A Deep Learning Model for Lake Ice Cover Forecasting

Samuel Johnston
Claude Duguay
Justin Murfitt

H2O Geomatics & University of Waterloo (Canada)

7th Workshop on "Parameterization of Lakes in Numerical Weather Prediction and Climate Modelling"
20-22 November 2024 - Milan, Italy

Background

- The presence (or absence) of ice cover and its extent affect socio-economic/recreational activities, climate and weather events (e.g., lake-effect snowfall, thermal moderation) locally and regionally
- EO observations of lake surface state (e.g., ice cover/open water and surface temperature) from multiple satellite missions can help improve the prediction of weather events from NWP models

Data Source: NASA Terra/MODIS image (11 February 2016) showing snow bands over the Great Lakes of North America.

Background

Lake-effect snowfall

November 2022 Great Lakes historic winter storm!

- 17-20 Nov. 2022: ~200 cm of snow fell in Buffalo area
- At least four deaths

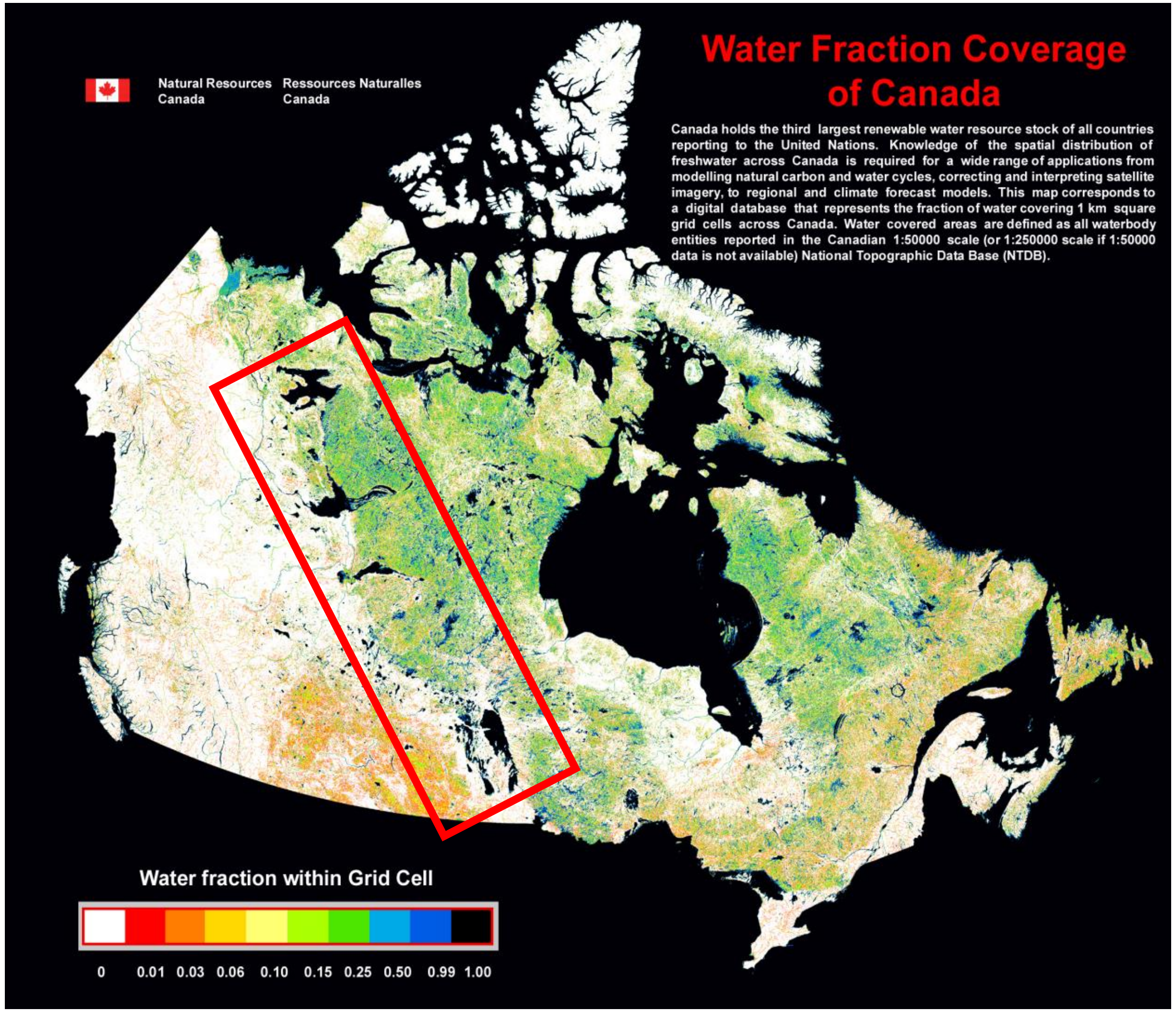
19-23 December 2022

- The storm lasted four days ~132 cm recorded in the Buffalo region — most fell over two days.
- 37 deaths reported (29 in City of Buffalo)
- Faced risk of flooding with rising temperatures and rain in forecast

Background

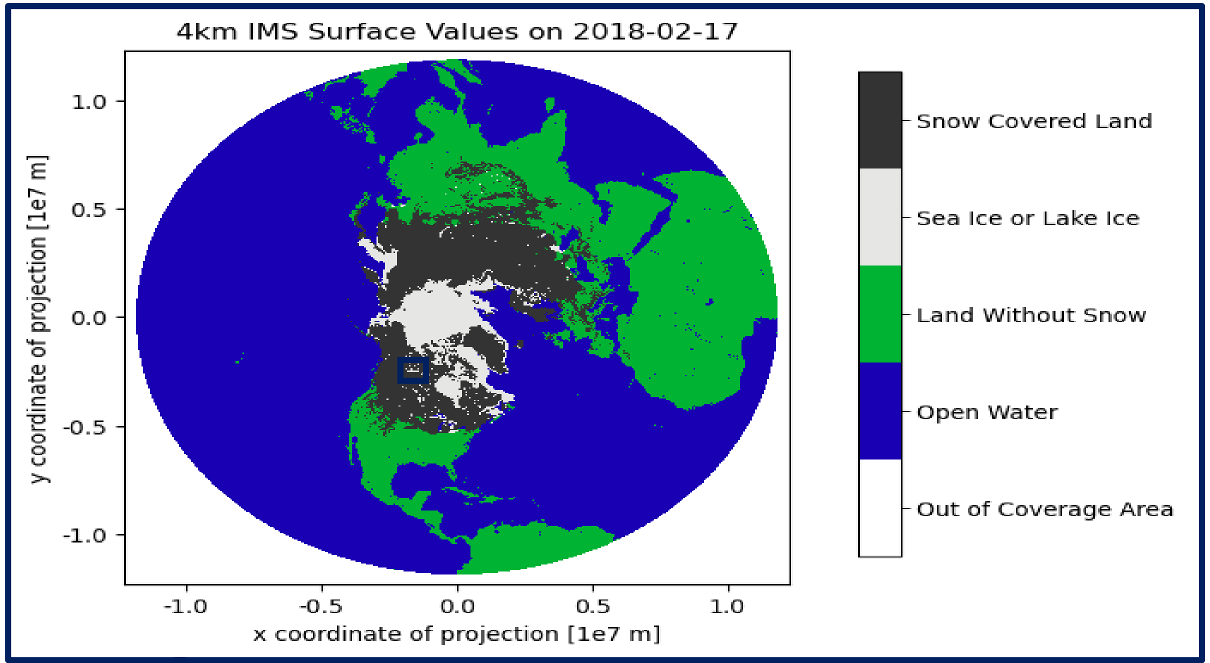
- Lakes comprise a significant proportion of the land surface at northern latitudes
- Existing lake models used as lake parameterization schemes in NWP and climate models are one-dimensional
- With recent advances in machine learning and the availability of longer historical satellite data records, **we *initiated a project on the development of a deep learning model for LIC forecasting (contemporary and future weather/climate conditions)***

