GSTS™
Navigate Tomorrow Today

ARCTIC SURVEILLANCE AND ICE PREDICTION:

Enhanced with OCIANA™

ARCTIC SURVEILLANCE AND ICE PREDICTION: ENHANCED WITH OCIANA™

GSTS was engaged by a Canadian port to develop an Ice Build Up Prediction and Management Solution. The results of this work are significant and evidenced in our findings. In summary, the benefits include:

- Prediction of ice build-up in waterways of interest to facilitate advanced planning of vessel movement up to 5 days in advance

- Notifications/alerts via email to ports, terminals, pilotages, and shipping lines to support vessel scheduling and facilitate JIT arrivals to reduce fuel consumption and greenhouse gas

Introduction

According to the ice climatology published by the Canadian Coast Guard (2012), in the St. Lawrence River, the first ice formation normally occurs during the second week of December. By the end of December, the south half of the estuary, west of a line from Pointe-des-Monts to Marsoui, is ice covered. Normally, freeze-up in the remainder of the river begins in early January. Particularly extensive areas of fast ice are found in Lake Saint-Pierre, in sections of the river between Lake St. Pierre and Montreal where islands hold the fast ice, and in the non-navigable channels between Montreal and Sorel. Dispersal of the ice begins in late February and is first evident in the Estuary near the mouth of the Saguenay River where ice concentrations fall to very open range. Breakup on the St. Lawrence River usually begins near the middle of March in leeward and thinner ice areas. The river is normally clear of all ice by the first week of April.

GSTS™
Navigate Tomorrow Today

Ice jam is prone to occur in Lake Saint-Pierre downstream from Montréal, blocking commercial navigation and causing floods that can involve significant damage. The water depth of Lake Saint-Pierre is shallow, except in the navigation channel, where it is maintained to a minimum 11.3 m below chart datum. Because of the lower water velocity and the presence of curves in the channel, this area is more subject to ice congestion than other sections in the river (Figure 1). Ice jams usually begin to form between Curve no. 1 (Curve Maskinongé) and Curve no. 2 (Curve Louiseville) (Morse et al., 2003).

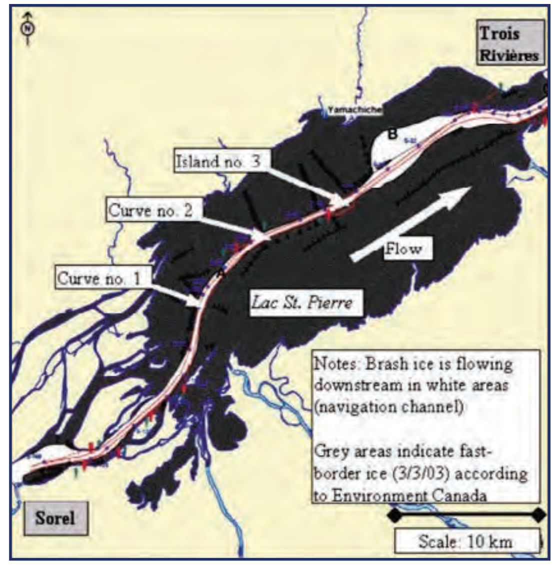

Figure 1. Navigation channel in Lake St. Pierre and key sections prone for ice jams (Morse et al., 2003).

Through the work with the Canadian port, GSTS collected ice jam event records of the Montréal-Québec City section of the St. Lawrence River by reviewing literature and checking other sources such as news. Scalabrini and Morse (2021) mentioned there are 4 recorded ice jam events since 1993 in this area, but no details of these events were given. According to a news report, the first ice jam event in 2019 was due to freeze-over causing a 13-km long ice jam near Île des Barques in the Lake St. Pierre. Canada, unlike the US, does not have an explicit ice-jam database; however, Canada has the Canadian Disaster Database (CDD), which includes records of significant hydrological-meteorological flood events. Literature and news reports in addition to CDD were explored to compile a list of ice jam events in the Montréal-Québec City section of the St. Lawrence River.

Figure 2. Selected points in Lake St. Pierre for ice condition forecast along with nearby weather stations.

