Enhanced Sea Ice Classification for ICESat-2 ICESat-2 Using Machine Learning

Wenxuan Liu[1], Michel Tsamados[2], Alek Petty[3], Taoyong Jin[1], Weibin Chen[2], Julienne Stroeve[2,4]
Julienne Stroeve[2,4]

1 Wuhan University,
2 University College London,
3 University of Maryland,
4 University of Manitoba

Motivation

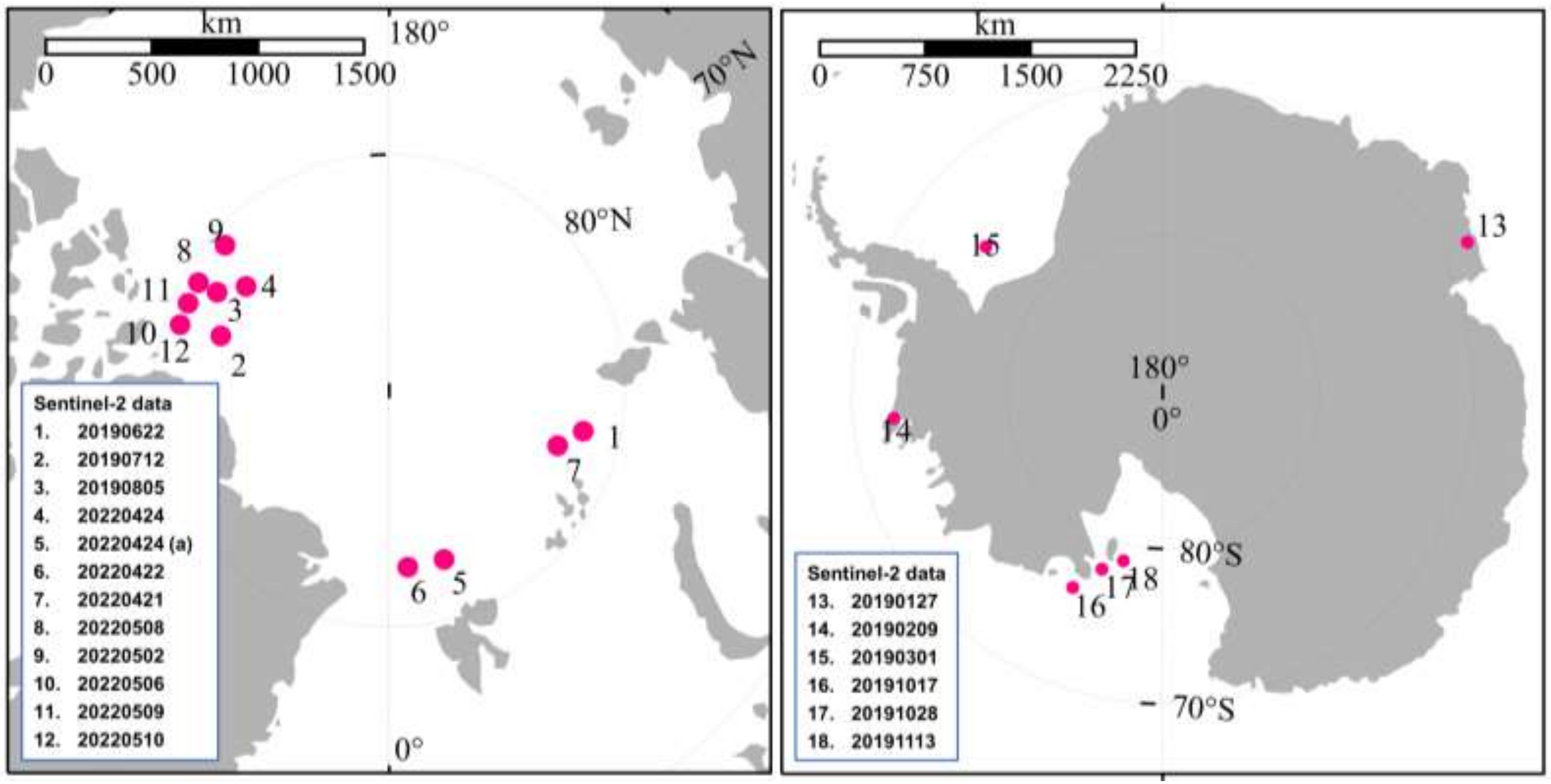
- sea ice freeboard estimate is relying on lead detection (classification) of altimetry

- The classification method of ATL07 is based on a decision tree algorithm with fixed thresholds, along with a local height filter
  - Lack of guidance from coincident imagery (although assessed by Petty et al., 2021)
  - Unreliable summer freeboard/thickness estimate due to uncertainties in surface type classification (ATL07/ATL10)

- Our goal: guided by coincident Sentinel-2 imagery, to improve surface type classification by leveraging both unsupervised and supervised machine learning methods

Data collection — Google Earth Engine

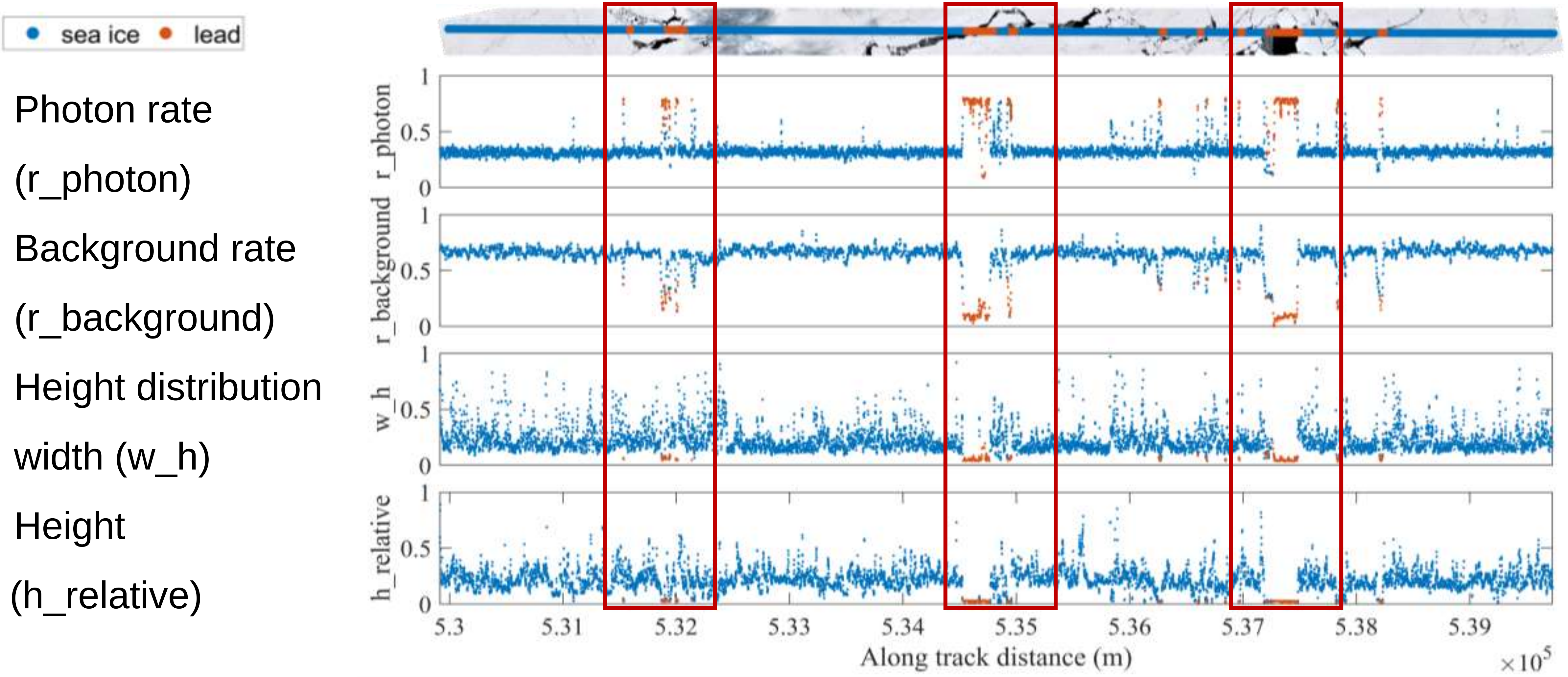
- 18 coincident scenes of ICESat-2 and Sentinel-2 images for Arctic and Antarctic
- Maximum time difference : 30 min → minimize the impact of sea ice drift
- This dataset formed the foundation for our machine learning model



location and acquisition date of coincident Sentinel-2 images

Classification parameters

- Photon rate (r_photon)
- Background rate (r_background)
- Height distribution width (w_h)
- Height (h_relative)



