study) in these regions in summer were in the range of 0.5–3.5 mg L^{-1}, and Vorobyev et al. (2017) showed that they were 0.2 to 1.8 mg L^{-1}, respectively. Compared with these data in the Ob River Basin, riverine dFe concentrations (<0.02–0.540 mg L^{-1}) in the Tyrma regions were relatively low. However, riverine DOC concentrations in the Tyrma region and those in sporadic permafrost regions in the Ob River Basin are almost in the same range: 7.4–29.5 mg L^{-1} in the Tyrma region and 6–33 mg L^{-1} in the Ob region (Pokrovsky et al., 2015; Vorobyev et al., 2017) . One of the reasons for the difference in dFe concentration among the Tyrma region and the Ob region (N60°–N62°) despite the same permafrost regime is likely soil profile in the active layer. Since soil thawing in summer allows water to interact with mineral soil horizon under the peat soil layer, this can increase dFe discharge into rivers as known in Siberian watersheds with sporadic or less permafrost distribution (Bagard et al., 2011; Pokrovsky et al., 2016; Vorobyev et al., 2017). In the Tyrma region, thick peat soil layers from the surface to near the permafrost table are formed in the wetland (Tashiro et al., 2020); therefore, water interaction with mineral soils may be restricted in the flow path from the permafrost wetlands (*Mari*) to rivers in summer.

4.2 Role of permafrost wetland (*Mari*) in supplying dFe and DOC to rivers

Our fine-scale landcover map enabled us to understand the distribution of permafrost wetland (*Mari*) in the Tyrma region and detect a relationship between its coverage and water chemistry. As shown in Figure 7, river water chemistry was greatly influenced by the coverage of wetland (*Mari*): dFe and DOC concentrations increased, but EC decreased with an increase in the wetland coverage. These results agree well with the findings in northern Sweden and northwestern Canada where permafrost is present under peatlands (Olefeldt et al., 2013, 2014). It is widely accepted that whether water predominantly passes through the peat

soil layer or the mineral soil layer is important in determing the river water chemistry. For example, DOC discharge generally reduces during transport through the mineral soil layer because of adsorption on clay minerals (Kothawala et al., 2012; Smedberg et al., 2006), which may also reduce the discharge of Fe-organic complexes to rivers. According to the soil profiles in the Tyrma region reported by Tashiro et al. (2020), peat soil accumulates more than 40 cm from the surface to the permafrost table in the wetlands (*Mari*), whereas peat soil occupies only 7–12 cm of the surface in the forests. Based on this fact and our findings, it follows that river water chemistry in the Tyrma region is greatly regulated by the combination of two water resources: the peat soil layer in the wetlands (*Mari*) and the mineral soil layer in the forests. Moreover, this study found a strong positive correlation between riverine dFe and DOC concentrations in the Tyrma region (Figure 7d), suggesting that wetland(*Mari*)-derived dFe comprises mainly Fe-organic complexes which are stable in river waters with neutral pH and can be transported to the ocean (Tipping, 2002). This is consistent with Levshina (2012) which investigated the ratio of Fe bounded with humic acids to riverine dFe concentration in the Amur-Mid Basin, and also with other previous studies in the boreal regions (Björkvald et al., 2008; Ingri et al., 2006).

The observed dFe concentration in the large rivers in the Tyrma region was 0.12–0.38 mg L^{-1} (Table 1), which is as high as the dFe concentration observed in July in the Bureya River and the Zeya River which are the representative large rivers in the Amur-Mid Basin (Nagao et al., 2007). Given the considerable water discharge of such large rivers, there will be no doubt that great quantities of dFe are supplied from Amur-Mid Basin to the Amur Main River, although riverine dFe concentration varies seasonally (Tashiro et al., 2020). According to the previous study and the existent wetland map in the Amur River Basin, wetlands are widely distributed in the Bureya River Basin and Zeya River Basin (Egidarev & Simonov, 2007; Egidarev et al., 2016). There is unfortunately little information about wetland types and permafrost distribution in these regions, it is thus difficult to understand how important the wetland (*Mari*) is as a dFe source for the whole of Amur-Mid Basin. From our findings, however, it should be emphasized that wetland (*Mari*) is major dFe source for rivers in the Tyrma region and can be important dFe source for rivers in other regions of the Amur-Mid Basin where permafrost is sporadically distributed. In addition, this study warns the possibility of change in riverine dFe concentration in the Amur-Mid Basin due to permafrost degradation under a warming climate. Recent studies pointed out that permafrost degradation will influence on biogeochemical cycle of iron because of change in redox conditions and water flow path induced by increase in active layer thickness (Patzner et al., 2022; Pokrovsky et al., 2016). In the Amur River Basin, the evidence of permafrost degradation was recently found by Winterfeld et al. (2018) using $\Delta14C$, and it may be more serious near the permafrost boundary like the Tyrma region. Considering the great contribution of wetland-derived dFe to the marine ecosystem in the Sea of Okhotsk (Nishioka et al., 2014; Shiraiwa, 2012; Suzuki et al., 2014), we will need to pay careful attention to the influence of permafrost degradation on riverine

dFe concentration in the Amur-Mid Basin.