Since an ice jam event in Lake Saint-Pierre is an extreme event there are limited records available for developing data-intensive forecasting models. Also, there is limited historical ice, meteorological and hydrometric data in this region. It is almost impossible to develop an AI model to predict the occurrence of such an extreme event with limited data. Therefore, we focused on the prediction of ice concentration as a proxy for ice jam event likelihood. We selected 10 points along the navigation channel in Lake Saint-Pierre. Figure 2 shows the 10 points along with 4 nearby weather stations.

Data Collection

Two types of data are needed to build an AI model for predicting ice conditions in Lake St. Pierre: target variable and predictor variables. Here, the target variable is ice concentration, and the predictor variables are meteorological and hydrometric variables.

Target Variable

The target variable is total ice concentration, which was obtained from ice charts. GSTS obtained electronic ice charts from the Data Science and Artificial Intelligence Team of National Research Council Canada (NRC) that cover the vicinity of Lake St. Pierre. NRC converted the original ice charts (Figure 3 as an example) into Shapefiles with the software, ICEgg, provided by CIS. The electronic daily ice charts range from December 2010 to December 2021 with a large number of missing days. In total, there were 743 daily ice charts of Lake St. Pierre and 53 daily ice charts downstream of Lake St. Pierre from mid-December to early April of each year.

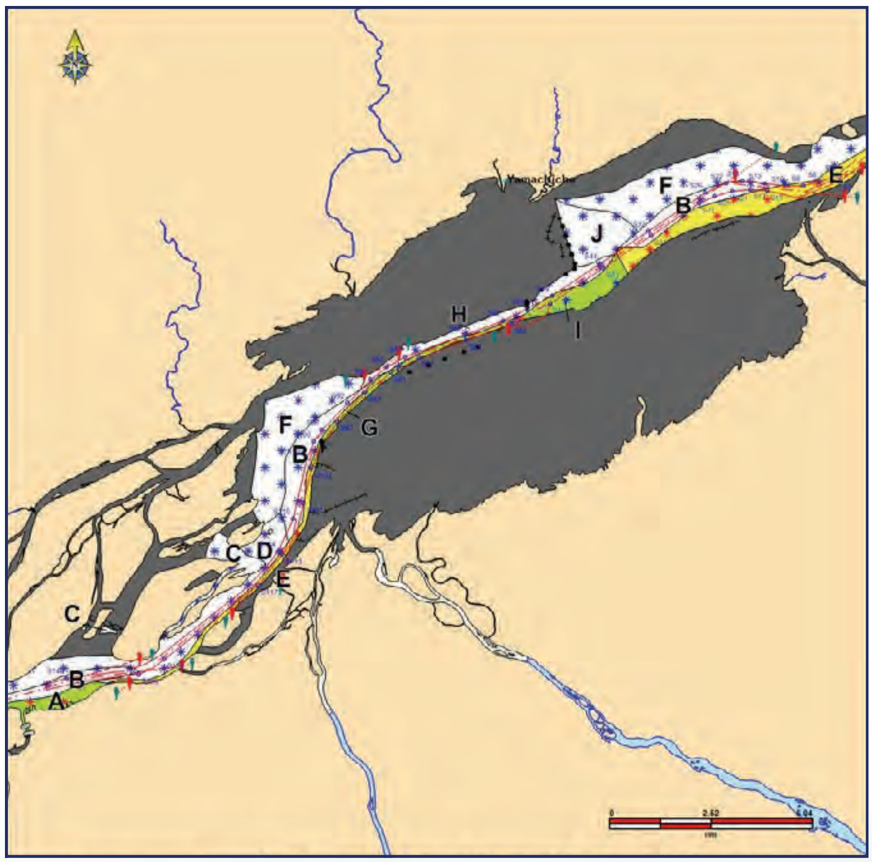

Figure 3. Example of ice chart of Lake St. Pierre in GIF format on February 5, 2020.

Predictor Variables

Through communications with the Canadian port, the following predictor variables were identified for forecasting ice conditions in Lake St. Pierre:

- freezing-degree days,
- wind,
- water flow changes upstream Montreal,
- water level changes downstream Montreal,
- water temperature,
- snow buildup,
- ice thickness,
- tides, and
- current.

