DEM Generation by Combining GEDI and ICESat-2 Data

Xiangxi Tian and Jie Shan
Purdue University
(Contact: Jie Shan at jshan@purdue.edu)

Outlines

1. Introduction on ICESat-2 and GEDI
2. Motivation
3. Study areas and dataset
4. Random forest spatial interpolation
5. Results and evaluation
6. Summary

ATLAS on ICESat-2 Satellite

- ATLAS: Advanced Topographic Laser Altimeter System
- The first spaceborne photon-counting laser altimetry launched in 2018
- Collected data timeframe: 2018-12 ~ 2022-06

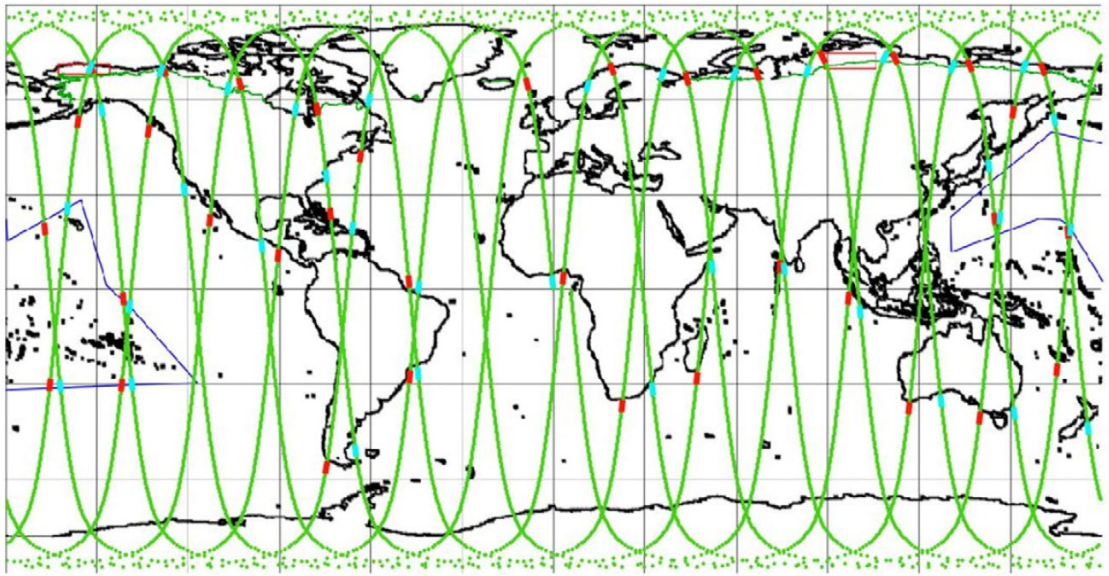
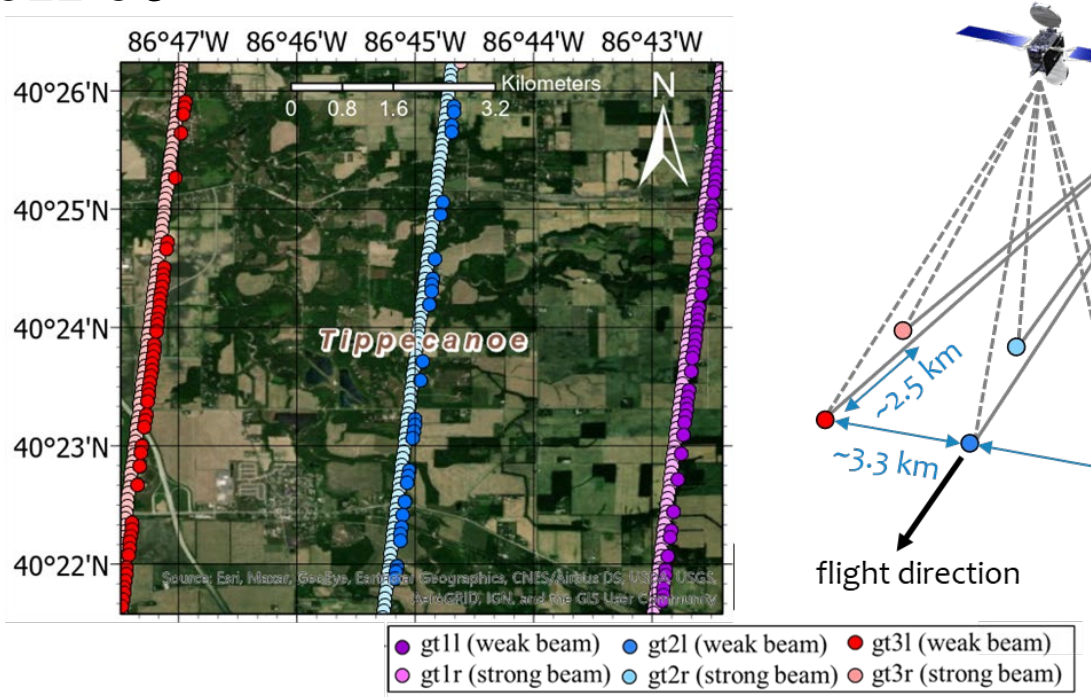


Illustration of one day of ICESat-2 orbits [1].