5 Conclusions

To assess the importance of permafrost wetlands, called *Mari*, as a dFe source for rivers, we made a landcover map with fine resolution (30 m) using Landsat-8 data and a machine learning technique (decision tree analysis). As a result, this study clearly demonstrated that river water chemistry in the Tyrma region was greatly influenced by the coverage of wetland (*Mari*): dFe and DOC concentrations increased, but EC decreased with an increase in the wetland coverage in the watershed. To our knowledge, this study is the first to discuss the role of permafrost wetlands (*Mari*) in river water chemistry and to show the direct evidence of the importance of permafrost wetlands (*Mari*) for riverine dFe and DOC concentrations in the Amur-Mid Basin. Given that permafrost degradation is predicted to occur due to ongoing climate change, further research will be required to assess the influence of permafrost degradation on the watershed hydrological cycle, iron dynamics, and riverine dFe cocentration in the Amur-Mid Basin because this issue may have the potential to change the amount of dFe discharged to the Sea of Okhotsk.

Open Research

Additional information is available in the Supporting Information.

Figure S1. Photos of wetland (*Mari*) in the Tyrma region.

Figure S2. Code for extracting JJA-median Landsat-8 data and calculating NDXI in Google Earth Engine

Figure S3. Sample display of random point clouds generation in QGIS.

Figure S4. Code for decision tree analysis by Python

Table S1. Underlying data for Figure 4 including the coordinates, the values of NDXI, and slope data on 270 point clouds.

References

Bagard, M. L., Chabaux, F., Pokrovsky, O. S., Viers, J., Prokushkin, A. S., Stille, P. et al. (2011). Seasonal variability of element fluxes

in two Central Siberian rivers draining high latitude permafrost dominated areas. *Geochimica et Cosmochimica Acta, 75*(12), 3335–3357. https://doi.org/10.1016/j.gca.2011.03.024

Bruland, K. W. & Lohan, M. C. (2003). 6.02 – Controls of trace metals in seawater: *Treatise on Geochemistry, 6,* 23–47, https://doi.org/10.1016/B0-08-043751-6/06105-3

Björkvald, L., Buffam, I., Laudon, H., & Mörth, C. M. (2008). Hydrogeochemistry of Fe and Mn in small boreal streams: The role of seasonality, landscape type and scale. *Geochimica et Cosmochimica Acta, 72(*12*),* 2789–2804. https://doi.org/10.1016/j.gca.2008.03.024

Crichton, R. (Ed.). (2001). *Inorganic biogeochemistry of iron metabolism: from molecular mechanisms to clinical consequences*, Hoboken, NJ: John Wiley.

Egidarev, E. & Simonov, E. (2007). *Wetlands, water infrastructure.* WWF Russia, Amur Branch, Vladivostok, Russia. http://amur-heilong.net/Gis_site/gis_index.html

Egidarev, E., Simonov, E., & Darman, Y. (2016). Amur-Heilong River Basin: Overview of Wetland Resources. In Finlayson, C., Milton, G., Prentice, R., & Davidson N. (Eds.), *The Wetland Book*(pp. 1–15). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-6173-5_7-2

Gao, B. (1996). NDWI A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment, 58,* 257–266. https://doi.org/10.24059/olj.v23i3.1546

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202,* 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Ingri, J., Malinovsky, D., Rodushkin, I., & Baxter, D. C. (2006). Iron isotope fractionation in river colloidal matter. *Earth and Planetary Science Letters, 245,* 792–798. https://doi.org/10.1016/j.epsl.2006.03.031

Kawahigashi, M., Kaiser, K., Kalbitz, K., Rodionov, A., & Guggenberger, G. (2004). Dissolved organic matter in small streams along a gradient from discontinuous to continuous permafrost. *Global Change Biology, 10(*9*),* 1576–1586. https://doi.org/10.1111/j.1365-2486.2004.00827.x

Kothawala, D. N., von Wachenfeldt, E., Koehler, B., & Tranvik, L. J. (2012). Selective loss and preservation of lake water dissolved organic matter fluorescence during long-term dark incubations. *Science of the Total Environment, 433,* 238–246. https://doi.org/10.1016/j.scitotenv.2012.06.029

Laglera, L. M., & Vandenberg, C. M. G. (2009). Evidence for geochemical control of iron by humic substances in seawater. *Limnology and Ocanography, 54(*2*),* 610–619. https://doi.org/10.4319/lo.2009.54.2.0610