We conducted data search for the above predictor variables that were in the temporal coverage of the ice charts (i.e., December 2010 to December 2021). Table 4 shows the data sources of available historical predictor variables. To the best of our knowledge, there is no historical data of water temperature, ice thickness, current, or river discharge in the vicinity of Lake St. Pierre. Water level change at Sorel consists of changes due to tides, upstream river discharge and meteorological factors (e.g., precipitation and wind); therefore, we did not consider tides separately but used water level change at Sorel as one predictor variable. To summarize, the final predictor variables for AI model development are:

- air temperature,
- wind,
- water level at Cornwall,
- water level at Sorel, and
- snow.

Table 4. List of available historical meteorological and hydrometric data sources for AI model development

Variables	Data Source	Notes
Hourly air temperature, wind	Environment and Climate Change Canada	Historical weather station data
Daily air temperature, wind, snow	Canadian Centre for Climate Services	Daily climate data derived from observations
Daily snow	Environment and Climate Change Canada	Historical adjusted daily rainfall and snowfall dataset for Canada
Water level and discharge	Water Office	Historical hydrometric data across Canada

GSTS™
Navigate Tomorrow Today

Both historical hourly data and climate daily data were downloaded for 4 weather stations. Nicolet has the most complete data that match the temporal coverage of the ice data, but there is no snow data. Limited snow data is available for Louiseville from the hourly and daily weather datasets. There is no precipitation/snow data for Louiseville from the historical adjusted daily rainfall and snowfall dataset, while snow data is available for Nicolet until 2018. The snow data for Louiseville from the daily weather data was combined with the daily snow data for Nicolet from the historical adjusted rainfall and snowfall dataset to get snow data from 2010 to 2021.

Daily water level data is available from 1919 to 2021 at Cornwall Canal (Station #: 02MC022; 45°00’53” N and 74°42’41” W) and from 1897 to 2021 at Lanoraie (Station #: 02OB011; 45°57’33” N and 73°12’52” W). There is no historical water level data available at Sorel, therefore, Lanoraie is used instead.

Finally, daily temperature and wind data at Nicolet, daily snow data at Nicolet supplemented with snow data at Louiseville for filling in missing values, and daily water levels at Cornwall Canal and Lanoraise were fused with the ice data as the training dataset for model development.

Machine Learning Model Development

Our aim of this study is to predict ice concentration (i.e., integer values from 0 to 10) at the 10 selected points along the navigation channel in Lake Saint-Pierre with meteorological (i.e., air temperature, wind speed, wind direction, and snow) and hydrometric (i.e., water level upstream and near the lake) variables. We solve this modelling task with two approaches: a multivariate time series and multi-class classification model, and a regression model. Finally, the Regression Model with XGBoost was chosen due to better predictability.

Regression Model with XGBoost

Extreme Gradient Boosting (XGBoost) is a scalable, distributed gradient-boosted decision tree machine learning library. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems. XGBRegressor was selected to develop a regression model. The following hyperparameters were tuned:

- Max Depth: The deeper the trees, the more complex they are. Tuning values range from 2 to 20 with an increment step of 4.
- Min Child Weight: If the tree splitting creates a node with a sum that is lower than this value, the model will stop splitting to avoid overly complex models. Tuning values range from 0 to 10 with an increment step of 2.
- Eta: It is the step size of the optimization that is used to prevent overfitting. Tuning values are 0.01, 0.05, 0.1, 0.15, 0.2, 0.3, and 0.5.
- Gamma: This is the minimum loss reduction that allows further splitting of a node. Tuning values range from 0 to 10 with an increment step of 2.

The model used 38 features: month, day, temperature, wind speed, wind direction, snow, and water levels at Cornwall Canal and Lanoraie. For each of the meteorological and hydrometric feature, 6 consecutive daily values were used.

K-Folds cross-validator was also used to train the model: The training dataset was split into 6 consecutive folds (with shuffling), and each fold was then used once as a validation while the 5 remaining folds formed the training set. The predicted ice concentration values were rounded to integers to obtain valid ice concentration values from 0 to 10. The model evaluation metrics were also root mean square error (RMSE) and mean absolute mean error (MAE).