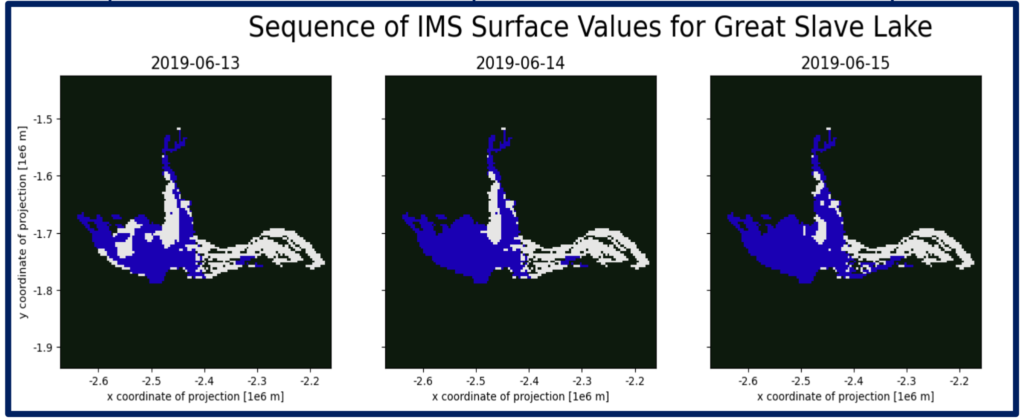
Johnston et al. (in preparation)

