

MACHINE LEARNING CLASSIFICATION OF THE PEAK LEVELS OF SPRING ICE RUN AT THE ARCTIC RIVERS

I. A. Lisenkov^{1,3,*} , S. A. Kainova¹ , E. O. Luppov^{1,2} , and V. A. Kuznetsov³ 

¹Geophysical Center of the Russian Academy of Sciences, Moscow, Russian Federation

²Russian University of Transport, Moscow, Russian Federation

³National Research Nuclear University MEPhI, Moscow, Russian Federation

*Correspondence to: I. A. Lisenkov, lisenkov@gcras.ru

Abstract: Floods during spring ice run are among the most hazardous hydrological phenomena at the rivers of the northern part of the European territory of Russia (ETR). This paper presents an original approach to classifying the peak levels of spring ice run against the critical water level thresholds (markers of adverse and hazardous phenomena) during spring ice run in the lower reaches of the Onega, the Northern Dvina, the Mezen, and the Pechora rivers. This approach is based on the consolidation of heterogeneous data in spatiotemporal neighborhoods and the application of machine learning methods. A comprehensive dataset, which includes long-term hydrological and meteorological observation series for 1961–2023, channel and catchment morphometric characteristics, and ice regime indicators, aggregated within a circular neighborhood with a radius of 25–250 km around 29 hydrometric stations, has been developed. On the basis of the analysis of hydrological data of 7 locations in the lower reaches of the studied rivers, the effectiveness of classical models of machine learning for binary classification of hydrological situations is demonstrated (“no adverse or hazardous phenomenon” / “adverse or hazardous phenomenon”). Five machine learning model training strategies have been tested (general training, spatial generalization across basins and stations, and isolated training). Isolated training was found to provide the best classification quality for three of the four studied basins (Onega, Northern Dvina, Mezen; F1-macro 0.684–1.000), while model transferring between river basins has demonstrated lower quality (F1-macro 0.255–0.608) due to the high specificity of local hydrological conditions. Spatial generalization at the station level within a single basin (F1-macro 0.684–0.774) is recommended when observations are limited. Feature importance analysis revealed the key role of temperature, precipitation, and hydrological predictors. The developed methods can be useful in planning emergency prevention measures.

Keywords: systems analysis, geospatial data, multiparametric analysis, data consolidation, ice run, ice jams, hydrological forecasting, risk management

Citation: Lisenkov I. A., Kainova S. A., Luppov E. O., and Kuznetsov V. A. (2026), Machine Learning Classification of the Peak Levels of Spring Ice Run at the Arctic Rivers, *Russian Journal of Earth Sciences*, 26, ES3016, EDN: ESQYLD, <https://doi.org/10.2205/2026es001147>

REVIEW

Received: June 2, 2026

Accepted: July 2, 2026

Published: August 31, 2026



Copyright: © 2026. The Authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Among the hazardous hydrological phenomena, ice-related flooding occupies a special place. Such ice-related flooding occurs during the cold season. Ice-related flooding is accompanied by ice overflowing onto the shore, and has additional danger since in addition to flooding it can cause damage to residential buildings, infrastructure, and hydraulic structures. As a result, the damage from ice-related floods often exceeds the one from open water floods (at similar water levels) [Kozlov *et al.*, 2015]. At the Russian Arctic rivers, the highest annual water levels are often observed during the breakup period, that is, during spring ice run and ice jams [Agafonova and Vasilenko, 2020].

The highest levels of spring ice run in the studied river sections depend largely on two main factors: water flow and ice conditions, i.e., they have runoff, ice run, and sometimes ice jam components [Regularities..., 2012]. Therefore, when developing forecasting methods, it is necessary to consider both the factors that shape spring runoff and the factors that determine ice breakup conditions.

The main factors shaping runoff during the snowmelt flood period (and the runoff component of water levels) are snow water equivalent, precipitation (liquid and solid), heat exchange determining the intensity of snowmelt and evaporation, and the water absorption capacity of the basin, which determines potential runoff losses. Traditional methods for forecasting flood characteristics are either based on water balance calculations or rely on empirical relationships with 2–3 selected predictors (snow water equivalent, soil freezing depth, and soil moisture). There are also forecasting methods based on runoff formation models [Simonov, 2023; Varentsova et al., 2022].

The ice component of the water level is determined by the conditions of the spring ice run, the thickness of the ice floes, and the presence of ice jams in the channel. In hydrological practice, methods are being developed for forecasting specifically ice jam levels. Forecasting ice run levels is a more general case, but the predictors of these phenomena are essentially the same. Such forecasts are based on empirical relationships, in modern practice by applying logistic regression or classical machine learning methods (artificial neural networks, decision trees, etc.) [Blaskey et al., 2024; Iglın et al., 2025; Malygin and Aleshin, 2022; Semenova et al., 2020; Sumachev et al., 2024; White, 2003]. The main predictors that should be taken into account were identified back in [Shulyakovsky, 1951]. A slightly revised list of predictors looks like this: water discharge at the ice edge (as a characteristic of the runoff component of the water level and the water conditions during spring ice run), water level at the onset of freeze-up (as a characteristic of freezing conditions and an indicator of hanging dams), and ice thickness (a characteristic of winter conditions and the volume of ice participating in ice run). Instead of water discharge, the rate of water level rise before breakup is sometimes used; instead of ice thickness, the accumulated freezing degree-days is used; instead of the freeze-up level, the duration of autumn ice run and slush run, or predictors based on air temperature data, are used. For rivers freezing due to accumulation of frazil ice and slush, the highest water level at the onset of freeze-up is used; for other sections, water discharges or water levels, as well as ice thickness before the onset of breakup, are used [Buzin, 2015; Madaeni et al., 2020].