The hyperparameters and evaluation metrics of the best trained models are shown in Table 5. The main conclusions are:

- The accuracy of the regression model is much better than that of the classification model.
- The RMSE ranges from 1.61 to 2.06 and MAE from 1.11 to 1.46.
- Point 8 is the least predictable with the smallest RSME and MAE among all the 10 points, while points 1 and 5 are the most predictable.
- Analysis of feature importance showed that generally temperature/snow was the most/least importance feature.

Table 5. Hyperparameters and evaluation metrics of the best trained regression models

Model Run	Parameters	Average RMSE	RMSE Standard Deviation	Average MAE	MAE Standard Deviation
xgb1-P01	(6, 4, 2, 0.1)	1.61	0.18	1.13	0.10
xgb1-P02	(6, 0, 6, 0.1)	1.68	0.14	1.20	0.12
xgb1-P03	(6, 4, 8, 0.05)	1.70	0.13	1.24	0.11
xgb1-P04	(6, 0, 8, 0.1)	1.78	0.15	1.23	0.09
xgb1-P05	(14, 4, 8, 0.1)	1.65	0.10	1.11	0.07
xgb1-P06	(10, 2, 8, 0.05)	1.71	0.12	1.16	0.08
xgb1-P07	(6, 6, 6, 0.1)	1.81	0.12	1.27	0.12
xgb1-P08	(10, 0, 8, 0.05)	2.06	0.10	1.46	0.07
xgb1-P09	(18, 4, 6, 0.05)	1.81	0.12	1.24	0.10
xgb1-P10	(18, 8, 8, 0.15)	1.66	0.16	1.21	0.11

As snow was the least important feature for the model performance. A new set of models were trained without snow as a training feature. The parameters and evaluation metrics of the best models are shown in Table 6. The new models had equal, if not better, evaluation metrics than the best models trained with all features. Therefore, the final best models are regression models with 32 features excluding snow.

Table 6. Hyperparameters and evaluation metrics of the best trained regression models with all features except snow

Model Run	Parameters	Average RMSE	RMSE Standard Deviation	Average MAE	MAE Standard Deviation
xgb2-P01	(14, 8, 0, 0.2)	1.64	0.14	1.18	0.12
xgb2-P02	(6, 0, 6, 0.1)	1.68	0.13	1.20	0.12
xgb2-P03	(6, 6, 8, 0.1)	1.70	0.11	1.23	0.10
xgb2-P04	(6, 0, 8, 0.1)	1.79	0.15	1.22	0.13
xgb2-P05	(14, 6, 8, 0.1)	1.63	0.12	1.11	0.08
xgb2-P06	(6, 0, 6, 0.15)	1.71	0.12	1.16	0.07
xgb2-P07	(18, 0, 6, 0.1)	1.79	0.10	1.24	0.08
xgb2-P08	(10, 0, 8, 0.05)	2.03	0.12	1.45	0.08
xgb2-P09	(14, 4, 6, 0.1)	1.81	0.13	1.25	0.10
xgb2-P10	(6, 0, 8, 0.05)	1.67	0.16	1.22	0.12

Analytical Model

We also developed an analytical model to issue warnings for potential ice jam event in Lake Saint-Pierre. The model is threshold value based. By analyzing the air temperature, wind and snow during the three ice jam events, we concluded that three conditions were met prior to the occurrence of the ice jams:

- 2 °C below normal air temperature,
- prevailing northeast wind, and
- buildup of snowfall.

As shown in Figure 7 prior to the occurrence of all three ice jams, air temperature was below historical means for at least two days, northeast wind above 10 m/s was observed, and snowfall was present as well.