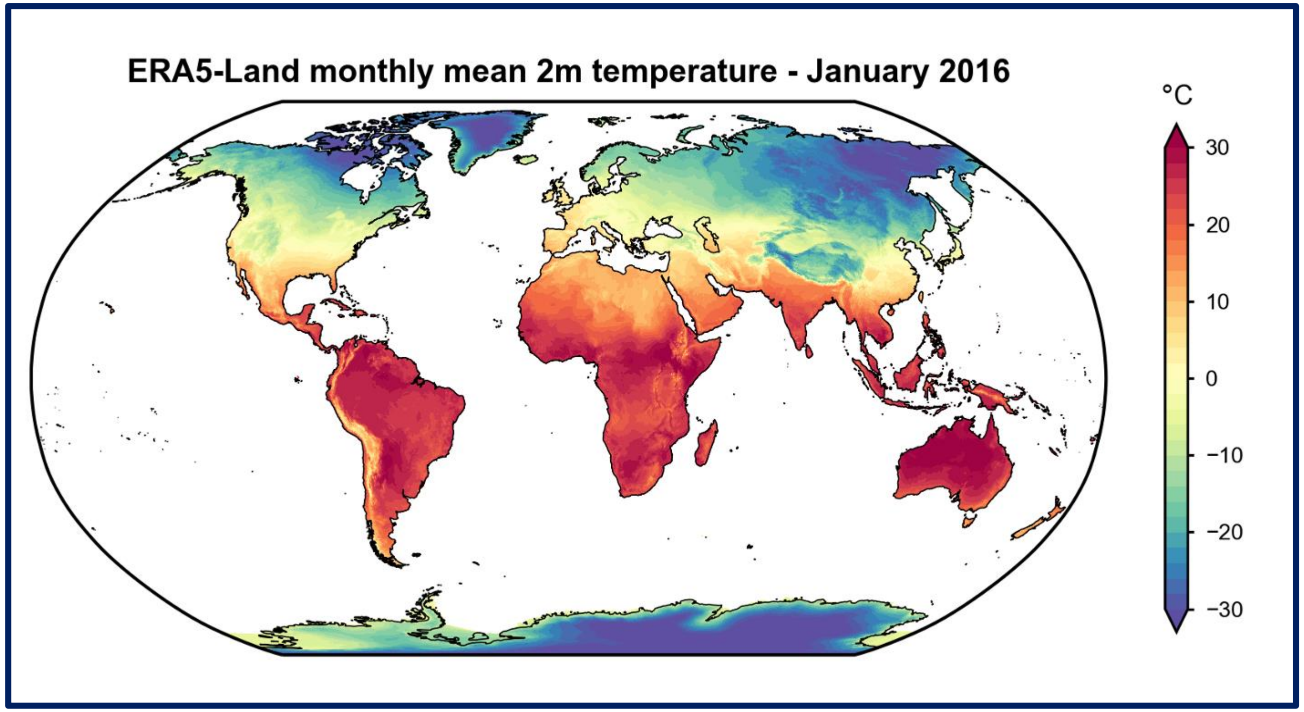


LIC is a thematic product of Lakes as an ECV

Data: IMS (training and validation) and ERA5



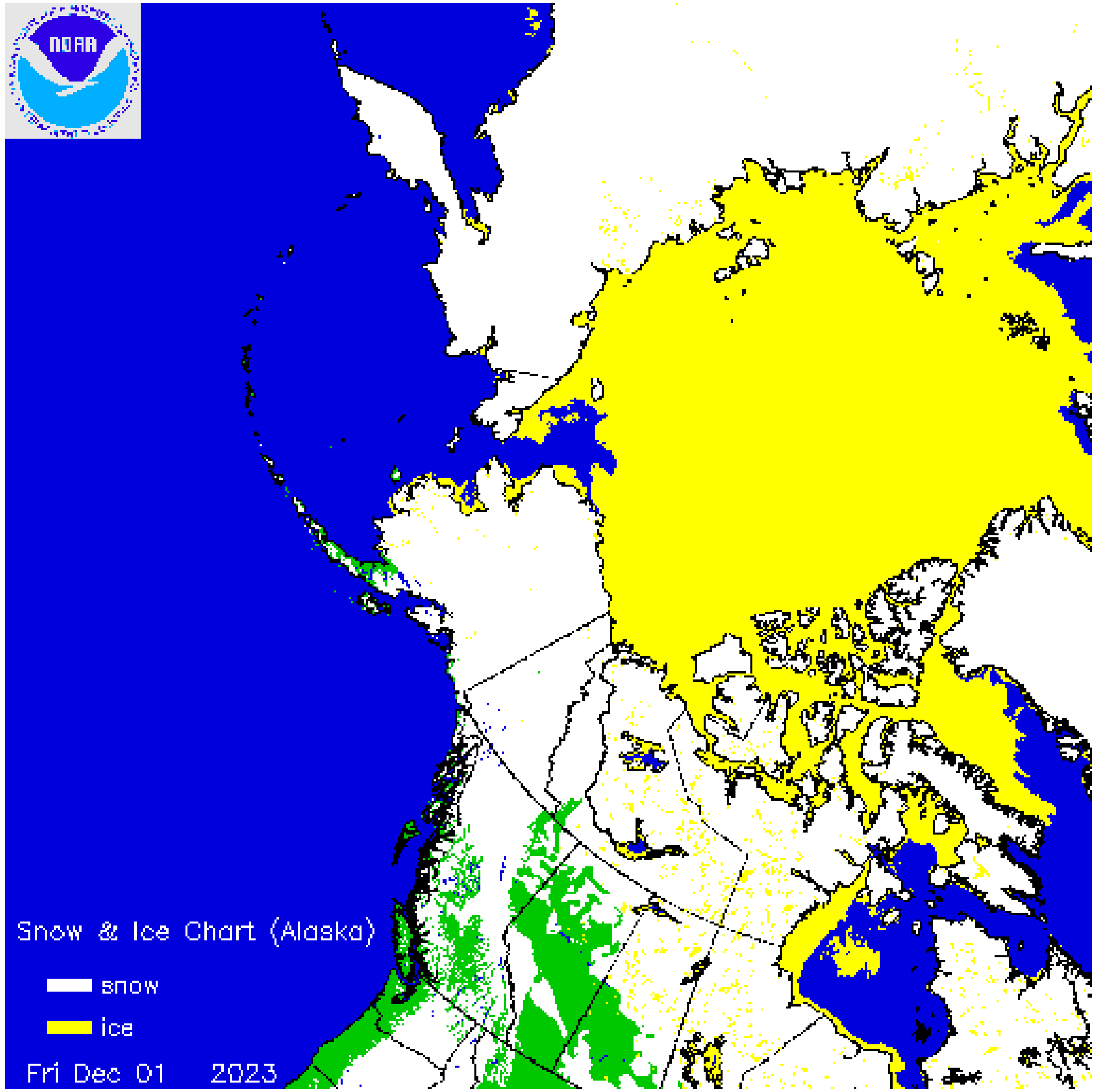


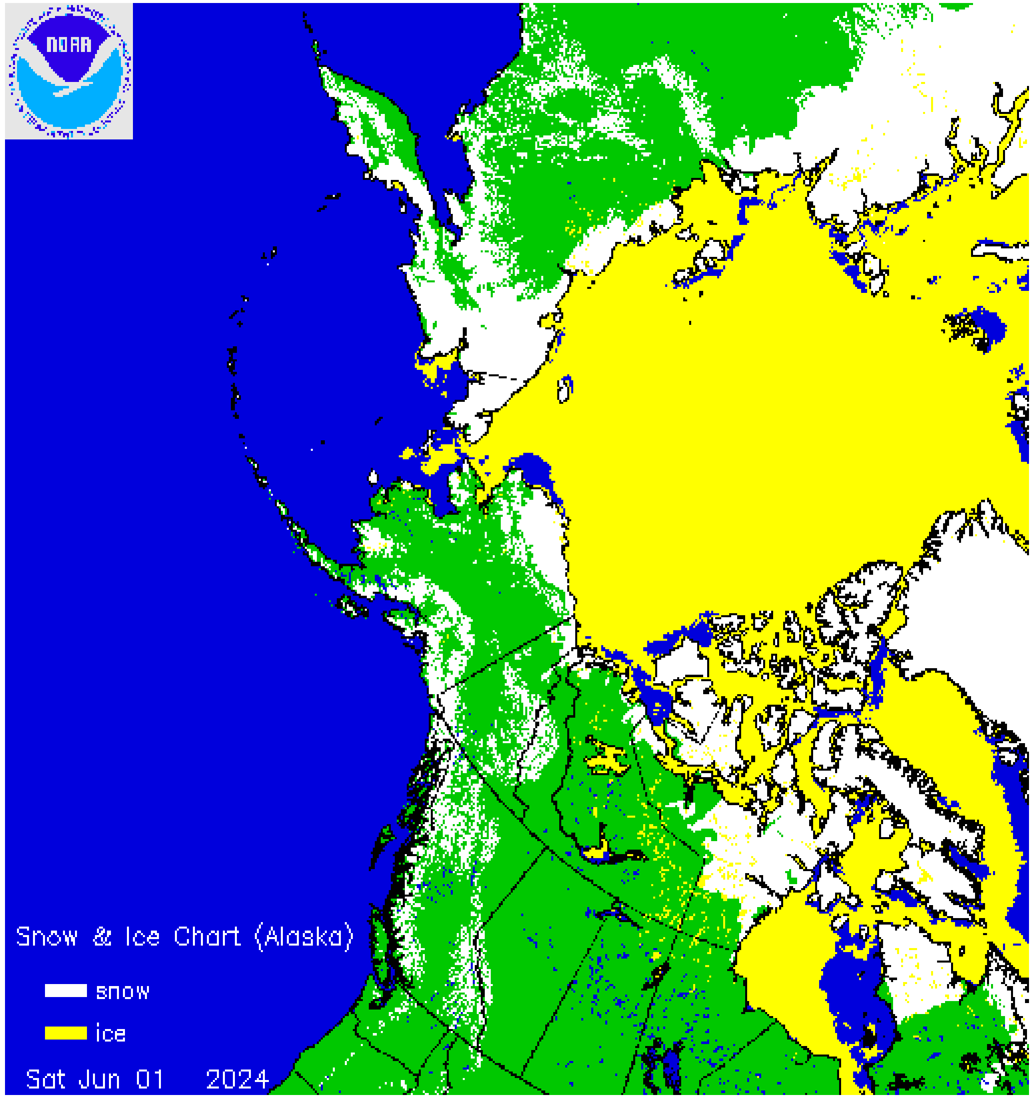


ERA5 Processing:

1. Calculate additional variables
2. Aggregate from hourly to daily
3. Reproject to match IMS
4. Interpolate onto 4-km IMS grid (Nearest Neighbours)

IMS Examples: Freeze-up and break-up (2023-2024)

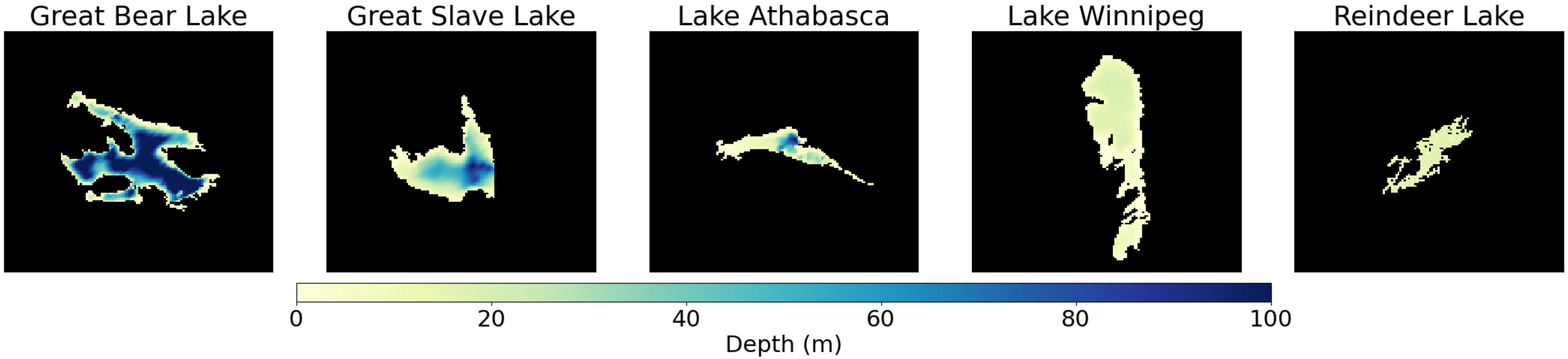




Data: ERA5

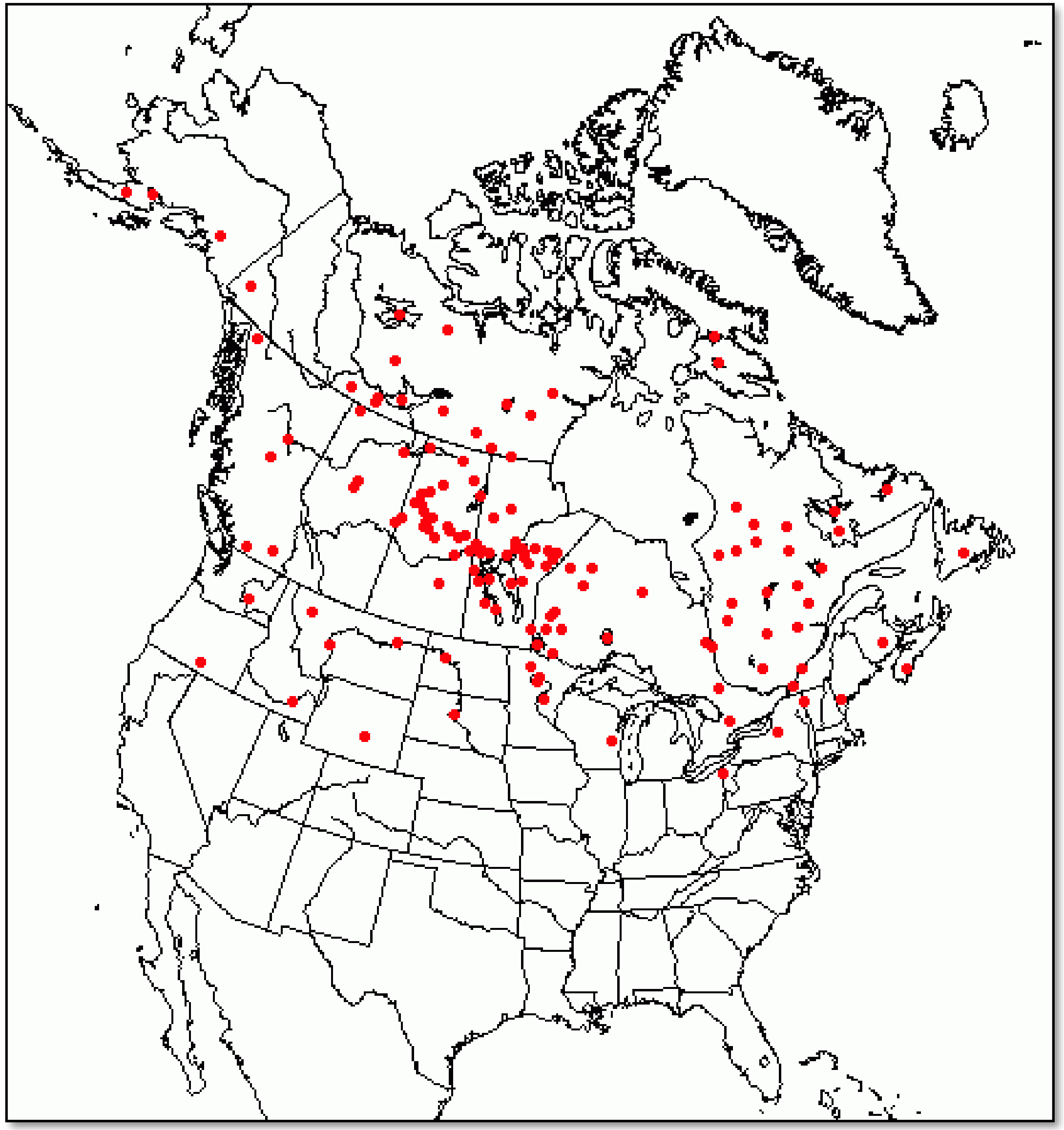
Variable Name	Source
Temperature_2m (Celsius)	ERA5-Land: Temperature at 2m (K)
Surface_solar_radiation_downwards_sum (Jm^{-2})	ERA5-Land: Surface Solar Radiation Downwards (Jm^{-2})
Wind_speed_10m (ms^{-1})	ERA5-Land: U-component of wind at 10m (ms^{-1}) and V-component of wind at 10m (ms^{-1})
Total_precipitation_sum (m)	ERA5-Land: Total Precipitation (m)
Total_cloud_cover (fraction)	ERA5: Total Cloud Cover (fraction)
Accumulated_freezing_degree_days (Celsius)	ERA5-Land: Temperature at 2m (K)
Accumulated_thawing_degree_days (Celsius)	ERA5-Land: Temperature at 2m (K)