In the review [Soloviev and Lisenkov, 2024] it is demonstrated that accounting for the full range of hydrometeorological characteristics and identifying nonlinear relationships between them using classical empirical approaches can be challenging. Systems analysis facilitates the consideration of a river basin as a single system with multiple interconnected characteristics, but it does not always provide the tools to automatically extract hidden patterns from large arrays of heterogeneous data. Machine learning methods, a key technology of artificial intelligence, overcome this limitation and make it possible to work with a large number of predictors and automatically identify hidden patterns and nonlinear dependencies. Thus, systems analysis provides a conceptual framework and meaningful interpretation of the processes under study, while machine learning serves as a universal tool for automatically identifying nonlinear patterns in large arrays of heterogeneous data. Their combined application opens up extensive opportunities in the Earth sciences.

As a basic approach to develop a method of classifying peak levels of spring ice runs against the critical water level thresholds, the method of building classification models on the basis of a deliberately redundant set of predictors has been used.

The aim of the present work is to identify the prerequisites of spring ice run when the water peak level attains critical thresholds on the basis of a wide spectrum of heterogeneous geospatial data, to develop and validate a methodology for their consolidation in spatial neighborhoods for constructing machine learning classification models with assessment of generalization capability under various spatial transfer strategies, as well as to identify

the most informative predictors and their combinations that describe the prerequisites for hazardous water levels during spring ice run in the lower reaches of northern ETR's rivers. The results of feature importance analysis provide the basis for constructing simplified, lower space dimension models applicable for short-term and long-term hydrological practical forecasting.

2. Materials and Methods

To develop methods for classifying peak spring ice run levels against the water level thresholds the following four river basins have been selected: the Onega, the Northern Dvina, the Mezen, and the Pechora river basins. The sections under study are the lower reaches of these four rivers. This region belongs to the subarctic and temperate climate zone. The climate of the territory is characterized by severe, long winters and short, cool summers (<https://nationalatlas.ru/>, <https://gis.gcras.ru/>).

Spring ice breakup for the Northern Dvina, the Mezen, and the Pechora (to a lesser extent for the Onega River) occurs primarily due to the impact of dynamic factors, that is, flood waves and increasing water flow. Ice breakup is often accompanied by the formation of ice jams. During ice run, ice jams can be formed in areas with insufficient ice-permeability (i.e., at sharp bends, narrowings, and channel forks).

The highest annual water levels at the studied river sections are often observed precisely during the period of ice breakup and clearing; for the lower reaches of the Northern Dvina, the Mezen, and the Pechora rivers, this occurs in 60–70% of cases (for the lower reaches of the Onega River, in 20% of cases) [Agafonova *et al.*, 2016].

The subject of the present research is the study of the highest water level during spring ice run at target hydrometric stations. Target stations are those selected for this study on the basis of the availability of hydrological data. Model training and performance assessment have been carried out for the target stations. Other stations, whose data have also been used to generate the input data, are designated as additional stations. Observations are organized within a hydrological year, spanning the period from October 1 to September 30 and incorporating all phases of the ice regime, from ice occurrence to ice clearing. At the rivers of the research area, the highest annual water level may occur both during spring ice run and after ice clearing, depending on the specific hydrological conditions of a given year. When the highest annual level is formed during the ice run period, it can reach critical levels established for a given post and corresponding to the criteria of adverse and hazardous hydrometeorological phenomena. As the critical thresholds for water level, the values established by Rosgidromet for each station were taken. The adverse and hazardous phenomena are defined as per criteria of “Criteria of Hazardous Hydrometeorological Phenomena” (Regulatory Document 52.04.563-2002):

- Adverse Hydrometeorological Phenomenon (AHMP) – is the hydrometeorological phenomenon that significantly hinders or impedes the work of business units and economic sectors but does not meet the criteria of the hazardous phenomenon;
- Hazardous Hydrometeorological Phenomenon (HHMP) – is the meteorological, agrometeorological, hydrological, or maritime hydrometeorological phenomenon and/or complex of hydrometeorological conditions that, due to its value, intensity, or duration, poses a threat to human safety and may cause significant damage to infrastructure facilities and the population.

The task is to provide the binary classification of the highest levels of water during the spring ice run according to the following: the highest ice run level belongs either to the “no AHMP/HHMP” class – if the level does not reach the critical threshold of AHMP, or to the “AHMP/HHMP” class – if this level is at or over-the-top of the critical threshold level of AHMP/HHMP (Figure 1).

To address the task of classification of hydrological phenomena, machine learning methods were used. The training was carried out on the basis of heterogeneous data related to the subject of research, including: meteorological (air temperature, precipitation) and hydrological (water levels and discharges, ice thickness, timing of ice phenomena) observa-