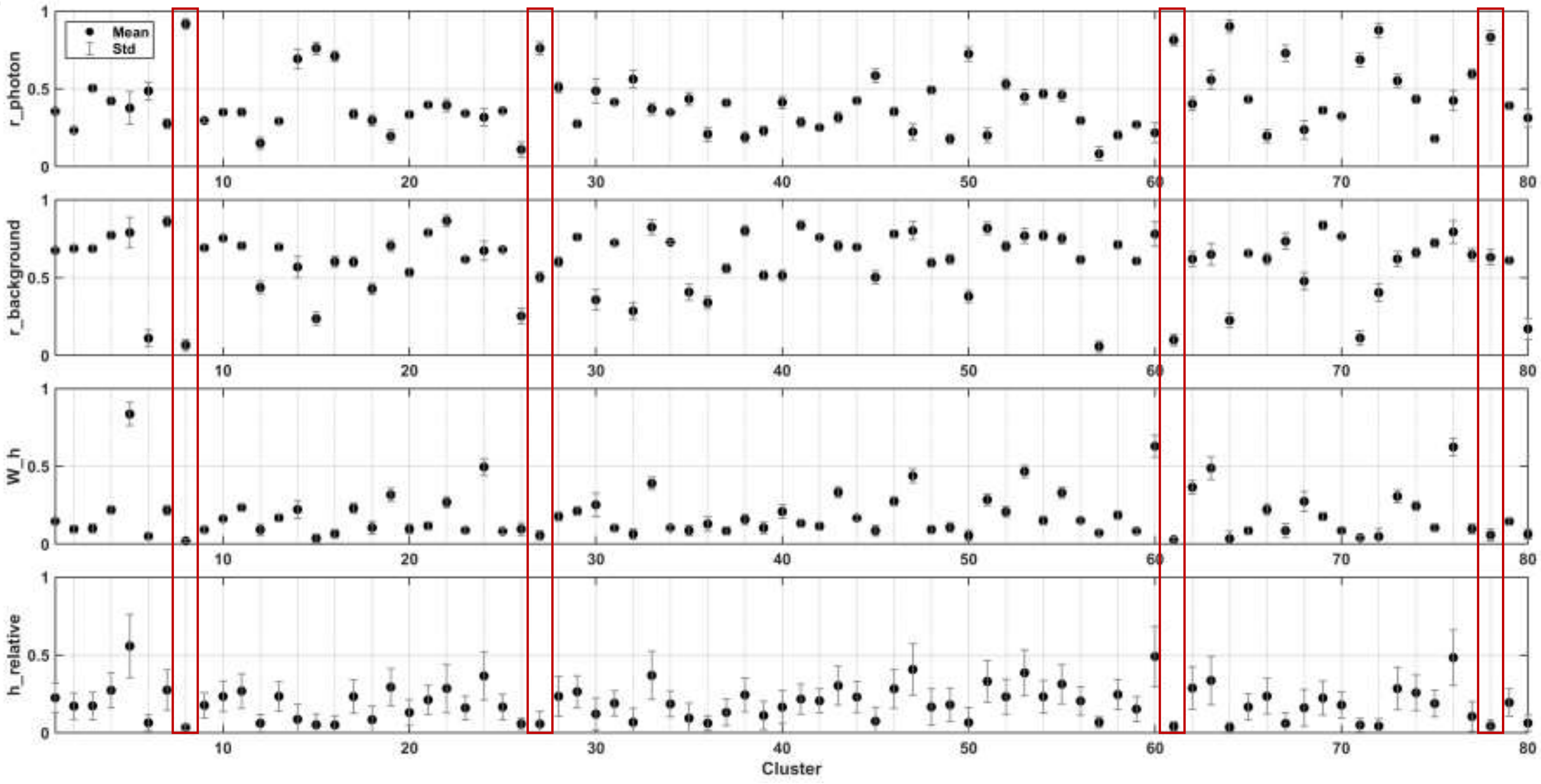
ATL07 ground track overlaid on Sentinel-2 RGB imagery and normalized parameters

Methodology — unsupervised clustering

(1) Use Gaussian Mixture Model clustering combined with visual interpretation to generate training data

- Group ICESat-2 segments into 80 clusters

- We don't know what surface type each cluster corresponds to

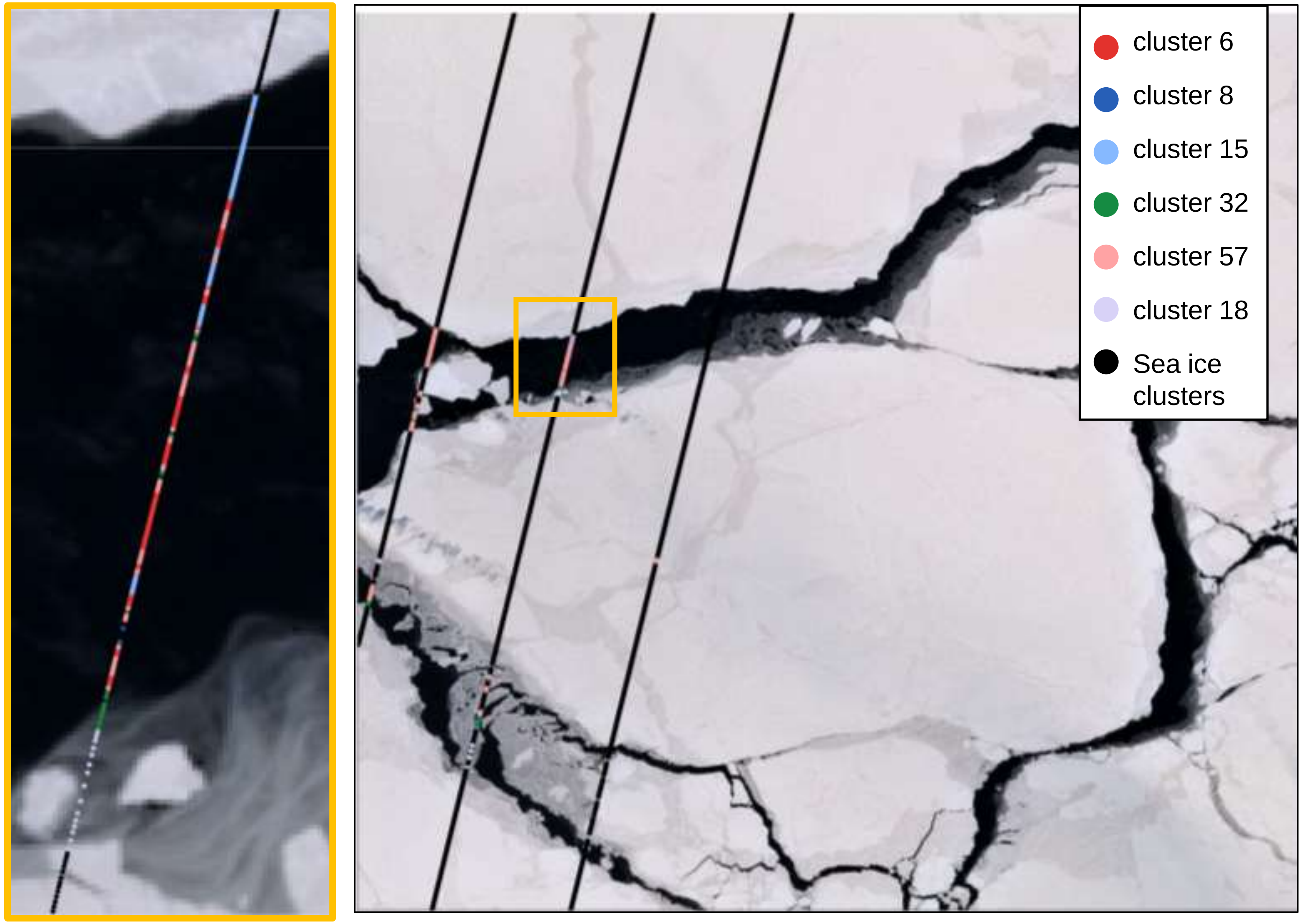


statistics (mean and standard deviation) for each cluster

Methodology — unsupervised clustering

(1) Use Gaussian Mixture Model clustering combined with visual interpretation to generate training data

- We overlaid all the cluster results on the coincident Sentinel-2 images to assign a certain surface type for each cluster



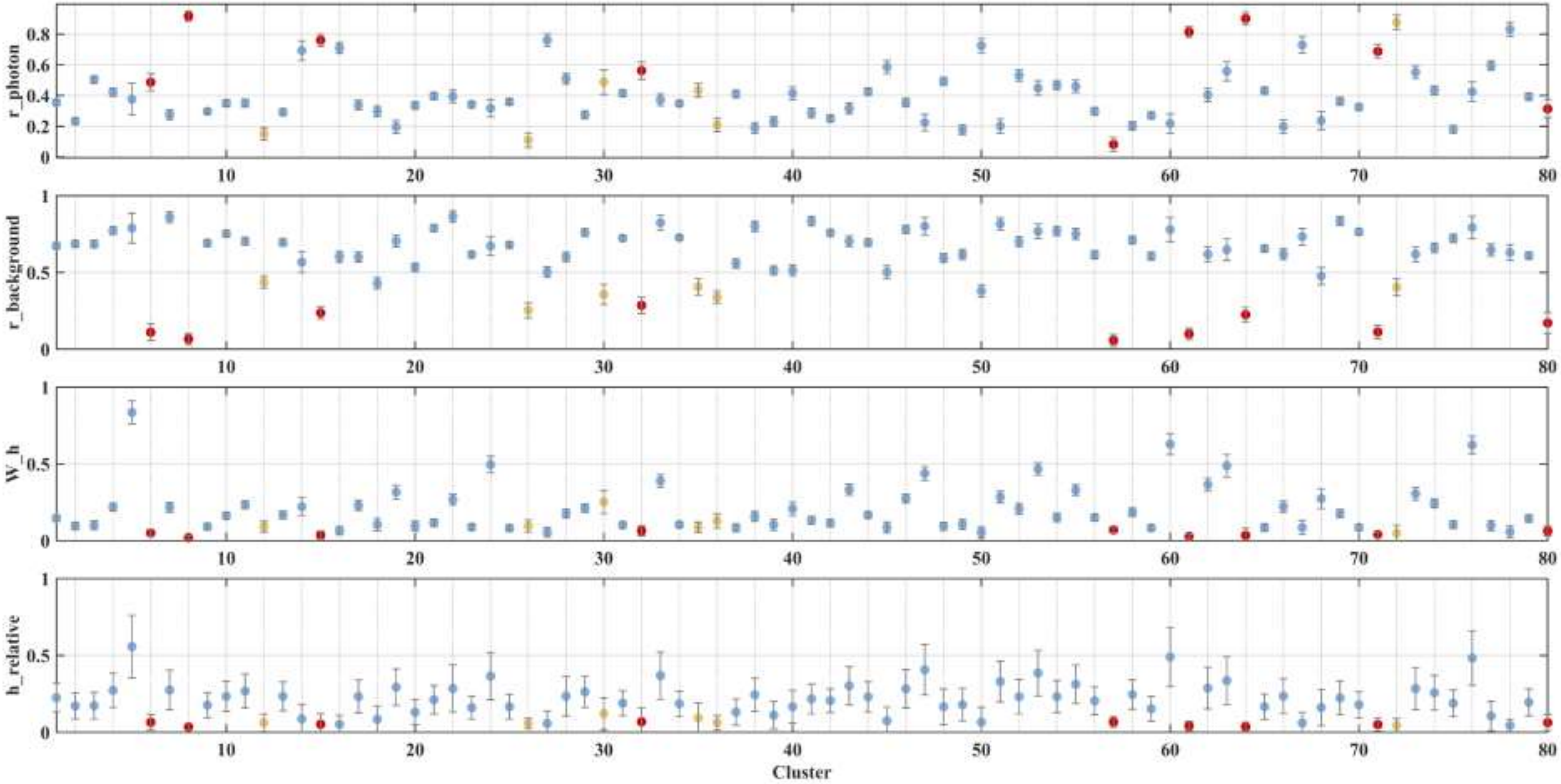
Methodology — unsupervised clustering

(1) Use Gaussian Mixture Model clustering combined with visual interpretation to generate training data