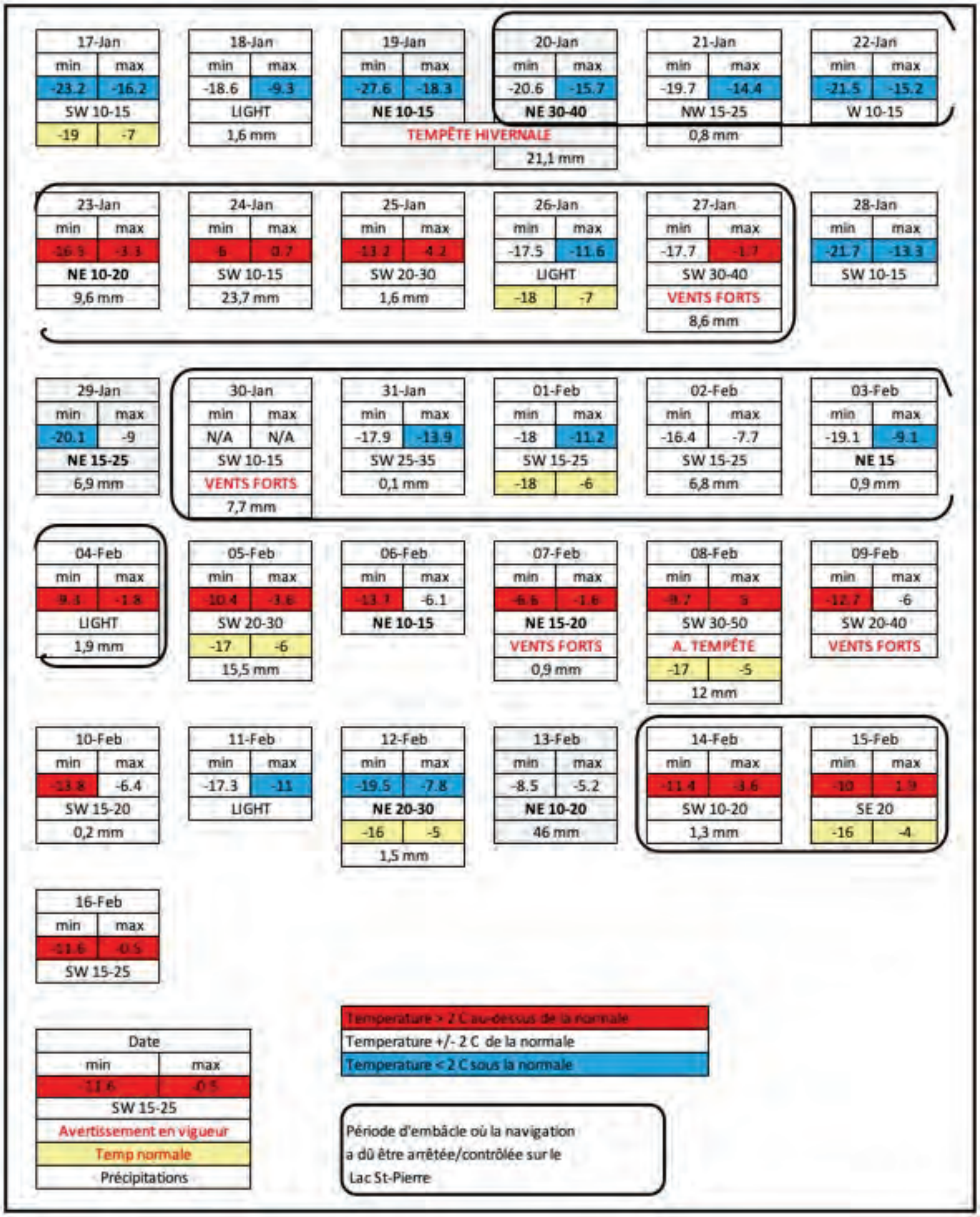

Figure 7. Air temperature, wind, and snow conditions throughout the three ice jam events in Lake Saint-Pierre in 2019 (Canadian Coast Guard).

Case Study – 2019 Ice Jam Events

The final XGBoost model was applied to study the ice concentrations from December 15, 2018 to March 15, 2019. The model was run with daily meteorological hydrometric data. An overall ice concentration was calculated by taking weighted averages of the predicted ice concentrations at the 10 points. The weights were proportional to the inverse of the MAE errors shown in Table 6. A threshold ice concentration of 8.5 was chosen to denote potential ice jam risk. In total, there are 6 days with ice concentration above this threshold value: January 13, 14, 21 and 31, February 19, and March 7. This model successfully predicted the occurrence of two of the three ice jams (i.e., January 21 and 31) as shown in Figure 8.

The analytical model was run to get potential ice jam warnings for the same time period. Two alerts were generated on two separate days: January 22 and February 20. The first warning corresponds to the first actual ice jam event (Figure 8).