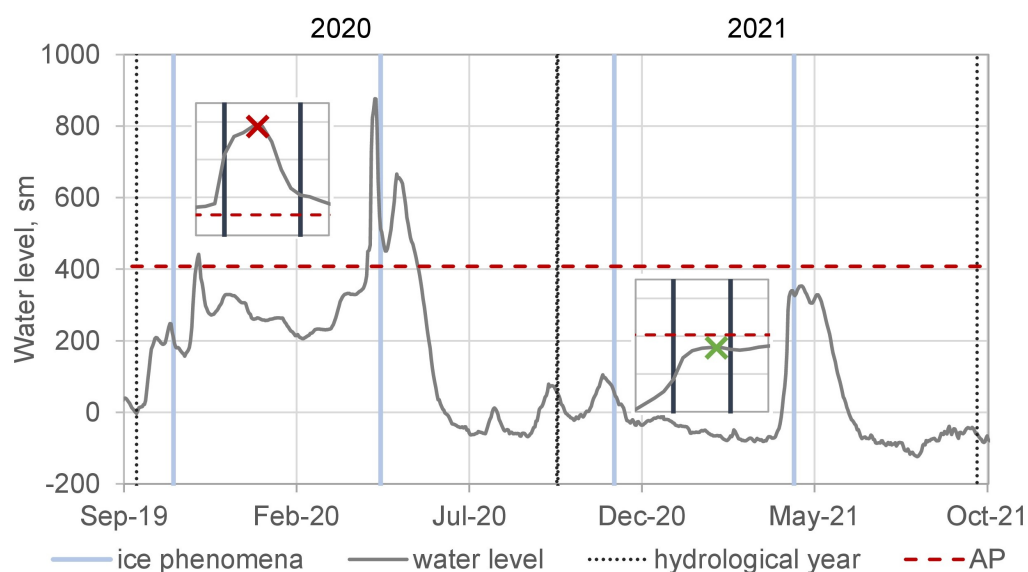


Figure 1. Water levels during the spring ice run on the Northern Dvina River – Ust-Pinega village: two hydrological-year scenarios (2020 → AHMP/HHMP, 2021 → “no AHMP/HHMP”); the level fluctuation during the spring ice run is provided in the insets.

tions at selected hydrometric and meteorological stations; morphometric characteristics of the river network and catchment basins (slopes, channel sinuosity, catchment area); hydrological vector maps IIASA (<https://gis.gcras.ru/>), constituting the sets of polygons with homogeneous hydrological indicators.

To consolidate these heterogeneous data and prepare them for use in machine learning models, the spatiotemporal neighborhood approach described in [Lisenkov et al., 2025] was applied. This approach is widely used in solving geophysical problems, in particular in morphostructural zoning (MSZ) methods used to identify seismically active areas. For example, in [Gvishiani et al., 2016, 2021; Soloviev et al., 2013]. In the present study, this approach is applied for the first time to the problem of hydrological classification.

When organizing the data preparation process for training and testing models, the specifics of spatial structure of hydrological and meteorological observation points were taken into account. River basins are relatively isolated hydrological systems, and therefore correlations between hydrological characteristics exists for stations within a single river basin only. At the same time, meteorological observations are not tied to river catchments. Flat areas are characterized by high spatial connectivity of meteorological fields. Therefore, data from meteorological stations located both within and outside such basins could be used to reconstruct hydrological conditions, provided the distance from the reference gauge section is taken into account. The applicability of data from a meteorological station is determined by its remoteness from the hydrometric station and its position relative to the basin boundaries.

The hydrological input data consisted of observations of water levels, discharges, ice phenomena, and ice thickness at 29 hydrometric stations. For each river basin, target stations were identified (70011, 70842, 70801, 70844, 70847, 70827, 70834). Classification of hydrological situations was implemented for these target stations (Figure 2).

When constructing the dataset for each target station, data from all stations located upstream of the target station in the same basin were used as additional input. If there were multiple target stations in one basin, data from upstream stations were included as additional predictors for downstream target stations. Data from downstream target stations were not used, as information about downstream river conditions cannot be used as a predictor for upstream sections.

Eleven meteorological stations located within the research area were selected to provide for meteorological observations. During the initial processing of meteorological

data, the proximity of the meteorological stations to the hydrometric station was accounted for. When determining the spatial neighborhood radius for data aggregation, the necessity to ensure complete coverage of the target and additional stations with meteorological observations was taken into account.

An analysis of the spatial distribution of the meteorological stations revealed that a radius of 125 km ensures that at least one meteorological station is located within the neighbourhood of each target and additional hydrometric stations. But to achieve stable averaging of meteorological data and account for the spatial variability of meteorological conditions it is imperative to have several meteo stations within the neighborhood of each hydrometric station. This is assured with a neighborhood radius of $R = 300$ km.

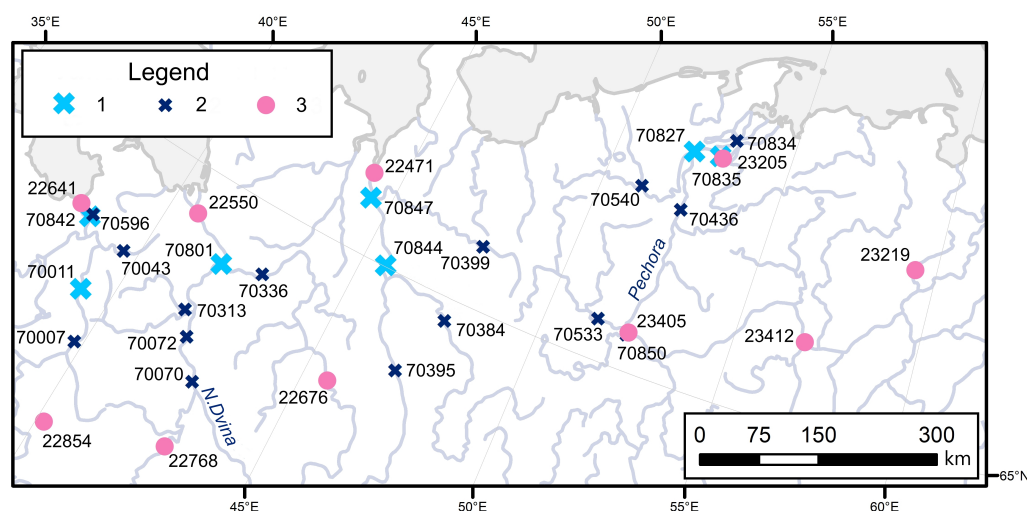


Figure 2. Geographical location of the hydrometric and meteorological stations selected for the research (1 – target hydrometric stations, 2 – additional hydrometric stations, 3 – meteorological stations).

The above-described list of hydrometeorological data was supplemented by the following geospatial data related to the research subject:

- FABDEM (Forest and Buildings removed Copernicus DEM) digital elevation model with a spatial resolution of 30 m in GeoTIFF raster format [Hawker *et al.*, 2022];
- the HydroRivers river network database in ESRI Shapefile vector format, containing information on the morphometric characteristics of river streams, including length, slope, catchment area, and stream order [Lehner and Grill, 2013];
- thematic hydrological maps of the International Institute for Applied Systems Analysis (IIASA) in ESRI Shapefile vector format, characterizing the spatial distribution of river runoff, snow, and ice cover, published on the geoinformation portal of the Geophysical Center of the Russian Academy of Sciences (GIS “Arctic”) (<https://gis.gcras.ru/>).