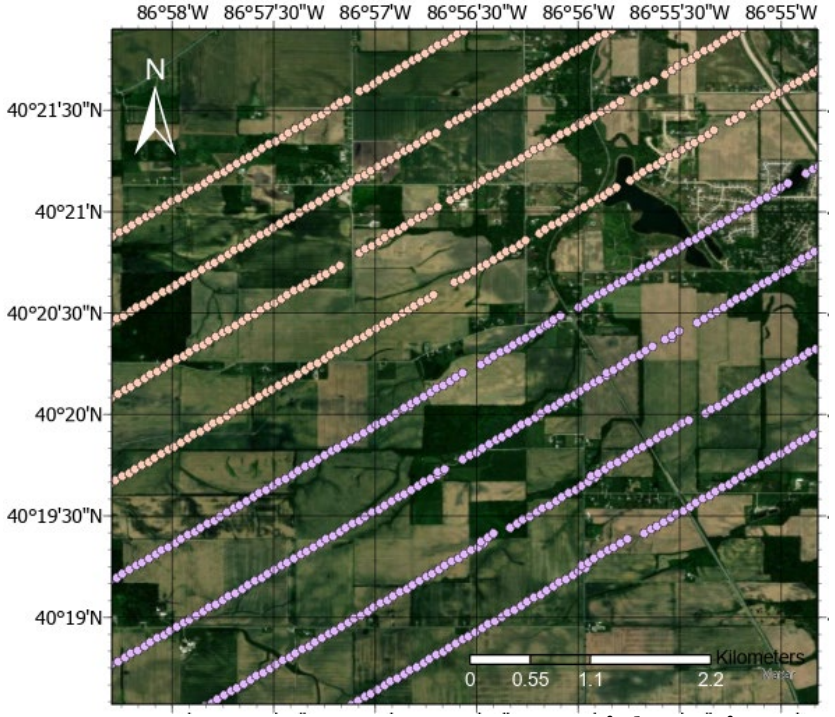


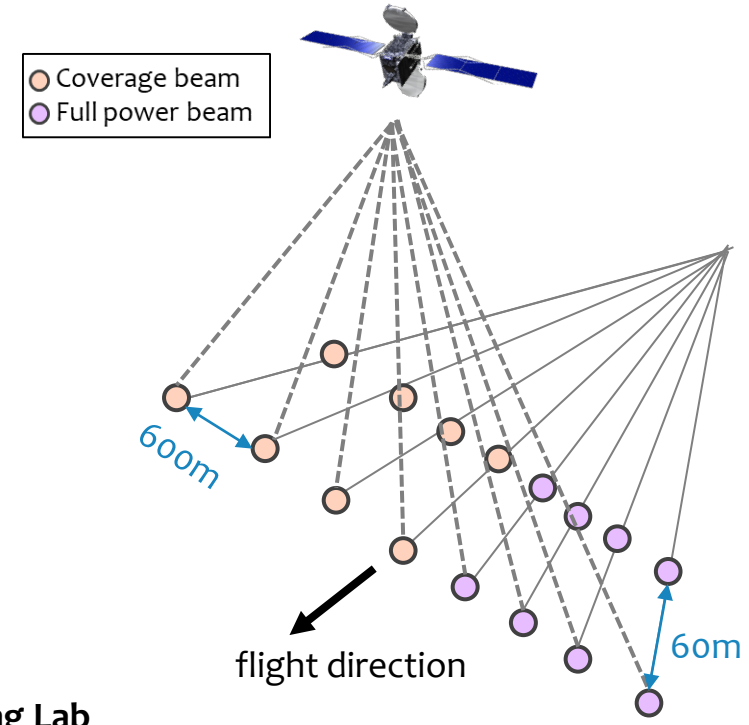
ATLAS ground track No.896 in Tippecanoe County and its beam footprints on ground [2].