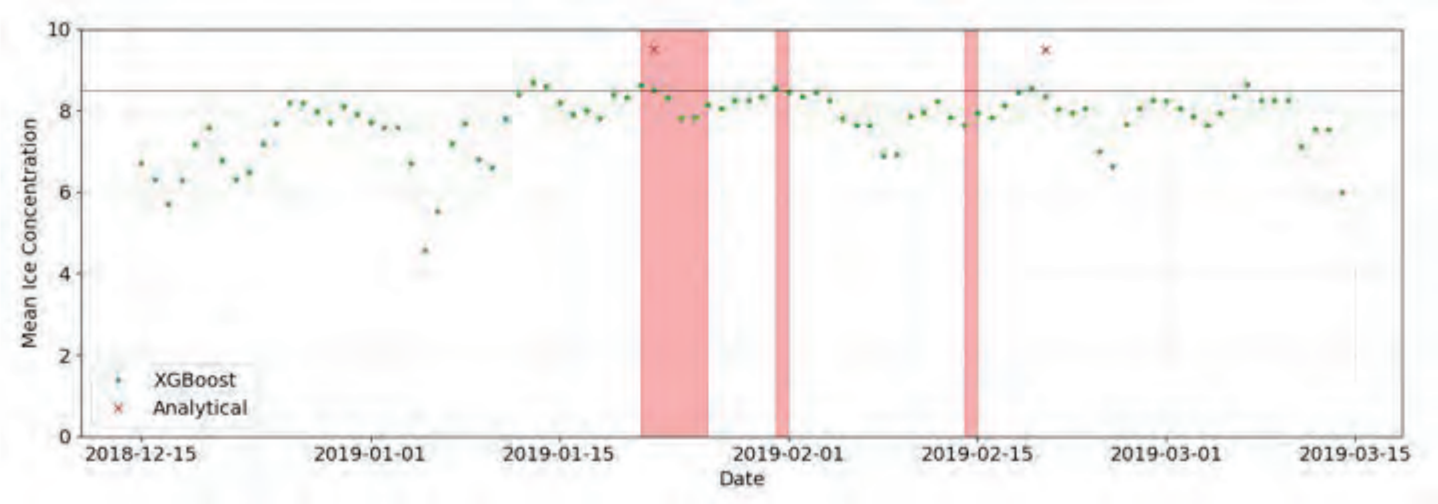

Figure 8. Case study of the 2018-2019 winter ice conditions: Ice concentration prediction with the final XGBoost model (green dots), ice jam warnings issued by the analytical model (red crosses). The red shaded areas correspond to the three ice jam events in 2019.

GSTS
Navigate Tomorrow Today

Next steps

An ice jam in Lake Saint-Pierre is an extreme event and does not happen frequently. There are limited records of past ice jam events in this region; also, there is limited historical ice and environmental data in this region. It is therefore almost impossible to develop an AI model to predict the occurrence of ice jam in this region. In this work, instead of predicting the occurrence of ice jams we developed machine learning models to predict ice concentration at 10 points along the navigation channel in Lake Saint-Pierre.

The best model was a regression model developed with XGBoost. This model used 32 features, including month, day, air temperature, wind speed, wind direction, water levels upstream and near Lake St. Pierre; lagged daily values of the meteorological and hydrometric variables were used. We also developed an analytical model by analyzing meteorological conditions throughout the ice jam events in 2019. Three criteria were observed prior to the occurrence of the ice jams: below normal air temperature, prevailing northeast wind, and buildup of snowfall. This model is supplementary to the XGBoost regression model and provides early warnings for potential occurrence of ice jams.

Next steps of this project include:

- conduct pipeline engineering and deploy models for operation, and
- integrate model results into OCIANA™ (e.g., development of an UI).

References

Canadian Coast Guard, 2012: Ice Navigation in Canadian Waters.

Morse B., M. Hessami, and C. Bourel, 2003: Characteristics of ice in the St. Lawrence River, Can. J. Civ. Eng. 30: 766–774.

Scalabrini, P. and B. Morse, 2021. An early Warning Ice Management System for the Saint Lawrence River: Increasing the Reliability of Winter Navigation to the Port of Montreal, 21st Workshop on the Hydraulics of Ice-Covered Rivers, Saskatoon, Saskatchewan, Canada, August 29-September 1, 2021.