Data: Lake Bathymetry



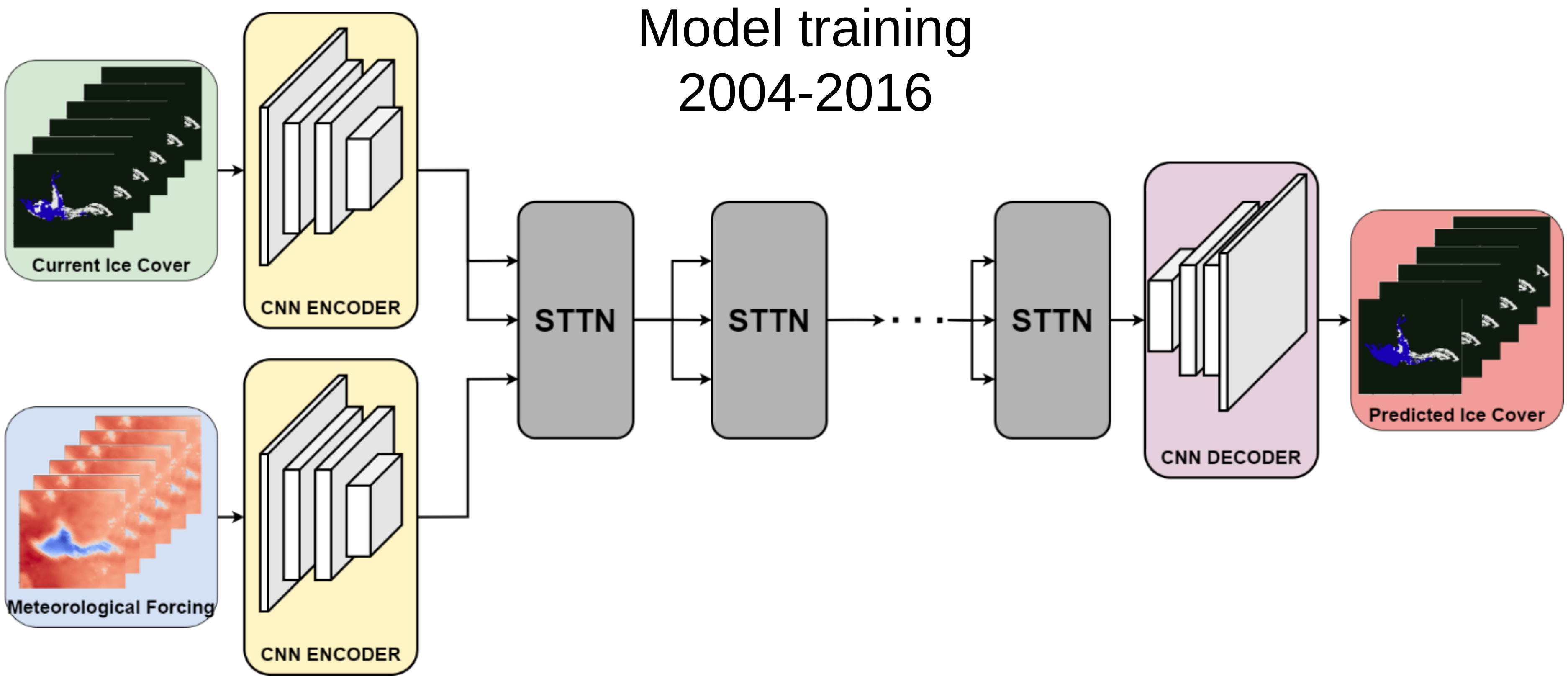
Toptunova, O., M. Choulga, and E. Kurzeneva, 2019. Status and progress in global lake database developments. Adv. Sci. Res., 16, 57–61, https://doi.org/10.5194/asr-16-57-2019

Data: Canadian Ice Service (validation)



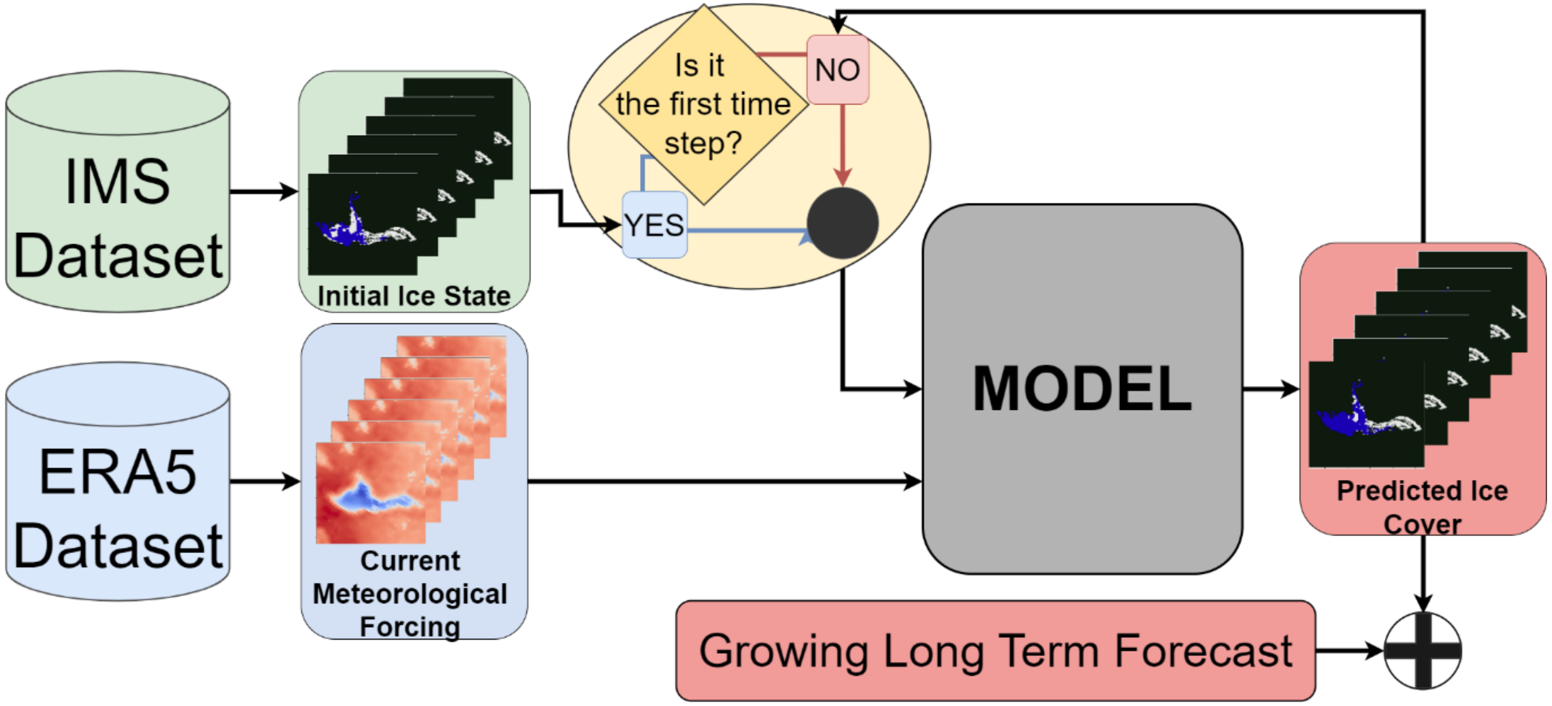
- Weekly ice fraction (concentration; 0 to 10 tenths) from visual interpretation of radar and optical imagery by ice analysts
- Single ice fraction value reported per lake (ca. 140 lakes across Canada and the northern US, excl. Laurentian Great Lakes – separate daily product)
- The product is used operationally at ECCC for weather forecasting
- CIS dataset was used for validation of the LIF-DL (Lake Ice Forecasting using Deep Learning) model output