In accordance with the approach of [Lisenkov *et al.*, 2025], heterogeneous geospatial data from HydroRivers, FABDEM, and thematic IIASA maps were aggregated into circular neighborhoods with a radius of 25 to 250 km, with a step of 25 km, centered at the locations of 29 hydrometric stations, and presented as a set of quantitative parameters. All possible combinations of neighborhoods’ radii were considered for each type of geospatial data (Figure 3).

The resulting feature vector, describing the research subject for each target station in a given hydrological year, contains three types of data: characteristics of the target station, characteristics of additional stations in the same river basin (up to nine stations), and characteristics of meteorological stations located in the neighborhood of both the target station and each of the additional stations. If there are multiple meteorological stations in the neighborhood of a single station, their readings are averaged.

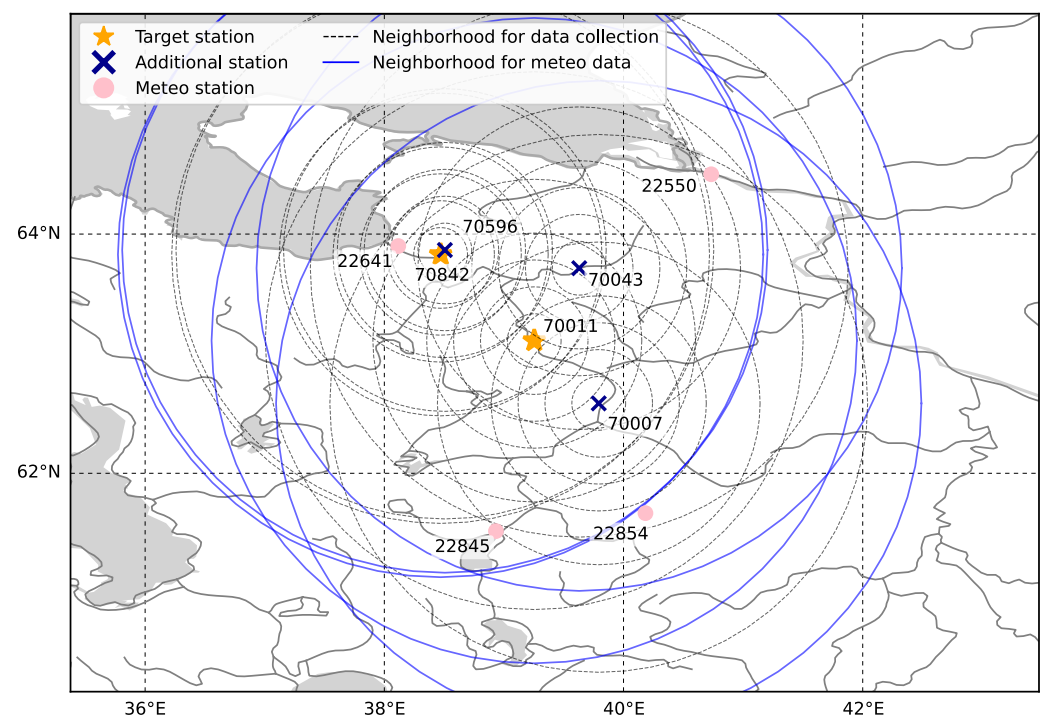


Figure 3. Formation of spatio-temporal neighborhoods for the aggregation of heterogeneous geospatial data (using the Onega River basin as an example).

Taking into account the specifics described above, the final dataset contains 10,625 examples (observation vectors), each characterized by 5,074 features (Table 1). The high dimensionality of the feature space is due to the inclusion of a complex of hydrological, meteorological, and geospatial characteristics for the target and additional stations, as well as the use of different neighborhood radii for geospatial data consolidation.

Table 1. Structure of features for the training data array

Group of Features	Quantity of Attributes	Description
Base	4	Hydrological year, River basin, Max ice run level (H_{ir}), event class ("no AHMP" and "AHMP/HHMP")
Target gauging stations (target_*)	20	Coordinates, radius, and neighborhood type for data consolidation, morphometry, hydrological indicators
Target and additional stations meteorological data (meteo_*_target, hydro_N_meteo_*)	590	[1 target + 9 additional] × 59 features each (including the neighborhood radius for obtaining meteo data)
Additional basin gauging stations (hydro_N_*)	180	Up to 9 slots × 20 features each (Coordinates, radius, and neighborhood type for geospatial data consolidation, morphometry, hydrological indicators)
FABDEM	410	[1 target + 9 additional] × 41 features each
HydroRivers	480	[1 target + 9 additional] × 48 features each
IIASA Hydrological maps	3390	[1 target + 9 slots] × 339 features each

Note. In feature names, the element target (prefix target_ or suffix _target) denotes a feature of the target station; the prefix hydro_N_ denotes the N-th additional (upstream) station within the same basin ($N = 1 \dots 9$); meteo_ denotes meteorological features aggregated within the neighborhood of the corresponding station.

Using the generated dataset, five series of experiments as described below were conducted to evaluate the generalization ability of machine learning models and their

applicability to various spatial conditions. These experiments differed in their training strategies and data partitioning into training and test sets.