Levshina, S. I. (2012). Iron distribution in surface waters in the Middle and Lower Amur basin. *Water Resources, 39(4),* 375–383. https://doi.org/10.1134/s0097807812040082

Martin, J. H., & Fitzwater, S. E. (1988). Iron deficiency limits phytoplankton growth in the North-east Pacific subarctic. *Nature, 336,* 403–405. https://doi.org/10.1038/331341a0

Martin, J. H., Gordon, R. M., Fitzwater, S. E., & Broenkow, W. W. (1989). Vertex: phytoplankton/iron studies in the Gulf of Alaska. *Deep Sea Research Part A, 36,* 649–680. https://doi.org/10.1016/0198-0149(89)90144-1

Martin, J. H., Gordon, R., & Fitzwater, S. E. (1990) Iron in Antarctic waters. *Nature, 345,* 156–158. https://doi.org/10.1038/345156a0

Martin, J. H., Coale, K. H., Johnson, K. S., Fitzwater, S. E., Gordon, R. M., Tanner, S. J. et al. (1994). Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. *Nature, 371,* 123–129. https://doi.org/10.1038/371123a0

Matsunaga, K., Nishioka, J., Kuma, K., Toya, K., & Suzuki, Y. (1998). Riverine input of bioavailable iron supporting phytoplankton growth in Kesennuma Bay (Japan). *Water Research, 32,* 3436–3442. https://doi.org/10.1016/S0043-1354(98)00113-4

Moore, J. K., & Braucher, O. (2008). Sedimentary and mineral dust sources of dissolved iron to the world ocean. *Biogeosciences, 5(3),* 631–656. https://doi.org/10.5194/bg-5-631-2008

Nagao, S., Motoki, T., Kodama H., Kim V. I., Shesterkin P. V., & Makhinov A. N. (2007). *Migration behavior of Fe in the Amur River Basin* (Report on Amur - Okhotsk Project No. 4, pp. 37–48.). Kyoto, Japan: Research Institute for Humanity and Nature.

Nishioka, J., Nakatsuka, T., Ono, K., Volkov, Y. N., Scherbinin, A., & Shiraiwa, T. (2014). Quantitative evaluation of iron transport processes in the Sea of Okhotsk. *Progress in Oceanography, 126,* 180–193. https://doi.org/10.1016/j.pocean.2014.04.011

Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H. H., Dashtseren, A. et al., (2019). Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km^2 scale. *Earth-Science Reviews, 193,* 299–316. https://doi.org/10.1016/j.earscirev.2019.04.023

Olefeldt, D., Roulet, N., Giesler, R., & Persson, A. (2013). Total waterborne carbon export and DOC composition from ten nested subarctic peatland catchments-importance of peatland cover, groundwater influence, and interannual variability of precipitation patterns. *Hydrological Processes, 27(16),* 2280–2294. https://doi.org/10.1002/hyp.9358

Olefeldt, D., Persson, A., & Turetsky, M. R. (2014). Influence of the permafrost boundary on dissolved organic matter characteristics in rivers within the Boreal

and Taiga plains of western Canada. *Environmental Research Letters, 9(3)*, 035005. https://doi.org/10.1088/1748-9326/9/3/035005

Patzner, M. S., Kainz, N., Lundin, E., Barczok, M., Smith, C., Herndon, E. et al., (2022). Seasonal Fluctuations in Iron Cycling in Thawing Permafrost Peatlands. *Environmental Science & Technology, 56*(7), 4620–4631. https://doi.org/10.1021/acs.est.1c06937

Petrone, K. C., Jones, J. B., Hinzman, L. D., & Boone, R. D. (2006). Seasonal export of carbon, nitrogen, and major solutes from Alaskan catchments with discontinuous permafrost. *Journal of Geophysical Research: Biogeosciences, 111(2),* 1–13. https://doi.org/10.1029/2005JG000055

Pokrovsky, O. S., Manasypov, R. M., Loiko, S., Shirokova, L. S., Krickov, I. A., Pokrovsky, B. G. et al., (2015). Permafrost coverage , watershed area and season control of dissolved carbon and major elements in western Siberian rivers. *Biogeosciences, 12,* 6301–6320. https://doi.org/10.5194/bg-12-6301-2015

Pokrovsky, O. S., Manasypov, R. M., Loiko, S. V., Krickov, I. A., Kopysov, S. G., Kolesnichenko, L. G. et al., (2016). Trace element transport in western Siberian rivers across a permafrost gradient. *Biogeosciences, 13(6),* 1877–1900. https://doi.org/10.5194/bg-13-1877-2016