LIF-DL Model: Architecture



Overview of the model which incorporates Convolutional Neural Networks (CNN) and Spatial-Temporal Transformer Networks (STTN) components

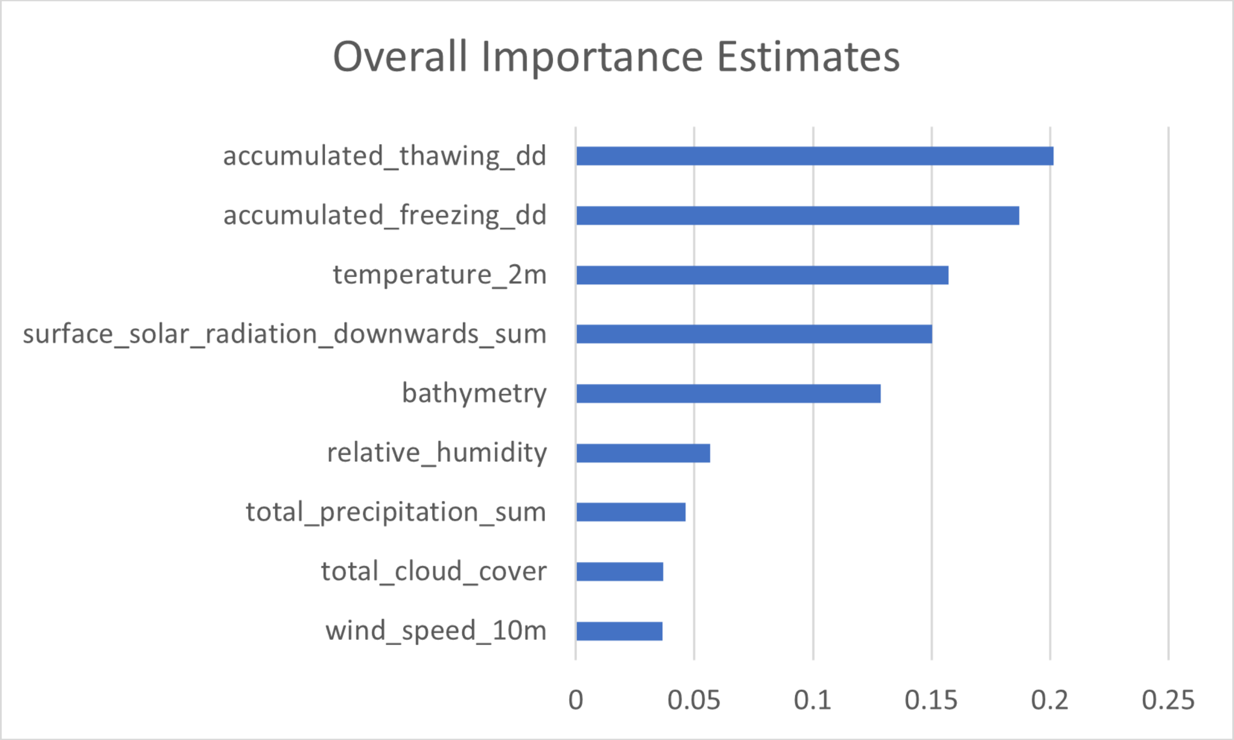
LIF-DL Model: Autoregressive Deployment



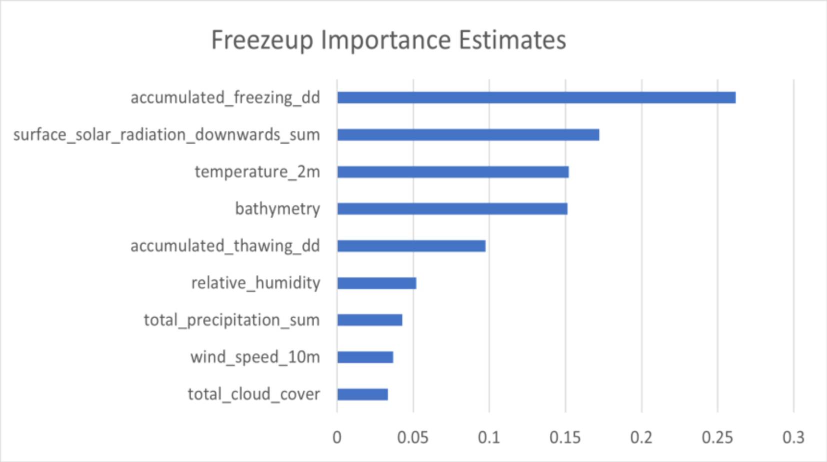
Overview of autoregressive method of producing long-term forecasts using the LIF-DL model. Initial ice states are used to produce the first prediction, after which model predictions are fed back in as input to continue forecasting forward through time

LIF-DL Model: Variable Importance Estimates

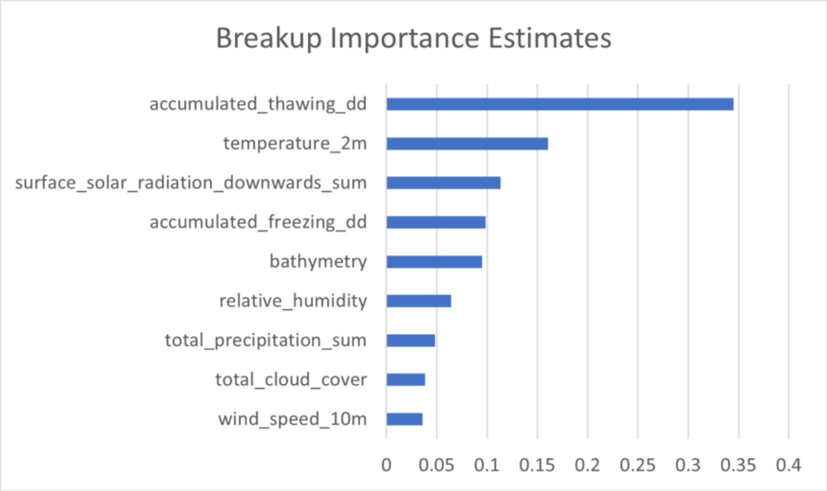
Overall variable importance estimates (across all dates)