- Experiment I (general training): All data were distributed in a ratio of 80% for training to 20% for testing, using stratification by river basin, target point, and event class (“no AHMP”, “AHMP/HHMP”).
- Experiment II (spatial generalization across the river basins): The model was trained on data from three river basins and tested on a fourth basin, which was not involved in the training. A total of four experimental scenarios were conducted using a “leave-one-out” approach, with each basin in turn assigned as a test basin.
- Experiment III (spatial generalization across target stations): The model was trained on data from six target stations and tested on a seventh station, which was not involved in the training. A total of seven experimental scenarios were conducted using a “leave-one-out” approach.
- Experiment IV (isolated training across basins): A separate model was created for each river basin. The data from each basin were divided into training (80%) and test (20%) sets, independently of the other basins, using stratification by target gauging stations and event class (“no AHMP”, “AHMP/HHMP”).
- Experiment V (isolated training by target gauging stations): A separate model was created for each target station. The data from each gauging station were divided into training (80%) and test (20%) sets, stratified by event class (“no AHMP”, “AHMP/HHMP”).

The binary classification of hydrological events was implemented using ensemble machine learning methods: ExtraTrees, RandomForest, GradientBoosting, Logistic Regression, and an MLP multilayer neural network. In the resulting dataset, observations for which critical thresholds had not been exceeded (“no AHMP/HHMP”) comprised 73% of the sample, while observations for which critical thresholds had been reached or exceeded (“AHMP/HHMP”) comprised 27%. To account for and counter-balance this class imbalance during training, the following methods were used: sample-based minority class expansion (SMOTE) and setting different class weights during training [Elreedy and Atiya, 2019].

The quality of the models was assessed using F1-score metrics (weighted and macro), as well as Precision and Recall for the “AHMP/HHMP” class. Detailed definitions and calculation formulas for these metrics are provided in [Lisenkov et al., 2025]. For interpretive convenience, the following indicative thresholds were introduced in this study: F1-macro ≥ 0.5 (satisfactory quality), F1-macro ≥ 0.7 (good quality). Preference was given to the F1-macro metric as the most informative for imbalanced classes, since it equally considers the classification quality of both classes, regardless of their ratio in the sample.

3. Results and Discussion

The results of classifying adverse and hazardous hydrological events demonstrated varying effectiveness of the developed models depending on the chosen training strategy and the characteristics of the target stations (Table 2).

General training on an integrated sample for all basins (Experiment I) demonstrated good classification quality (F1-macro in the range of 0.455–1.000) and practical applicability. However, the results are significantly determined by the class balance: at sites with the “AHMP/HHMP” proportion of less than 10%, the model tends to degenerate into predicting the dominant class (“no AHMP/HHMP”), which artificially inflates the F1-weighted value (up to 0.886) at a low F1-macro (0.455–0.480). The general training strategy does not ensure stable classification quality at sites with sharply different class proportions and is recommended as a baseline for subsequent experiments only.

Spatial generalization across river basins (Experiment II) did not yield the expected results: low F1-macro values (0.254–0.608) indicate the high specificity of hydrological conditions and the impossibility of successfully transferring the models between different catchments without significant loss of quality.

In contrast to Experiment II, spatial generalization across target stations (Experiment III) demonstrated potential for practical application. For the Pechora and the Onega basins,

Table 2. Results of the binary classification of hydrological events: “no AHMP” vs “AHMP/HHMP”

Experiment	River Basin	Station Code	Precision	Recall	F1 Macro	F1 Weighted	Sample Size	Best Model
I (General)	Onega	70011	0.827	0.843	0.659	0.815	370	ExtraTrees
		70842	1.000	1.000	1.000	1.000	320	ExtraTrees
	Northern Dvina	70801	0.694	0.833	0.455	0.758	300	—
		70827	0.933	0.800	0.688	0.838	250	—
	Pechora	70834	0.833	0.833	0.455	0.833	300	—
		70844	0.852	0.923	0.480	0.886	325	—
	Mezen	70847	0.833	0.833	0.455	0.833	300	—
II (Cross-Basin)	Onega	70011	0.688	0.738	0.535	0.704	1550	RandomForest
		70842	0.907	0.900	0.608	0.904	1575	RandomForest
	Northern Dvina	70801	0.861	0.270	0.255	0.344	1575	DecisionTree
		70827	0.635	0.588	0.513	0.607	1275	GradBoost
	Pechora	70834	0.794	0.783	0.555	0.789	1500	GradBoost
		70844	0.875	0.880	0.572	0.877	1575	RandomForest
	Mezen	70847	0.773	0.805	0.607	0.771	1575	RandomForest
III (Cross-Station)	Onega	70011	0.809	0.823	0.685	0.799	1550	GradBoost
		70842	0.957	0.865	0.703	0.894	1575	ExtraTrees
	Northern Dvina	70801	0.861	0.270	0.255	0.344	1575	DecisionTree
		70827	0.884	0.863	0.774	0.843	1275	ExtraTrees
	Pechora	70834	0.900	0.812	0.717	0.837	1500	ExtraTrees
		70844	0.900	0.892	0.656	0.896	1575	GradBoost
	Mezen	70847	0.788	0.794	0.676	0.791	1575	DecisionTree
IV (Isolated Basin)	Onega	70011	0.884	0.865	0.684	0.833	370	ExtraTrees
		70842	1.000	1.000	1.000	1.000	320	ExtraTrees
	Northern Dvina	70801	0.833	0.833	0.700	0.833	300	GradBoost
		70827	0.788	0.700	0.412	0.741	250	DecisionTree
	Pechora	70834	0.933	0.667	0.556	0.741	300	DecisionTree
		70844	1.000	1.000	1.000	1.000	325	LogReg
	Mezen	70847	0.825	0.750	0.429	0.786	300	LogReg
V (Isolated Station)	Onega	70011	0.865	0.865	0.780	0.865	370	GradBoost
		70842	1.000	1.000	1.000	1.000	320	ExtraTrees
	Northern Dvina	70801	0.833	0.833	0.700	0.833	300	GradBoost
		70827	0.800	0.800	0.444	0.800	250	ExtraTrees
	Pechora	70834	0.833	0.833	0.455	0.833	300	ExtraTrees
		70844	1.000	1.000	1.000	1.000	325	GradBoost
	Mezen	70847	0.845	0.833	0.491	0.839	300	ExtraTrees

F1-macro values in the range of 0.685–0.774 were achieved; however, the transferability of the approach to other monitoring stations requires further validation at the subsequent stages of the research.