Sea ice

Gray ice

Lead



- 717,009 segments from strong beams
- 702,843 segments from weak beams

Methodology — supervised classifier

(1) Use Gaussian Mixture Model clustering combined with visual interpretation to generate training data

(2) Train a K-Nearest Neighbor (KNN) classifier for classification

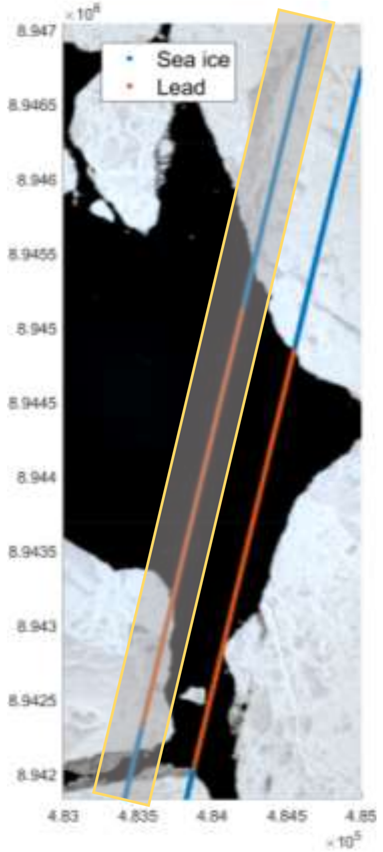
Methodology — external validation

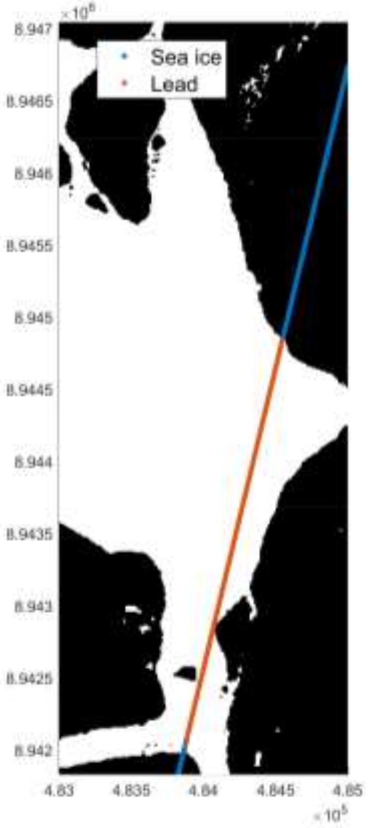
(1) Use Gaussian Mixture Model clustering combined with visual interpretation to generate training data

(2) Train a K-Nearest Neighbor (KNN) classifier for classification

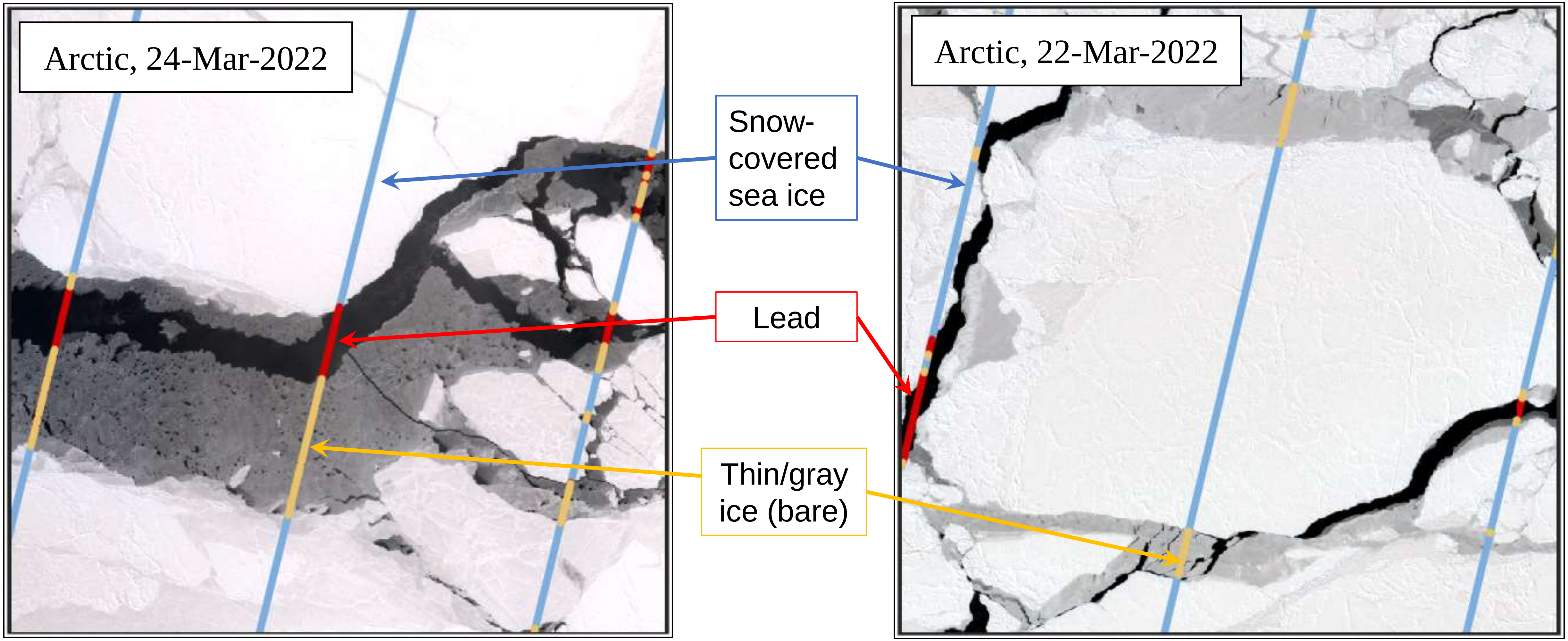
(3) The classification results are compared with “ground-truth” data from Sentinel-2 imagery

- Shift original ICESat-2 tracks (in yellow box) to align well with Sentinel-2 images
- Convert RGB image to binary image (0 for lead and 1 for non-lead)
- Generate independent validation data (each validation segment is assigned as either lead or non-lead based on the majority label of the five closest pixels)



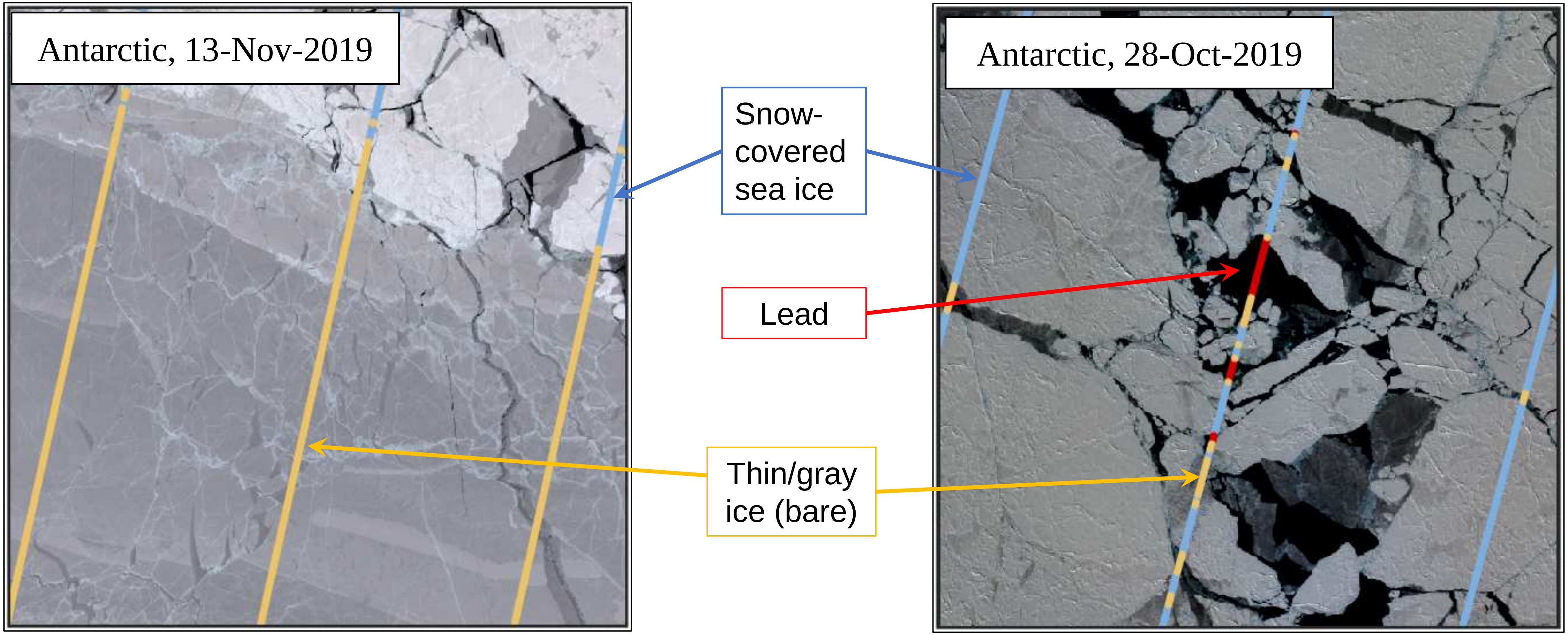


Result — visual comparison



Our method offers a more detailed surface classification that includes an additional category for gray/thin ice

Result — visual comparison



Our method offers a more detailed surface classification that includes an additional category for gray/thin ice

Result — lead detection performance

- Non-lead type includes : snow-covered sea ice and thin/gray ice (bare)
- Classifying a non-lead segment as a lead segment is more problematic than the reverse

Strong beam		Ground truth from Sentinel-2			Precision
		Non-lead	Lead	All	
Our surface type classification	Non-lead	685,830	2,507	688,337	99.6%
	Lead	378	28,294	28,672	98.6%
	All	686,208	30,801	717,009	
Recall		99.9%	91.8%		**99.7%**

Weak beam		Ground truth from Sentinel-2			Precision
		Non-lead	Lead	All	
Our surface type classification	Non-lead	672,276	2,898	675,174	99.7%
	Lead	692	26,977	27669	97.5%
	All	672,968	29,875	702,843	
Recall		99.9%	90.3%		**99.4%**