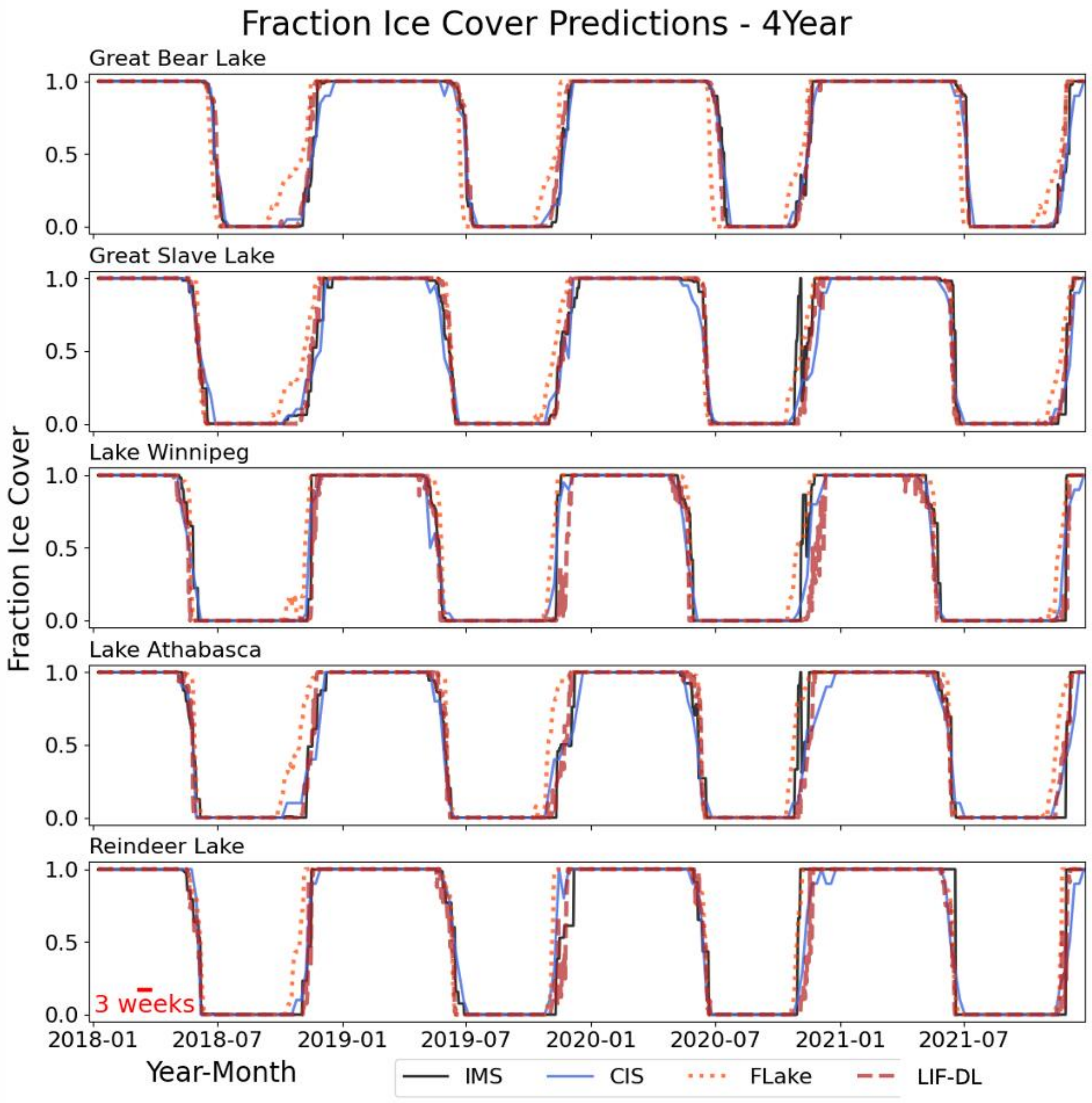
Freeze-up (September to February)



Break-up (March to August)