Isolated training across individual river basins (Experiment IV) and target stations (Experiment V) provided the best classification quality for three of the four studied basins (Onega, Northern Dvina, and Mezen; F1-macro 0.684–1.000). For river basins with sufficient observation data volumes, isolated training is preferable. It ensures high accuracy due to maximum adaptation to local hydrological conditions.

Among the classification algorithms, ExtraTrees demonstrated the most stable results and can be recommended as the base algorithm for this task. GradientBoosting is an effective alternative, demonstrating competitive results in Experiments III and V. In certain cases with linearly separable classes, simple models can also be applied: for example, logistic regression achieved $F1=1.0$ for gauge 70844 in Experiment IV.

Feature importance analysis using the Permutation Feature Importance (PFI) approach [Breiman, 2001] showed that the water level classification task is solved primarily on the basis of hydrological and meteorological indicators (Figure 4). The composition of leading features varies significantly depending on the basin. For the Onega, the most informative parameter is the average air temperature 10 days before ice breakup ($T_mean_br_minus_10$), followed by liquid precipitation in February ($X_gt_0_II$) at the target and adjacent stations. For the Northern Dvina, liquid precipitation in February ($X_gt_0_II$) at upstream stations and solid precipitation in November–April predominate. For the Pechora, the leading indicators are the ice jam onset date (D_i_j) at adjacent stations, the accumulated positive degree-day in March ($APDD_III$), and liquid precipitation in February ($X_gt_0_II$). For the Mezen, the most informative parameters are liquid precipitation in February ($X_gt_0_II$) at adjacent stations, maximum ice thickness, and the highest water level during the first 30 days of the ice cover period. All basins have the accumulated freezing degree-day during the ice cover period at the target station ($AFDD_ic_target$).

Geospatial data (FABDEM, HydroRivers, IIASA) are not presented among the most important features in any of the basins, which confirms their limited applicability for this task due to the static nature of this data.

The composition of the most significant indicators is somewhat consistent with the composition of the indicators in similar studies for rivers in the northern European part of Russia (this includes such indicators as water level at the onset of freeze-up, ice thickness, and/or the accumulated freezing degree-days during the ice cover period). However, unlike in these studies, the leading indicators included meteorological characteristics before ice breakup (average air temperature 10 days before ice breakup, total liquid precipitation in February, etc.). This may be due to a bias toward spring indicators in the overall list of meteorological characteristics.

Table 3. Description of the identifiers presented in Figure 4

Identifier	Description
$T_mean_br_minus_10$	Average air temperature 10 days before ice breakup at target station, °C
$X_gt_0_II$	Sum of liquid precipitation in February, mm
$X_lt_0_I/IV/XI/XII$	Sum of solid precipitation in Jan / Apr / Nov / Dec, mm
$AFDD_I/II/XI$	Accumulated freezing degree-days in Jan / Feb / Nov, °C·day
$AFDD_minus_15$	Accumulated freezing degree-days below -15°C during ice cover period, °C·day
$AFDD_ic_target$	Accumulated freezing degree-days during the ice cover period at target station, °C·day
H_ic	Water level on the day of freeze-up onset, cm
$APDD_III/V$	Accumulated positive degree-days in Mar / May, °C·day
$APDD_max_br_target$	Accumulated max positive degree-days within 30 days before ice breakup at target station, °C·day
$Hmax_ic$	Highest water level in the first 30 days of ice cover period, cm
IT_max	Maximum ice thickness, cm
D_i_j	Ice jam start date
N_fr	Duration of freezing (between ice appearance and freeze-up onset dates)
N_frr	Duration of autumn ice run and slush run, days
$catchment_area_km2$	Catchment area, km^2

However, the most informative indicator – the average air temperature 10 days before ice breakup – is operational in nature, becoming available only 1–5 days before the event,