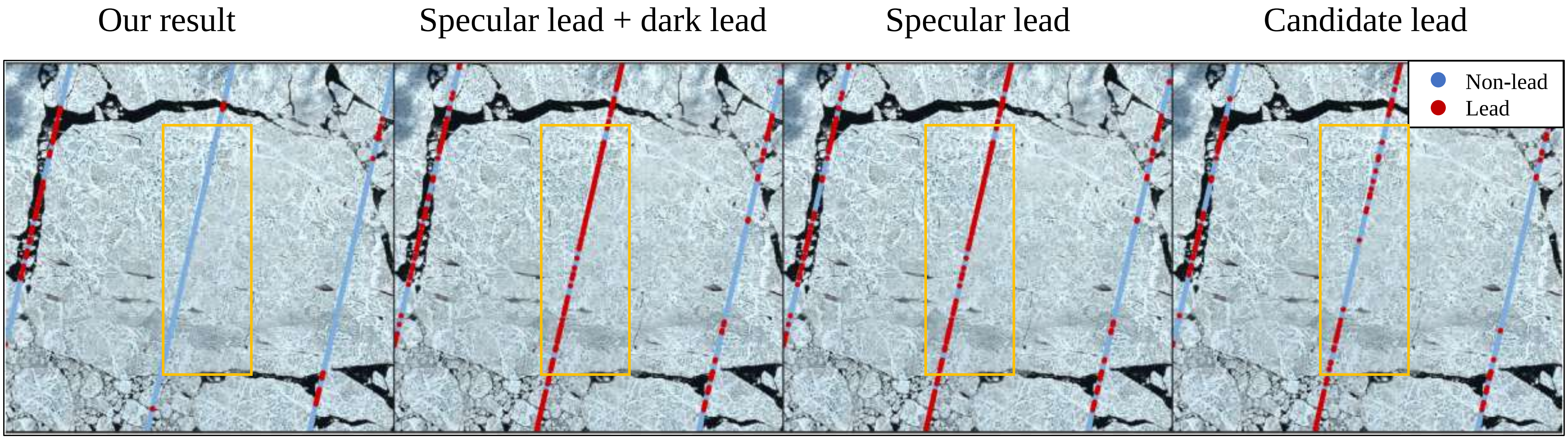
Result — comparison with ATL07 (summer)

- Each ATL07 segment is assigned a type by decision tree algorithm：

  Specular lead (specular lead/pond mixture), dark lead (dark lead/pond mixture), ice (pond/ice mixture)

- Segments are further classified by local height filter:

  Candidate lead (used to derive the local reference sea surface height and freeboard), ice



Although our method also can not include a specific surface type for melt pond, it can be excluded from the lead type, contributing to more reliable local sea surface height and sea ice freeboard retrieval.

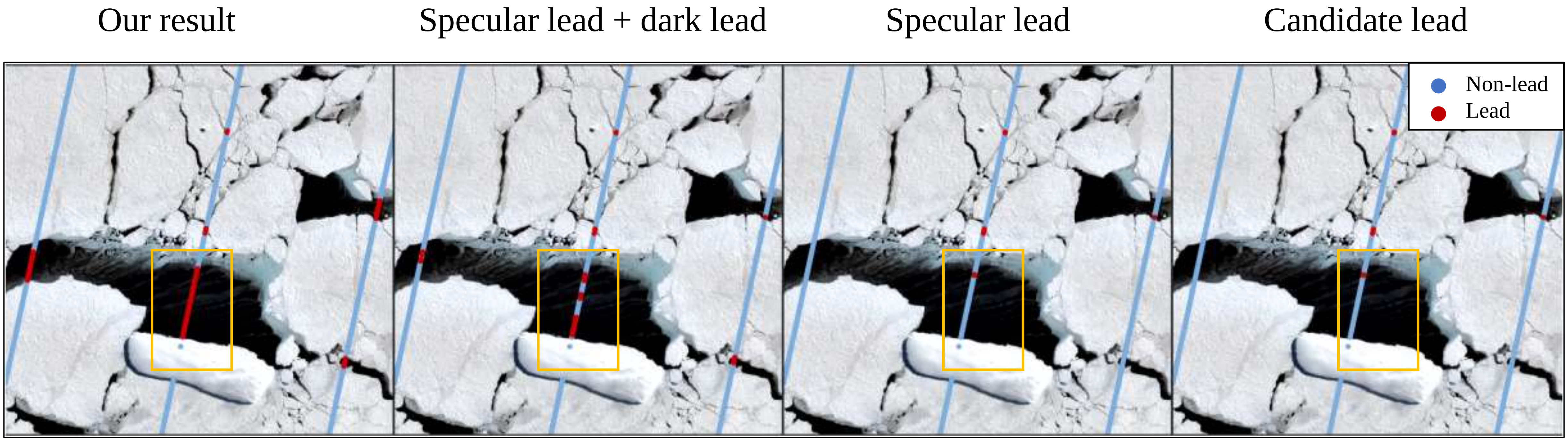
Result — comparison with ATL07 (winter)

- Each ATL07 segment is assigned a type by decision tree algorithm：

  Specular lead (specular lead/pond mixture), dark lead (dark lead/pond mixture), ice (pond/ice mixture)

- Segments are further classified by local height filter:

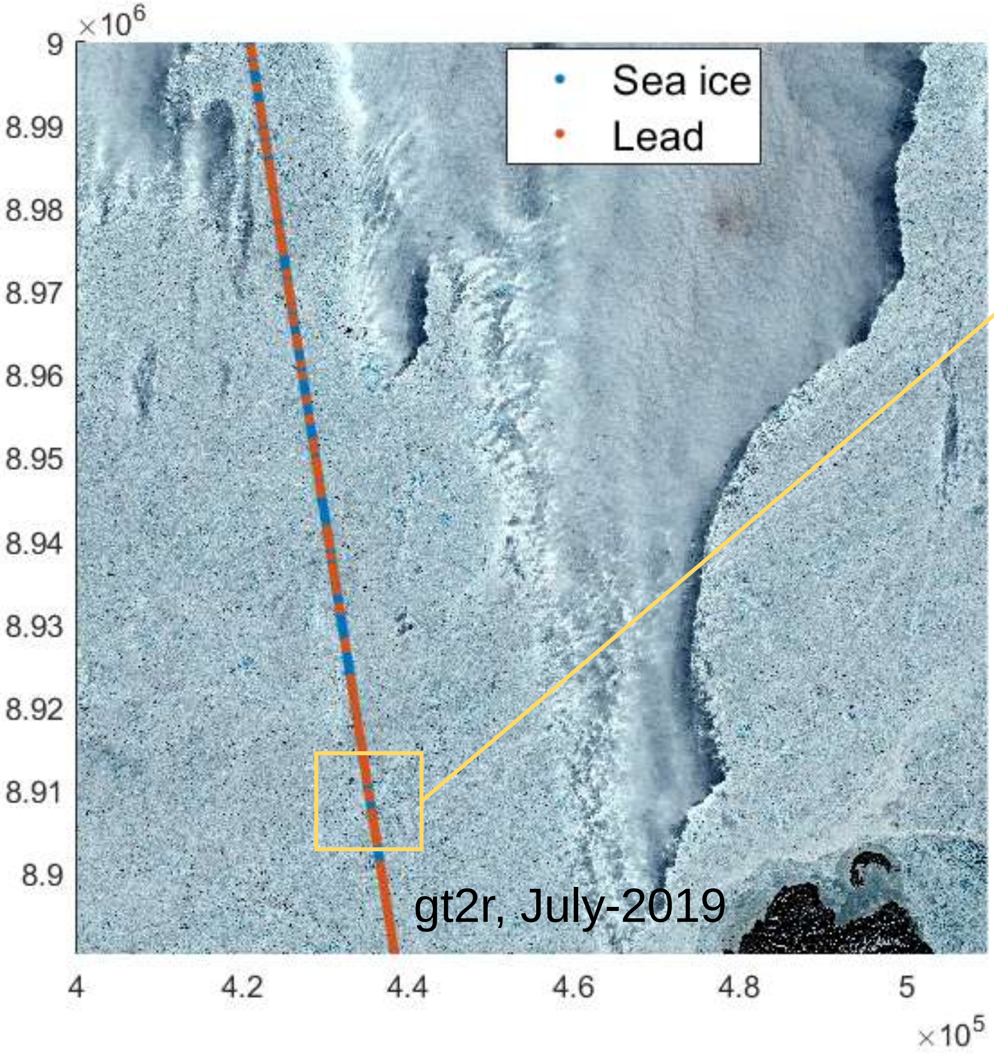
  Candidate lead (used to derive the local reference sea surface height and freeboard), ice