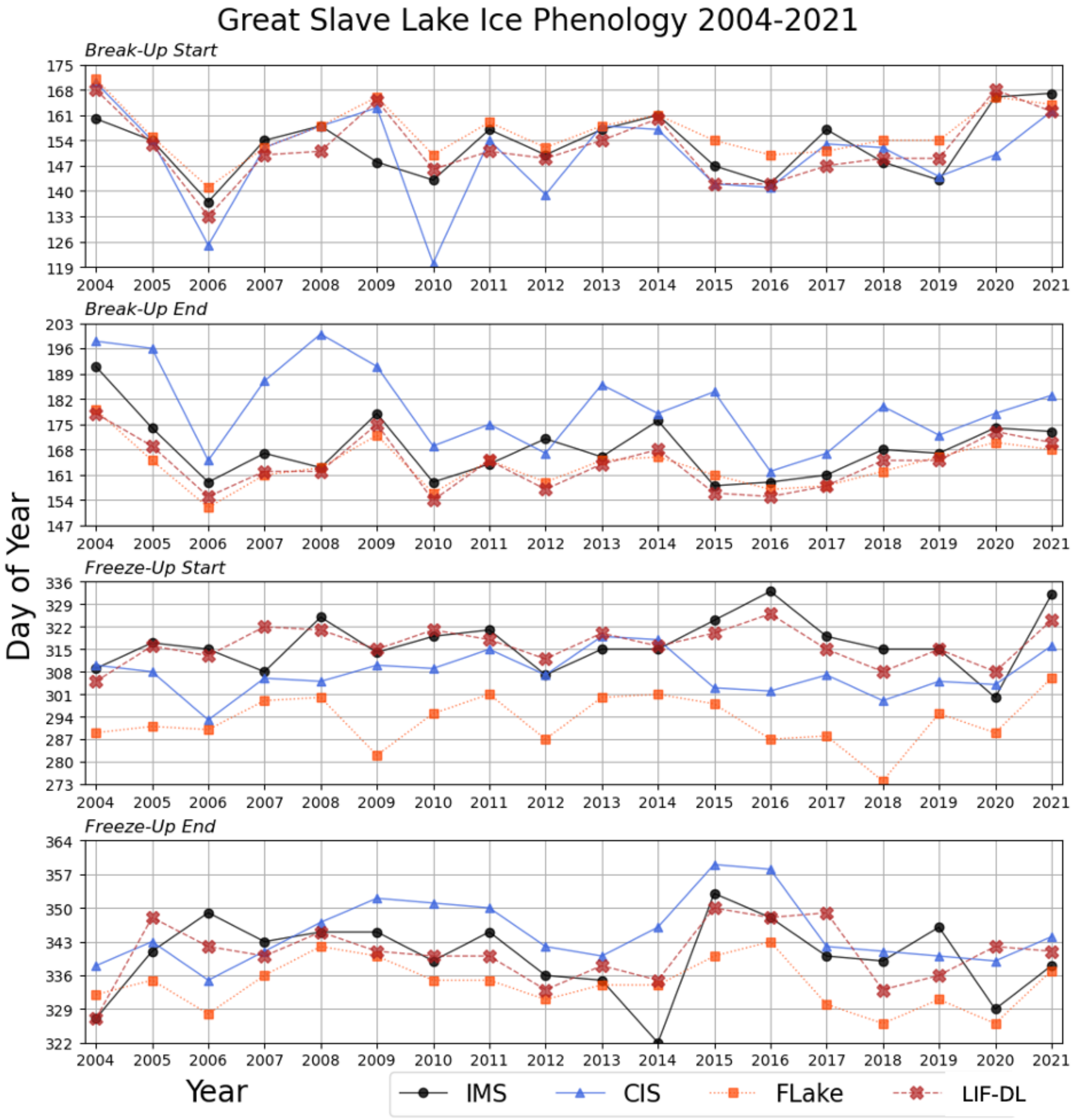
Results: Ice Cover Fraction



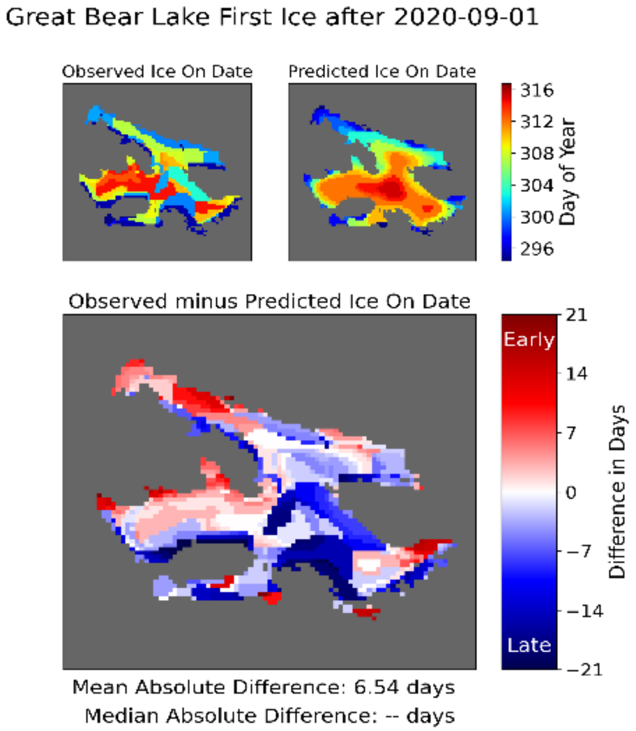
Lake-wide ice cover fraction over the testing (validation) period (2018 – 2021) from observation (IMS and CIS) and models (LIF-DL and FLake)

Comparison	Break-up Start (days)	Break-up End (days)	Freeze-up Start (days)	Freeze-up End (days)
LIF-DL, IMS	6.6	4.8	5.0	5.7
FLake, IMS	7.1	6.1	20.9	10.4
LIF-DL, CIS	8.1	11.3	7.1	7.6
FLake, CIS	7.9	11.8	14.8	14.4

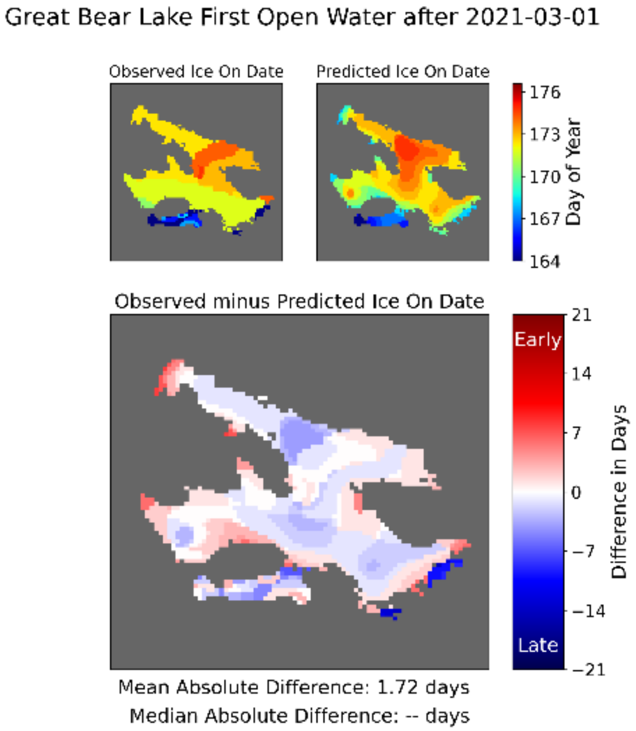
Results: Ice Phenology



Freeze-up start (Great Bear)



Break-up start (Great Bear)



Conclusions

- Deep learning and data-driven approaches have the capacity to:
  - Learn relationships between climate and ice-cover extent
  - Learn spatial patterns of freezing/thawing
  - Forecast over long time periods without significant error accumulation
- Limitations: Dataset quality/bias
  - ERA5 values are diagnostic – they are affected by the ERA5 ice model
  - IMS temporal gaps due to cloud cover – leads to punctuated changes in ice-cover classifications
  - IMS data contains ‘artifacts’ (erroneous pixels) – plan to integrate ESA CCI+ Lakes 1-km LIC gap-filled product in future
  - Bathymetry – for example, the depth of the east arm of Great Slave Lake
- Future work will investigate the incorporation of more physical understanding into the model design, conduct further validation to improve interpretability (more lakes), and use CCI+ Lakes gap-filled product to forecast LIC over ca. 1,500 lakes (possibly more) globally

Thank you for your attention!

Contact: Sam Johnston or Claude Duguay

samuel.johnston@h2ogeomatics.com
or
crduguay@uwaterloo.ca
claude.duguay@h2ogeomatics.com