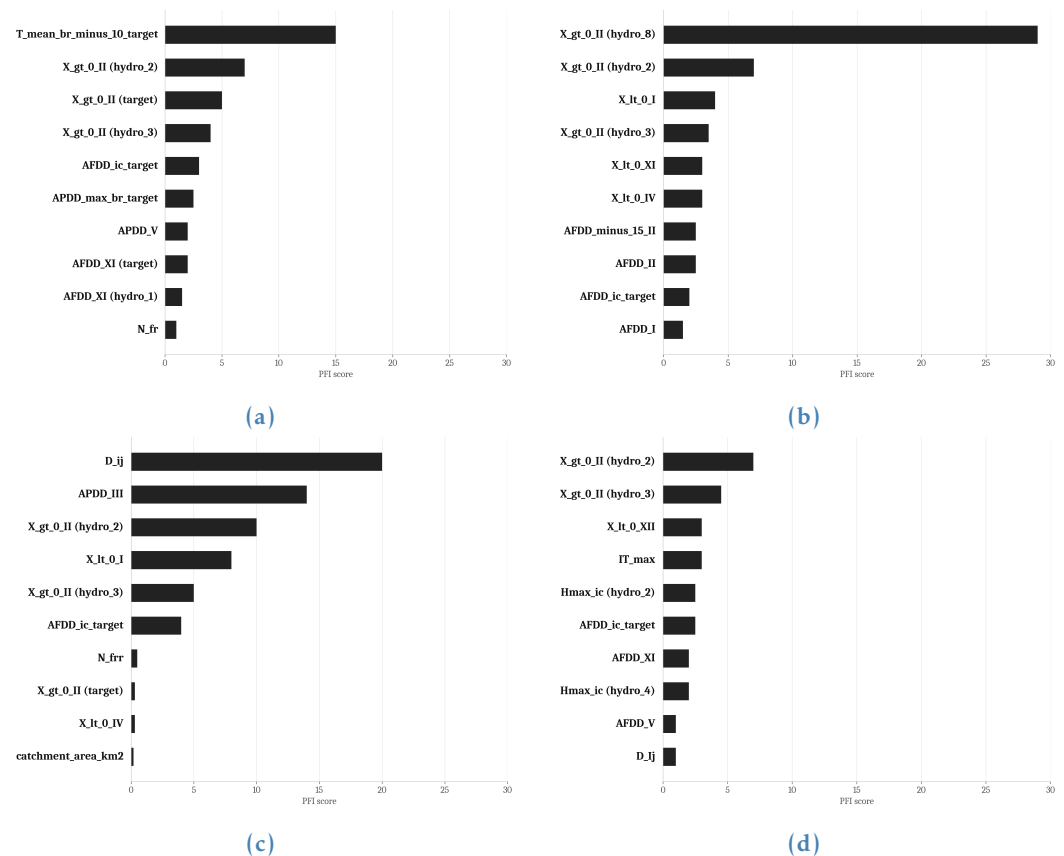


Figure 4. Feature importance analysis for ice event classification. The description of identifiers is provided in Table 3. (a) Onega; (b) Northern Dvina; (c) Pechora; (d) Mezen.

limiting the forecast lead time when used. Indicators providing a lead time of 50–75 days (AFDD_II, X_gt_0_II, AFDD_minus_15) are available by the end of February and can serve as the basis for a long-term forecast. Hydrological indicators – the highest level during the first 30 days of the ice cover period (Hmax_ic) and the level on the day of freeze-up onset (H_ic) – are significant primarily for the Mezen and Pechora basins and provide a lead time of 90–150 days.

The present study addresses the problem of classifying the highest level of spring ice run during the current hydrological year relative to critical threshold levels for adverse or hazardous phenomena without specifying a specific moment of occurrence. Consequently, it seems appropriate to use the developed classification models in conjunction with the river breakup date forecasting models [Limber *et al.*, 2025], which would further refine the time horizon of the expected event and expand the possibilities for preventive decision-making.

4. Conclusions

A comprehensive methodology for generating datasets for machine learning models is proposed, consolidating heterogeneous hydrological and meteorological observation series (1961–2023), morphometric characteristics extracted from the FABDEM digital elevation model and the HydroRivers database, and IIASA hydrological maps within circular spatio-temporal neighborhoods with a radius of 25–250 km around 29 hydrological stations, including 7 target ones (10,625 observations, 5,074 features), for binary classification of spring ice run levels (“no AHMP/HHMP” / “AHMP/HHMP”) in the lower reaches of the Onega, Northern Dvina, Mezen, and Pechora rivers, where the class “no AHMP/HHMP” corresponds to cases in which the peak ice run level does not attain the critical threshold of an adverse hydrometeorological phenomenon (AHMP), and the class “AHMP/HHMP”

corresponds to cases in which it is attained or exceeded, meeting the criteria of an AHMP or a hazardous hydrometeorological phenomenon (HHMP).

Testing five training strategies revealed that isolated learning provides the best classification quality for most of the studied basins; spatial generalization at the station level within the studied basins is recommended when observation series are limited; cross-basin model transfer demonstrates lower and inconsistent quality due to the high specificity of local hydrological conditions. The most informative were the dynamic hydrometeorological predictors: average air temperature 10 days before ice breakup, accumulated positive degree-days in February–March, and water levels at the onset of freeze-up. The contribution of morphometric and geospatial data (FABDEM, HydroRivers, IIASA) proved insignificant.

The obtained results could be used while forecasting the characteristics of the ice regime with the lead times that range from 1–5 days when using air temperature data before ice breakup (both actual and forecast), to 50–75 days when relying on February precipitation data, and up to 90–150 days when using autumn–winter hydrological parameters data. It opens up the potential applicability of the approach (in hazardous ice phenomena at the rivers of the European sector of the Arctic zone of Russia early warning systems) across several forecast horizons – from operational response to long-term preventive planning.

Acknowledgments. This work employed facilities and data provided by the Shared Research Facility “Analytical Geomagnetic Data Center” of the Geophysical Center of RAS (<http://ckp.gcras.ru/>).

Funding. This work was conducted in the framework of budgetary funding of the Russian University of Transport (Project No. 103-00001-26-00, dated January 15, 2026).