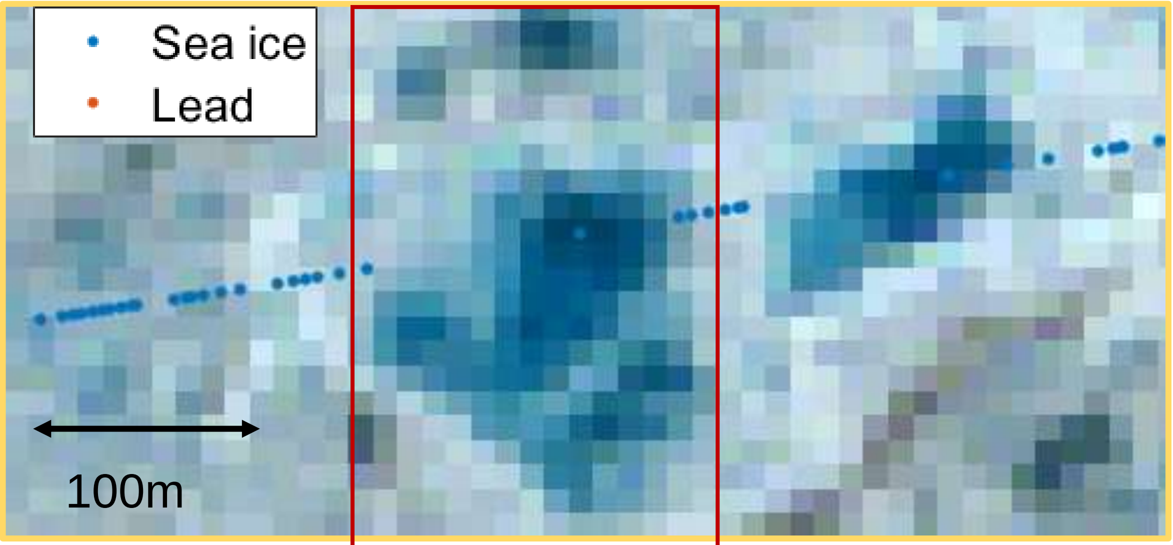


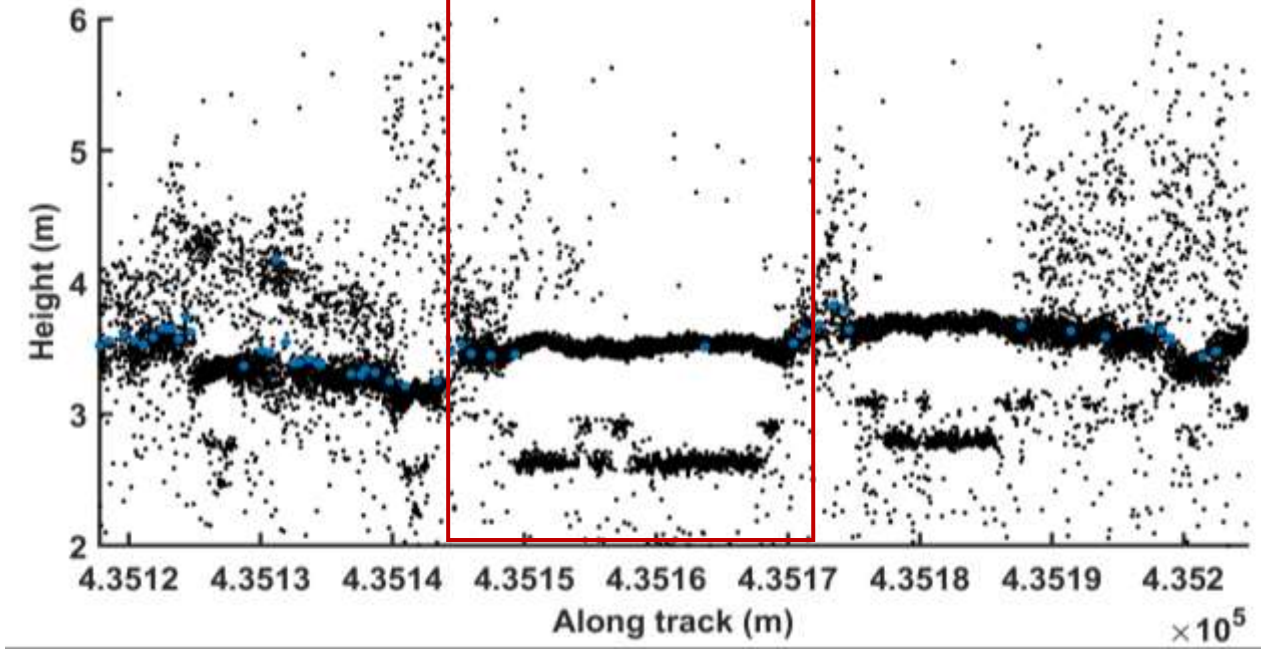
Compared to ATL07, our method can identify more leads and open water