Price, N. M., Ahner, B. A., and Morel, F. M. M. (1994). The equatorial Pacific Ocean: Grazer-controlled phytoplankton populations in an iron-limited ecosystem. *Limnology and Oceanography, 39,* 520–534. https://doi.org/10.4319/lo.1994.39.3.0520

Raudina, T. V., Loiko, S. V., Lim, A. G., Krickov, I. V., Shirokova, L. S., Istigechev, G. I. et al., (2017). Dissolved organic carbon and major and trace elements in peat porewater of sporadic, discontinuous, and continuous permafrost zones of western Siberia. *Biogeosciences, 14(14),* 3561–3584. https://doi.org/10.5194/bg-14-3561-2017

Raudina, T. V., Loiko, S. V., Lim, A., Manasypov, R. M., Shirokova, L. S., Istigechev, G. I. et al., (2018). Permafrost thaw and climate warming may decrease the CO2, carbon, and metal concentration in peat soil waters of the Western Siberia Lowland. *Science of the Total Environment, 634,* 1004–1023. https://doi.org/10.1016/j.scitotenv.2018.04.059

Shamov, V. V, Onishi, T., & Kulakov, V. V. (2014). Dissolved iron runoff in Amur Basin rivers in the late XX century. *Water Resources, 41(2),* 201–209. https://doi.org/10.1134/S0097807814020122

Shiraiwa, T. (2012). "Giant Fish-Breeding Forest": A New Environmental System Linking Continental Watershed with Open Water. In Taniguchi, M. & Shiraiwa, T. (Eds.), *The Dilemma of Boundaries*(pp. 73–85). Tokyo, Japan: Springer. http://dx.doi.org/10.1007/978-4-431-54035-9_8

Smedberg, E., Mörth, C. M., Swaney, D. P., & Humborg, C. (2006). Modeling hydrology and silicon-carbon interactions in taiga and tundra biomes from a

landscape perspective: Implications for global warming feedback. *Global Biogeochemical Cycles, 20(2).* https://doi.org/10.1029/2005GB002567

Sunda, W. G. (2012). Feedback interactions between trace metal nutrients and phytoplankton in the ocean. *Frontiers in Microbiology, 3,* 204. https://doi.org/10.3389/fmicb.2012.00204

Suzuki, K., Hattori-Saito, A., Sekiguchi, Y., Nishioka, J., Shigemitsu, M., Isada, T. et al. (2014). Spatial variability in iron nutritional status of large diatoms in the Sea of Okhotsk with special reference to the Amur River discharge. *Biogeosciences, 11(9),* 2503–2517. https://doi.org/10.5194/bg-11-2503-2014

Takeda, S., & Obata, H. (1995). Response of equatorial Pacific phytoplankton to subnanomolar Fe enrichment. *Marine Chemistry, 50,* 219–227. https://doi.org/10.1016/0304-4203(95)00037-R

Takeuchi, W., & Yasuoka, Y. (2004). Development of normalized vegetation, soil and water

indices derived from satellite remote sensing data. *Journal of the Japan Society of Photogrammetry and Remote Sensing, 43(6),* 7–19. In Japanese. https://doi.org/10.4287/jsprs.43.6_7

Tashiro, Y., Yoh, M., Shiraiwa, T., Onishi, T., Shesterkin, V., & Kim, V. (2020). Seasonal variations of dissolved iron concentration in active layer and rivers in permafrost areas, Russian Far East. *Water, 12(9),* 2579. https://doi.org/10.3390/w12092579

Tipping, E. (Ed.). (2002). *Cation Binding by Humic Substances (Cambridge Environmental Chemistry Series)*, Cambridge, MA: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511535598

Tucker, C. J., Townshend, J. R. G., & Goff, T. E. (1985). African landcover classification using satellite data. *Science, 227(4685),* 369–375. https://doi.org/10.1126/science.227.4685.369

Vorobyev, S. N., Pokrovsky, O. S., Serikova, S., Manasypov, R. M., Krickov, I. V., Shirokova, L. S. et al. (2017). Permafrost boundary shift in Western Siberia may not modify dissolved nutrient concentrations in rivers. *Water*, *9*(12), 985. https://doi.org/10.3390/w9120985

Wang, L., Yan, B., Pan, X., & Zhu, H. (2012). The spatial variation and factors controlling the concentration of total dissolved iron in rivers, Sanjiang Plain. *Clean – Soil, Air, Water, 40(7),* 712–717. https://doi.org/10.1002/clen.201100251

Winterfeld, M., Mollenhauer, G., Dummann, W., Köhler, P., Lembke-Jene, L., Meyer, V. D. et al., (2018). Deglacial mobilization of pre-aged terrestrial carbon from degrading permafrost. *Nature Communications, 9*, 3666. https://doi.org/10.1038/s41467-018-06080-w