References

- Agafonova S. A., Frolova N. L., Vasilenko A. N., et al. Ice regime and dangerous hydrological phenomena on rivers of the Arctic zone of European Russia // *Lomonosov Geography Journal*. — 2016. — No. 6. — P. 41–49. — (In Russian).
- Agafonova S. A. and Vasilenko A. N. Hazardous Ice Phenomena In Rivers Of The Russian Arctic Zone Under Current Climate Conditions And The Safety Of Water Use // *Geography, Environment, Sustainability*. — 2020. — Vol. 13, no. 2. — P. 43–51. — <https://doi.org/10.24057/2071-9388-2020-12>
- Blaskey D., Racine I., Harlan M., et al. Using Remote Sensing, Statistical, and Machine Learning Techniques to Assess Alaskan River Ice Phenology and Thickness. — ESS Open Archive [Preprint], 2024. — <https://doi.org/10.22541/essoar.172954196.61576186/v1>
- Breiman L. Random Forests // *Machine Learning*. — 2001. — Vol. 45, no. 1. — P. 5–32. — <https://doi.org/10.1023/a:1010933404324>
- Buzin V. A. Ice Frazil and Ice Jams on the Rivers of Russia. — St. Petersburg : Es Pe Kha Publ., 2015. — 240 p. — (In Russian).
- Elreedy D. and Atiya A. F. A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance // *Information Sciences*. — 2019. — Vol. 505. — P. 32–64. — <https://doi.org/10.1016/j.ins.2019.07.070>
- Gvishiani A. D., Dzeboev B. A. and Agayan S. M. FCaZm intelligent recognition system for locating areas prone to strong earthquakes in the Andean and Caucasian mountain belts // *Izvestiya, Physics of the Solid Earth*. — 2016. — Vol. 52, no. 4. — P. 461–491. — <https://doi.org/10.1134/s1069351316040017>
- Gvishiani A. D., Dzeboev B. A., Agayan S. M., et al. Fuzzy Sets of High Seismicity Intersections of Morphostructural Lineaments in the Caucasus and in the Altai-Sayan-Baikal Region // *Journal of Volcanology and Seismology*. — 2021. — Vol. 15, no. 2. — P. 73–79. — <https://doi.org/10.1134/s0742046321020032>
- Hawker L., Uhe P., Paulo L., et al. A 30 m global map of elevation with forests and buildings removed // *Environmental Research Letters*. — 2022. — Vol. 17, no. 2. — P. 024016. — <https://doi.org/10.1088/1748-9326/ac4d4f>
- Iglin S. M., Moreydo V. M. and Golovnin K. I. Forecasting rare hydrological events by machine learning methods: case study of ice jams on the Pechora river // *Lomonosov Geography Journal*. — 2025. — Vol. 80, no. 1. — P. 87–97. — <https://doi.org/10.55959/msu0579-9414.5.80.1.6> — (In Russian).

- Kozlov D. V., Buzin V. A., Frolova N. L., et al. Hazardous Ice Phenomena on Rivers and Reservoirs of Russia / ed. by D. V. Kozlov. — Moscow : Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, 2015. — 348 p. — (In Russian).
- Lehner B. and Grill G. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems // *Hydrological Processes*. — 2013. — Vol. 27, no. 15. — P. 2171–2186. — <https://doi.org/10.1002/hyp.9740>
- Limber R., Hoffman F. M., Schwenk J., et al. Long Short-Term Memory Model to Forecast River Ice Breakup Throughout Alaska USA // *Water Resources Research*. — 2025. — Vol. 61, no. 9. — e2025WR040635. — <https://doi.org/10.1029/2025wr040635>
- Lisenkov I. A., Soloviev A. A., Kuznetsov V. A., et al. Generalized dataset of geological and geophysical information on the eastern sector of the Russian Arctic for machine learning-based analysis // *Russian Geology and Geophysics*. — 2025. — Vol. 66, no. 2. — P. 210–223. — <https://doi.org/10.2113/rgg20244747>
- Madaeni F., Lhissou R., Chokmani K., et al. Ice jam formation, breakup and prediction methods based on hydroclimatic data using artificial intelligence: A review // *Cold Regions Science and Technology*. — 2020. — Vol. 174. — P. 103032. — <https://doi.org/10.1016/j.coldregions.2020.103032>
- Malygin I. V. and Aleshin I. M. Forecasting Ice Jams on the Lena River Using Machine Learning Methods // *Izvestiya, Atmospheric and Oceanic Physics*. — 2022. — Vol. 58, no. 10. — P. 1218–1225. — <https://doi.org/10.1134/s0001433822100061>
- Regularities of hidrological processes / ed. by N. I. Alekseevsky. — Moscow : GEOS, 2012. — 733 p. — (In Russian).
- Semenova N. K., Sazonov A. A. and Krylenko I. N. Predicting the possibility of ice jams formation using machine learning methods // *Proceedings of international scientific conference in memory of outstanding Russian scientist Yury Vinogradov. IV Vinogradov Conference Hydrology: From Learning to Worldview*. — St. Petersburg : Izdatelstvo VVM, 2020. — P. 358–361. — (In Russian).
- Shulyakovsky L. G. On ice jams and jam water levels during river breakup // *Meteorologiya i Gidrologiya*. — 1951. — No. 7. — P. 45–49. — (In Russian).
- Simonov Yu. A. Systems for preparing and issuing hydrological forecasts at hydrometcenter of Russia // *Meteorologiya i Gidrologiya*. — 2023. — No. 12. — P. 27–36. — <https://doi.org/10.52002/0130-2906-2023-12-27-36> — (In Russian).
- Soloviev A. A. and Lisenkov I. A. Review and prospects of applying modern approaches to comprehensive geodata analysis for predicting the spatial distribution of geological and geophysical parameters // *Geophysical Research*. — 2024. — Vol. 25, no. 2. — P. 20–45. — <https://doi.org/10.21455/gr2024.2-2> — (In Russian).
- Soloviev A. A., Novikova O. V., Gorshkov A. I., et al. Recognition of potential sources of strong earthquakes in the Caucasus region using GIS technologies // *Doklady Earth Sciences*. — 2013. — Vol. 450, no. 2. — P. 658–660. — <https://doi.org/10.1134/s1028334x13060159>
- Sumachev A. E., Banshchikova L. S. and Griga S. A. Using the Methods of Neural Network Learning for Peak Water Level Prediction: A Case Study for the Rivers in the Dvina-Pechora Basin // *Russian Meteorology and Hydrology*. — 2024. — Vol. 49, no. 4. — P. 354–362. — <https://doi.org/10.3103/s1068373924040095>
- Varentsova N. A., Kireeva M. B., Kharlamov M. A., et al. Spring river runoff in the European part of Russia: main factors and their estimation. I. Review of previous studies // *Hydrometeorological Research and Forecasting*. — 2022. — Vol. 2. — P. 92–116. — <https://doi.org/10.37162/2618-9631-2022-2-92-116> — (In Russian).
- White K. D. Review of prediction methods for breakup ice jams // *Canadian Journal of Civil Engineering*. — 2003. — Vol. 30, no. 1. — P. 89–100. — <https://doi.org/10.1139/102-047>