Variable ponds in ATL07

(1) No valid data

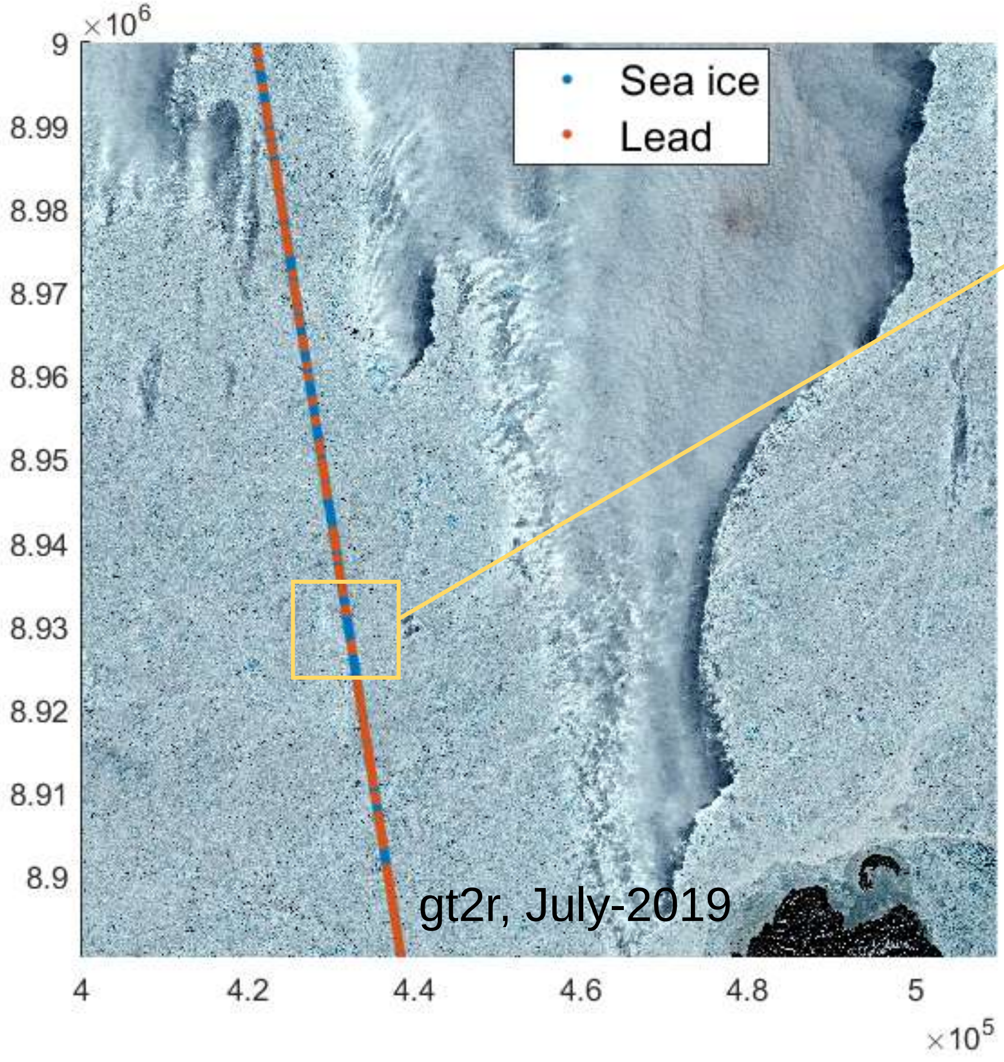


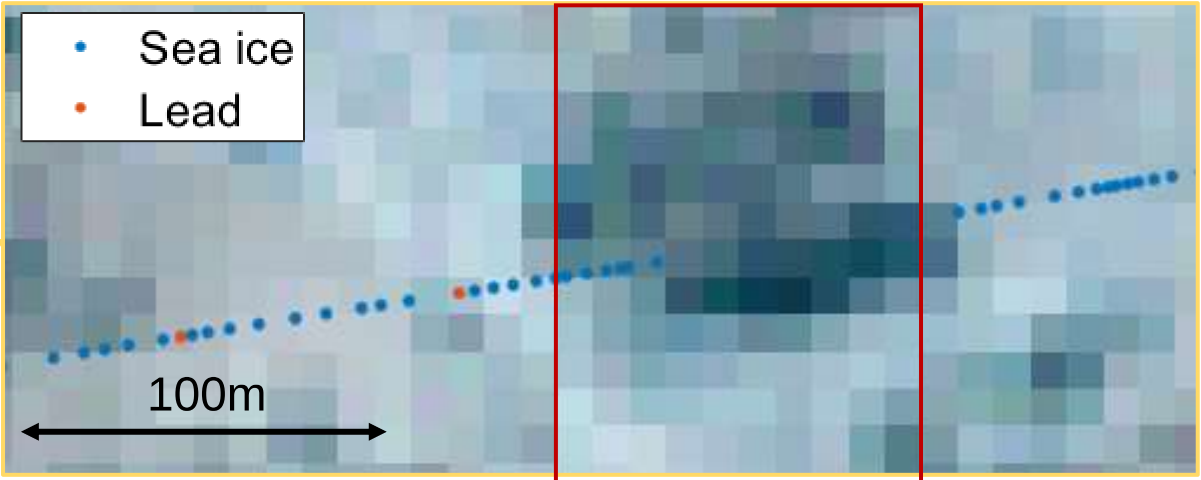


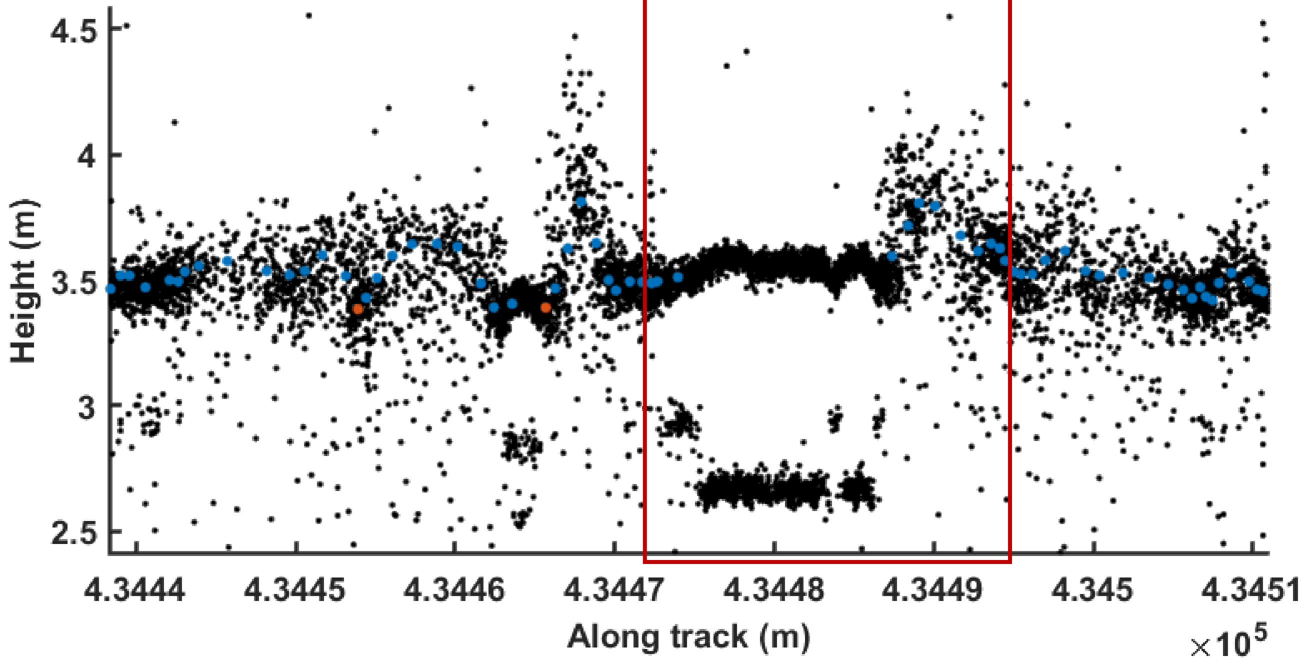


Variable ponds in ATL07

(1) No valid data

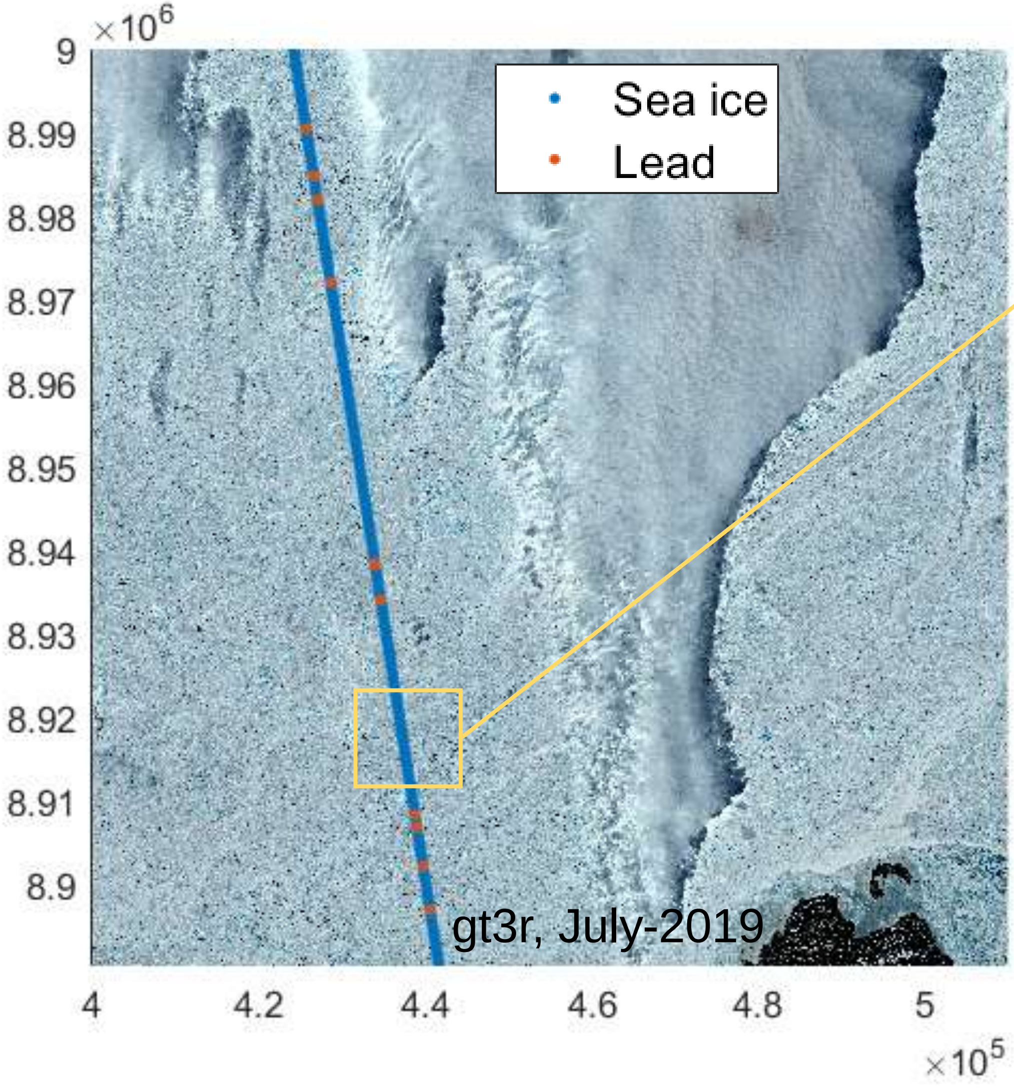


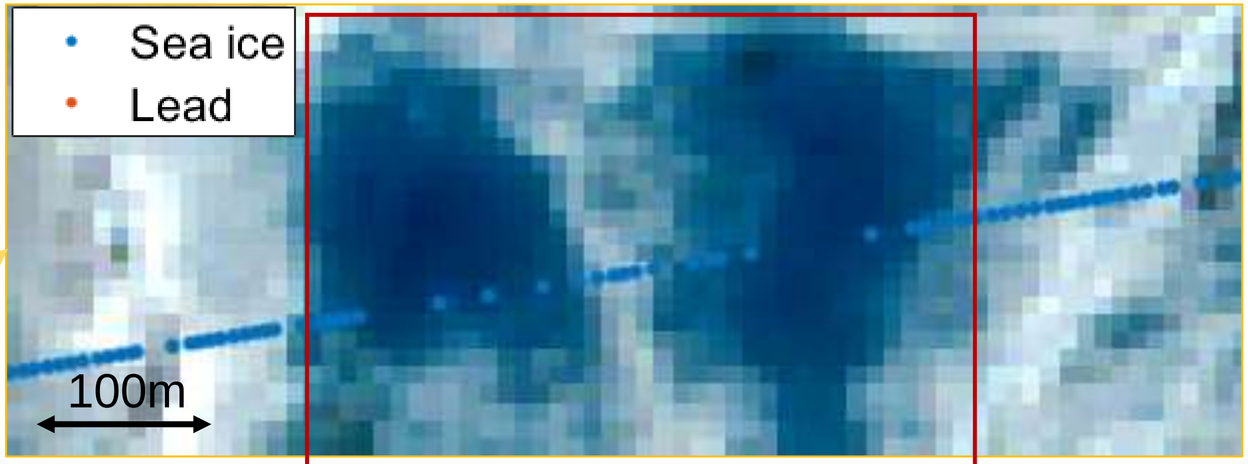


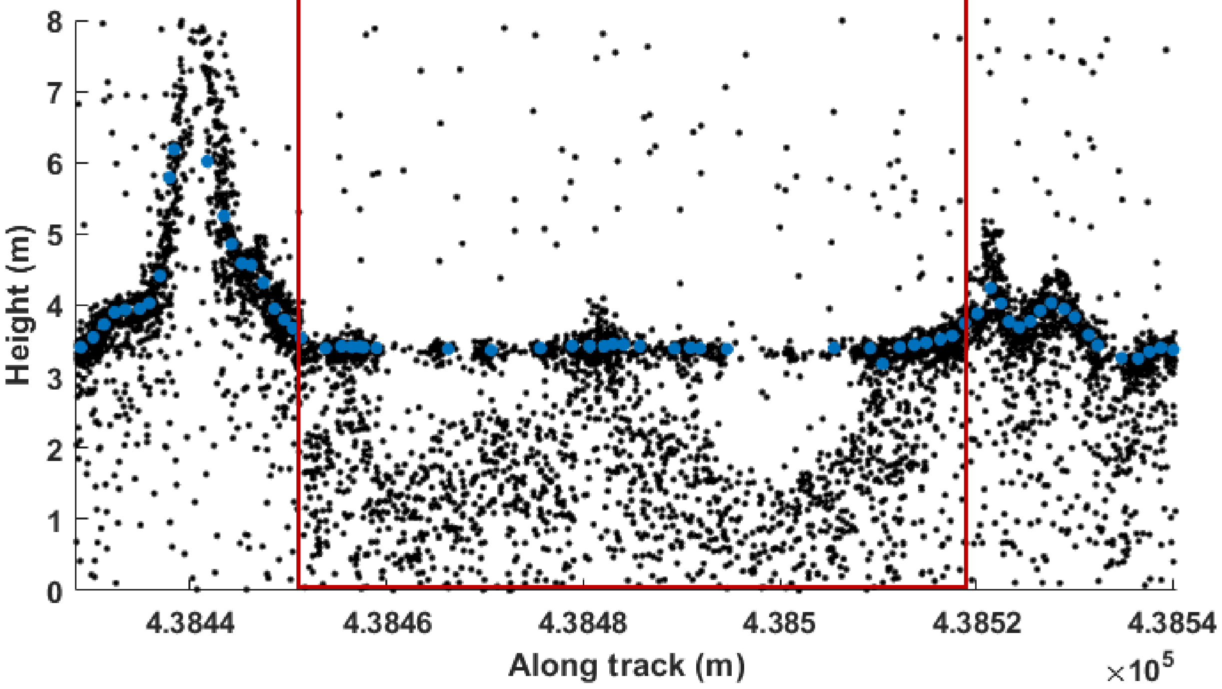


Variable ponds in ATL07

(2) ATL07 data is from the pond surface

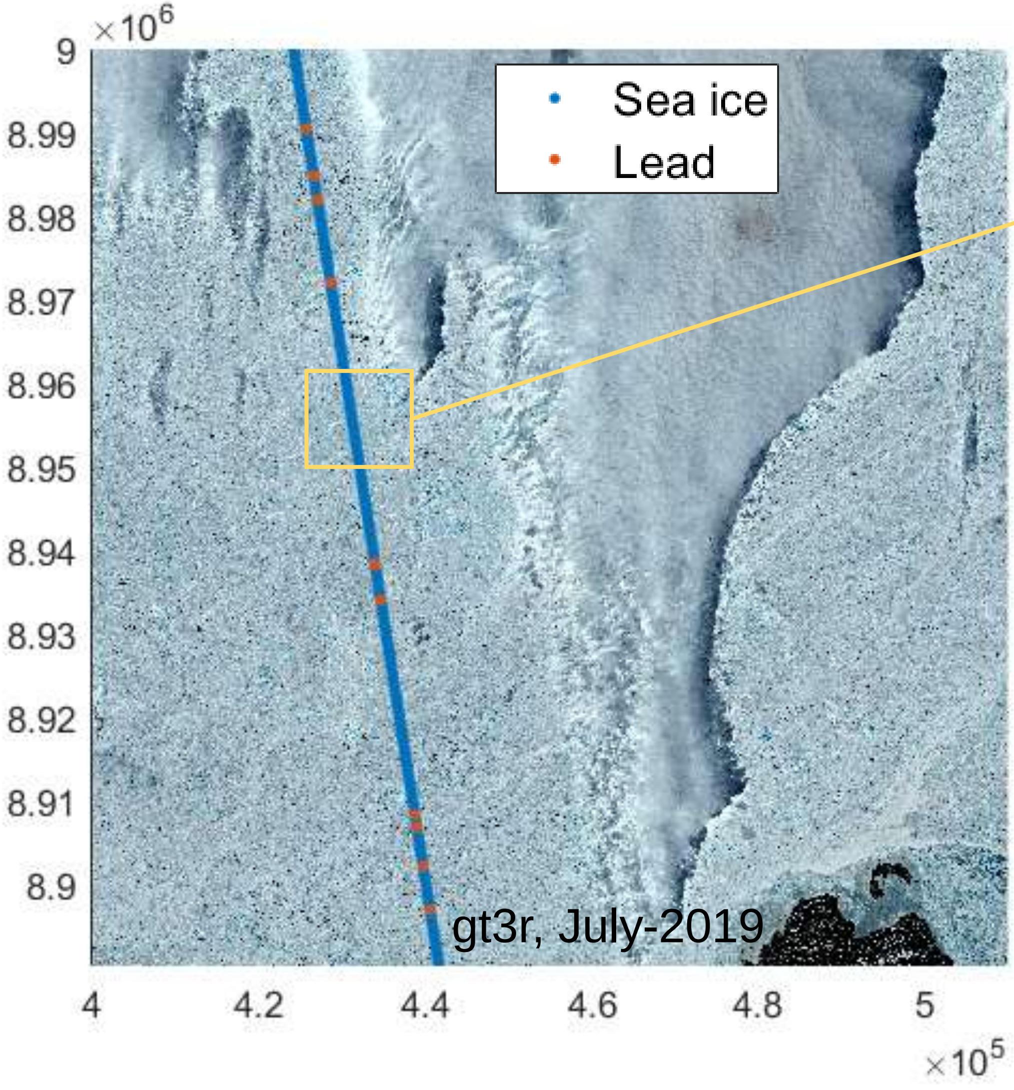


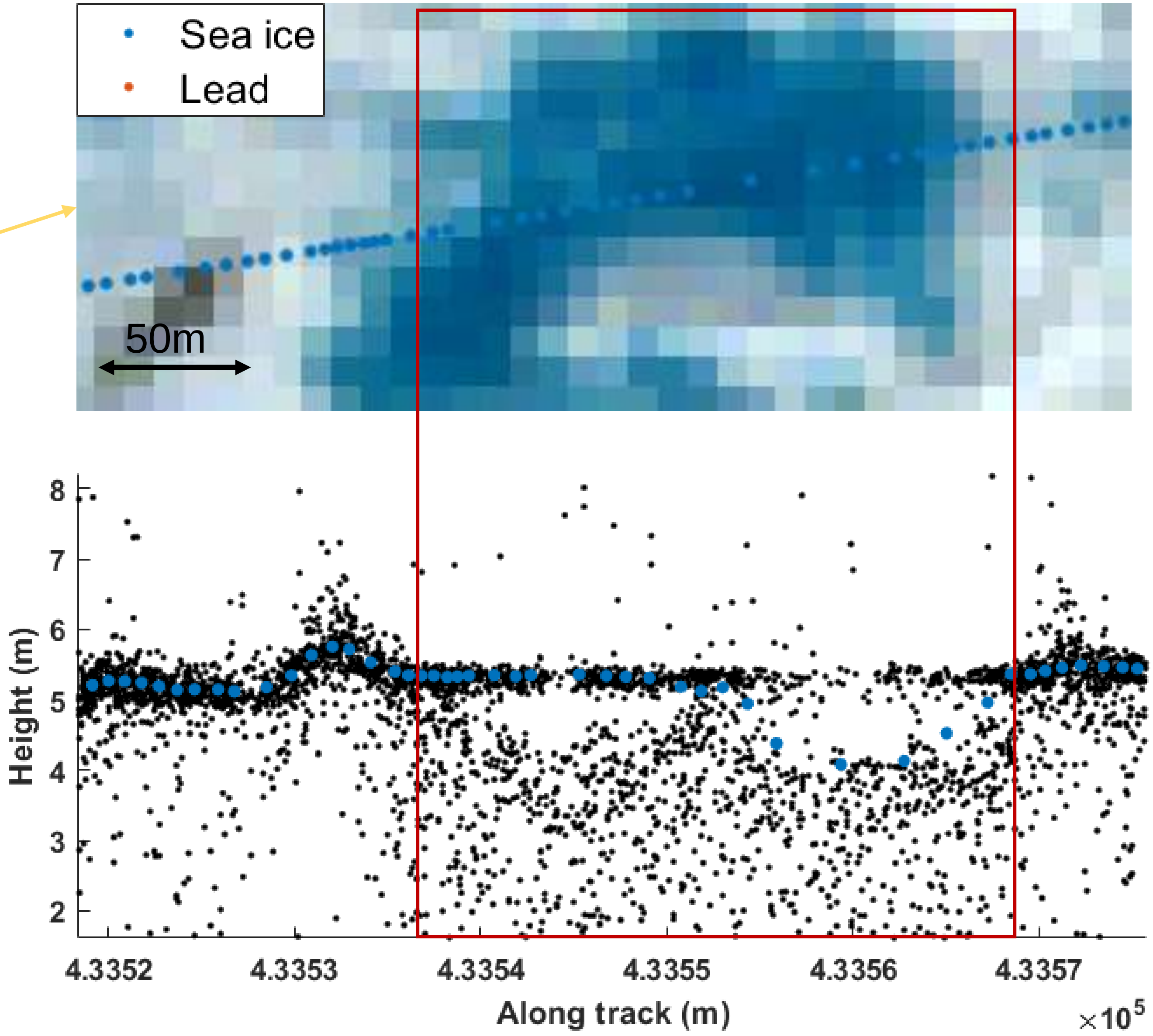




Variable ponds in ATL07

(3) ATL07 data is from the pond bottom

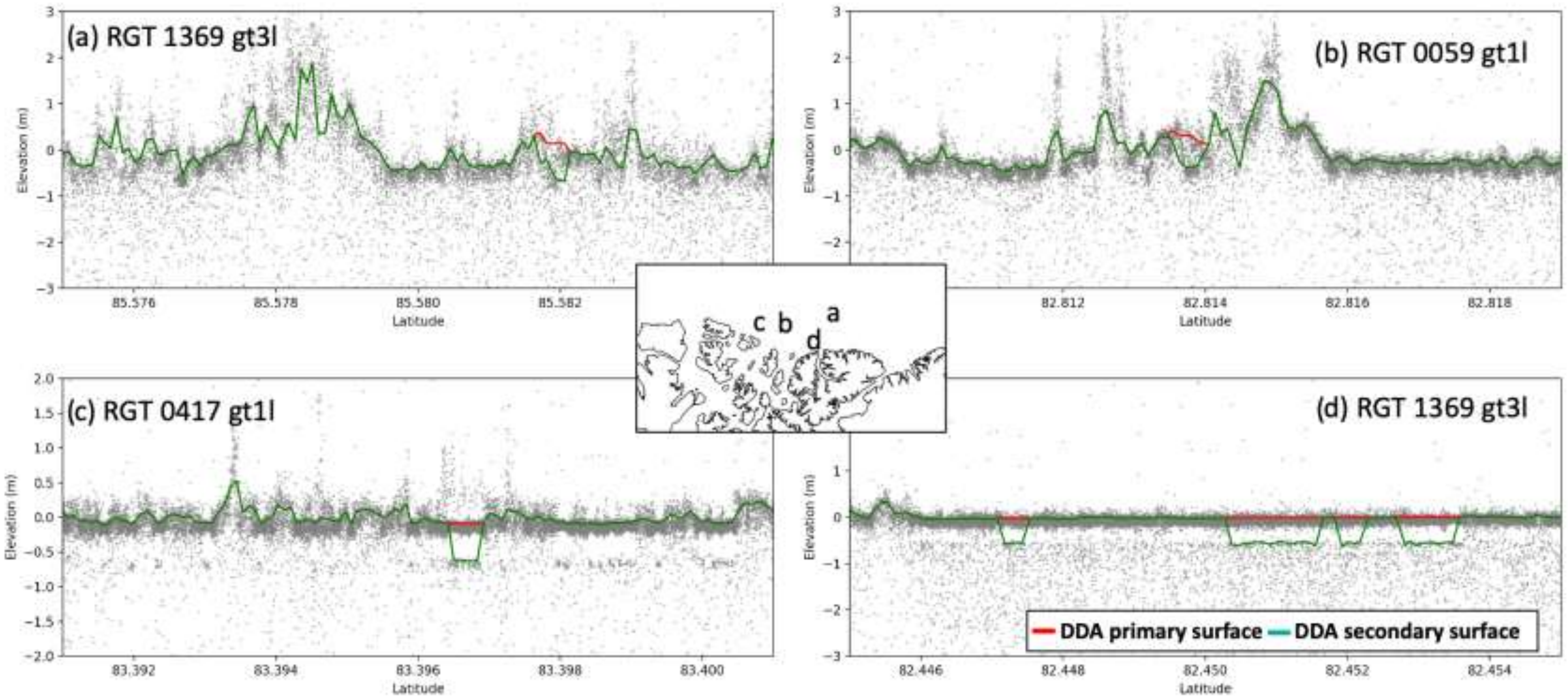




Potential for Obtaining Melt Pond from ATL03

Pioneering work by Herzfeld et al., (2023) and Buckley et al., (2023)

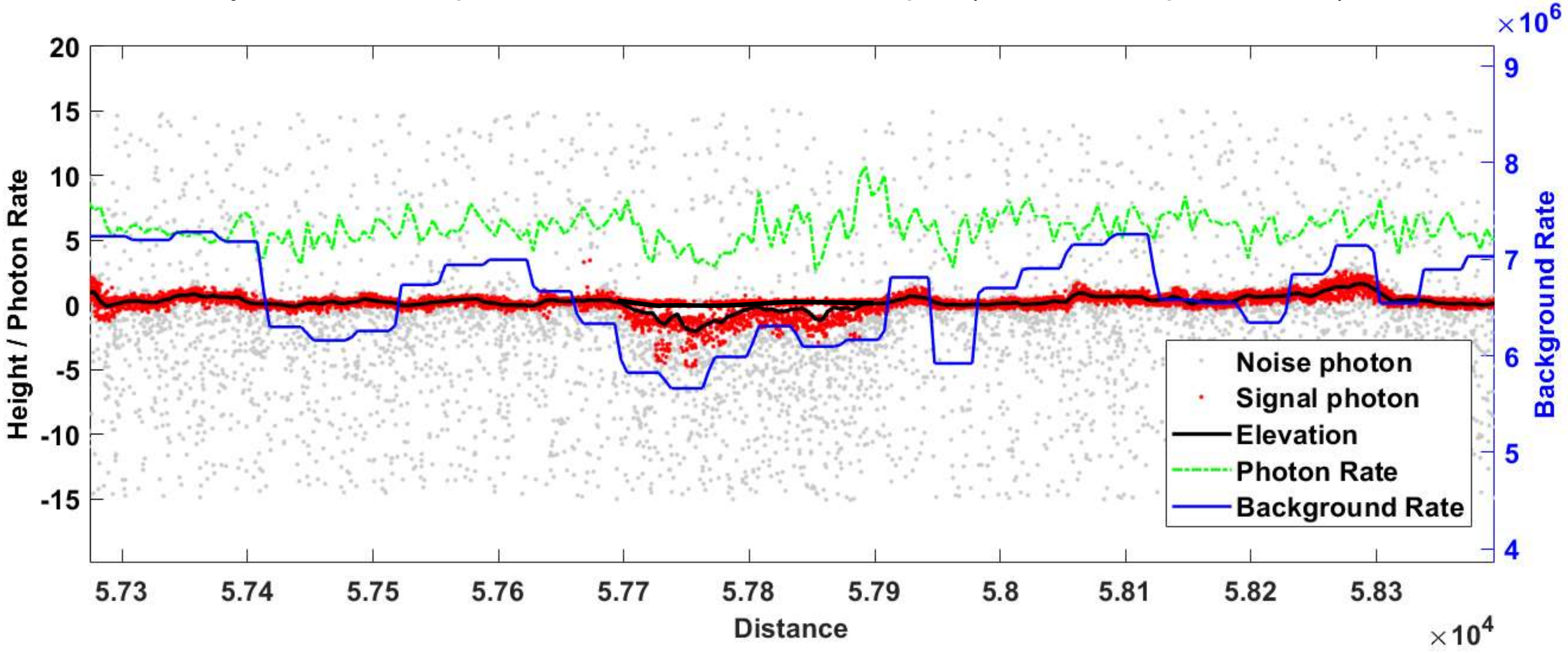
- Use DDA-Bifurcate-Sea ice algorithm to detect melt pond automatically
- Dataset has been released
- Without more classification (sea ice, lead……)



Potential for Obtaining Melt Pond from ATL03

We improve and apply the AC-KDE algorithm (Liu et al., 2023) to ATL03:

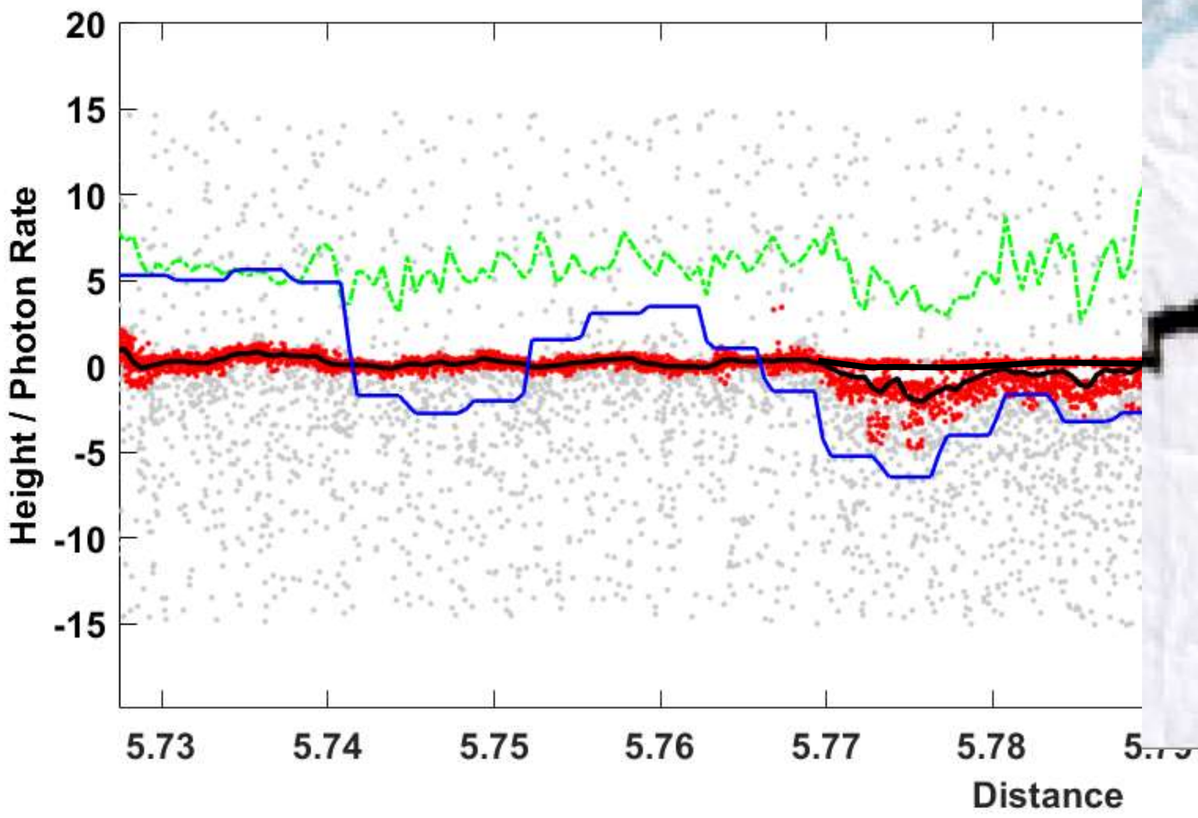
- Provide the sea ice profile in 3m resolution (identify ice ridges……)
- Derive the parameters used to classify (detect lead……)
- Automatically Locate melt ponds and measure their depth (detect melt ponds……)

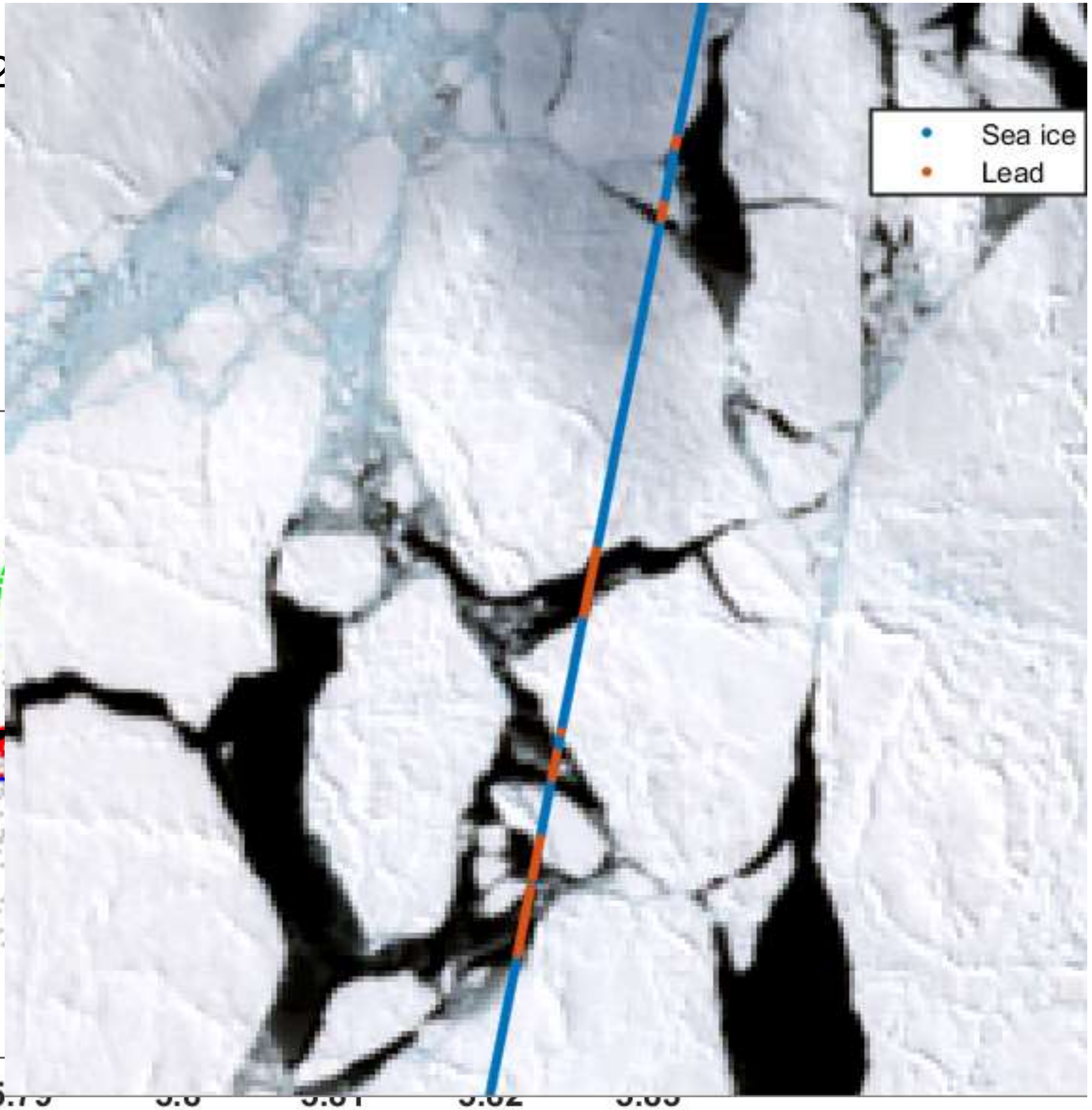


Potential for Obtaining Melt Pond from ATL03

We improve and apply the AC-KDE algorithm (Liu et al., 2

- Provide the sea ice profile in 3m resolution (identify
- Derive the parameters used to classify (detect lead.
- Automatically Locate melt ponds and measure their





Key takeaways

- Guided by the coincident Sentinel-2 imagery, we propose a combined unsupervised - supervised framework for enhancing ATL07 surface type classification.
- New Thin/gray ice (bare) type is included into the ATL07 surface type classification (need more quantitative validation efforts)
- We Improve the lead detection accuracy, especially avoid melt ponds being misclassified as lead
- The coincident dataset provide a valuable opportunity to assess sea ice product of ICESat-2

Future work

- Apply the enhanced classification results to year-round ATL07 data to estimate summer freeboard
- A new sea ice retrieval method for ATL03, with ability to derive more accurate sea ice information automatically, is under development