[1] Markus et al., 2020. "Accuracy Assessment of GEDI Terrain Elevation and Canopy Height Estimates in European Temperate Forests: Influence of Environmental and Acquisition Parameters." Remote Sensing
[2] Tian, Xiangxi, and Jie Shan. 2021. "Comprehensive Evaluation of the ICESat-2 ATL08 Terrain Product." IEEE Transactions on Geoscience and Remote Sensing

GEDI on ISS

- GEDI: global ecosystem dynamics investigation
- Full waveform laser altimetry equipped on ISS since Dec 2018 [1]
- Near-circular orbit, altitude of ~ 400 km, inclination = 51.6° [1]
- Collected data timeframe: 2019-03 ~ 2022-06





[1] Dubayah et al. 2020. "The Global Ecosystem Dynamics Investigation: High-Resolution Laser Ranging of the Earth's Forests and Topography." Science of Remote Sensing.

Motivation

1. Current spaceborne laser altimetry is an efficient way to generate an up-to-date large-scale DEM.
2. Limitations on existing global DEMs.
3. Dense coverage can be gained by combining two datasets.

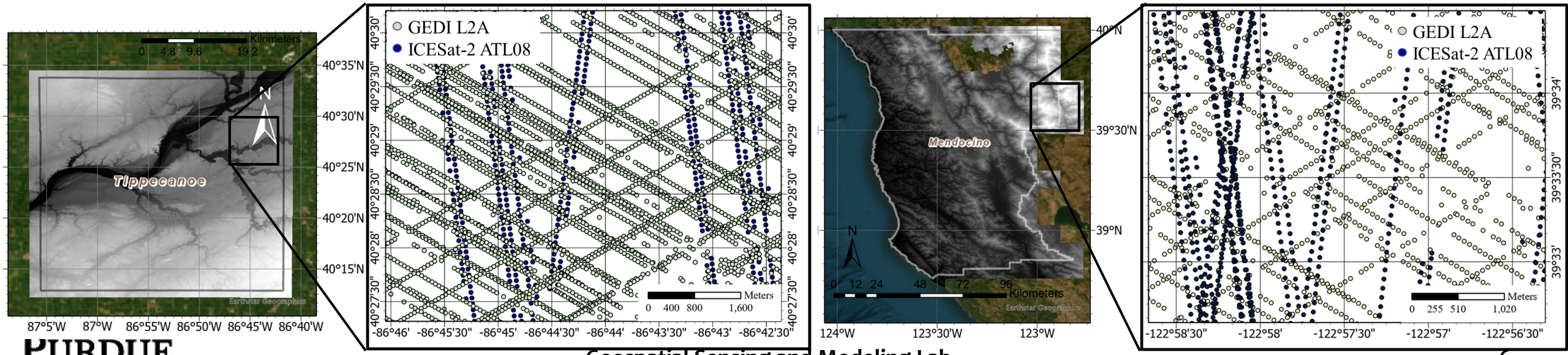
Study Areas and Data

Tippecanoe County, IN

- area ≈ 1,303.4 km^2
- average slope < 2°
- maximum elevation variance ≈ 110 m
- GEDI L2A footprints: 179,2839
- ICESat-2 ATL08 segments: 28,363
  - Accuracy w.r.t 3DEP DEM: $\mu = -0.56m, \sigma = 1.00m$

Mendocino County, CA

- area ≈ 10,040 km^2
- average slope > 6°
- maximum elevation variance ≈ 2100 m
- GEDI L2A footprints: 145,6036
- ICESat-2 ATL08 segments: 162,214
  - Accuracy w.r.t 3DEP DEM: $\mu = -1.73m, \sigma = 5.69m$



Random Forest Spatial Interpolation (RFSI)

- GEDI footprints filtering criteria:

$$|h_{GEDI} - h_{SRTM}| \leq 50\ m; \qquad 1 > sensitivity > 0.9$$

- Terrain height prediction at s_0 [1]:

$$\hat{h}(s_0) = f\{X(s_0), Y(s_0), \bar{h}_{GEDI}(s_0), \bar{h}_{ATL08}(s_0)\}$$

X, Y: projected coordinates at s_0;

$\bar{h}_{GEDI}$: NN interpolated height from nearby GEDI footprints;

$\bar{h}_{ATL08}$: NN interpolated height from nearby ATL08 segments.

- Features:

$$X_0, Y_0, \bar{h}_{GEDI}, \bar{h}_{ATL08}$$

- Target variable:

terrain height at s_0 from nearest ATL08 segment within the grid ($\hat{h}_{ATL08}$)

- RFSI hyperparameters:
  - Tree depth: no limitation (until all leaves contain less than 2 samples);
  - Number of trees: 1000

[1] Sekulić A, Kilibarda M, Heuvelink GBM, Nikolić M, Bajat B. Random Forest Spatial Interpolation. Remote Sensing. 2020

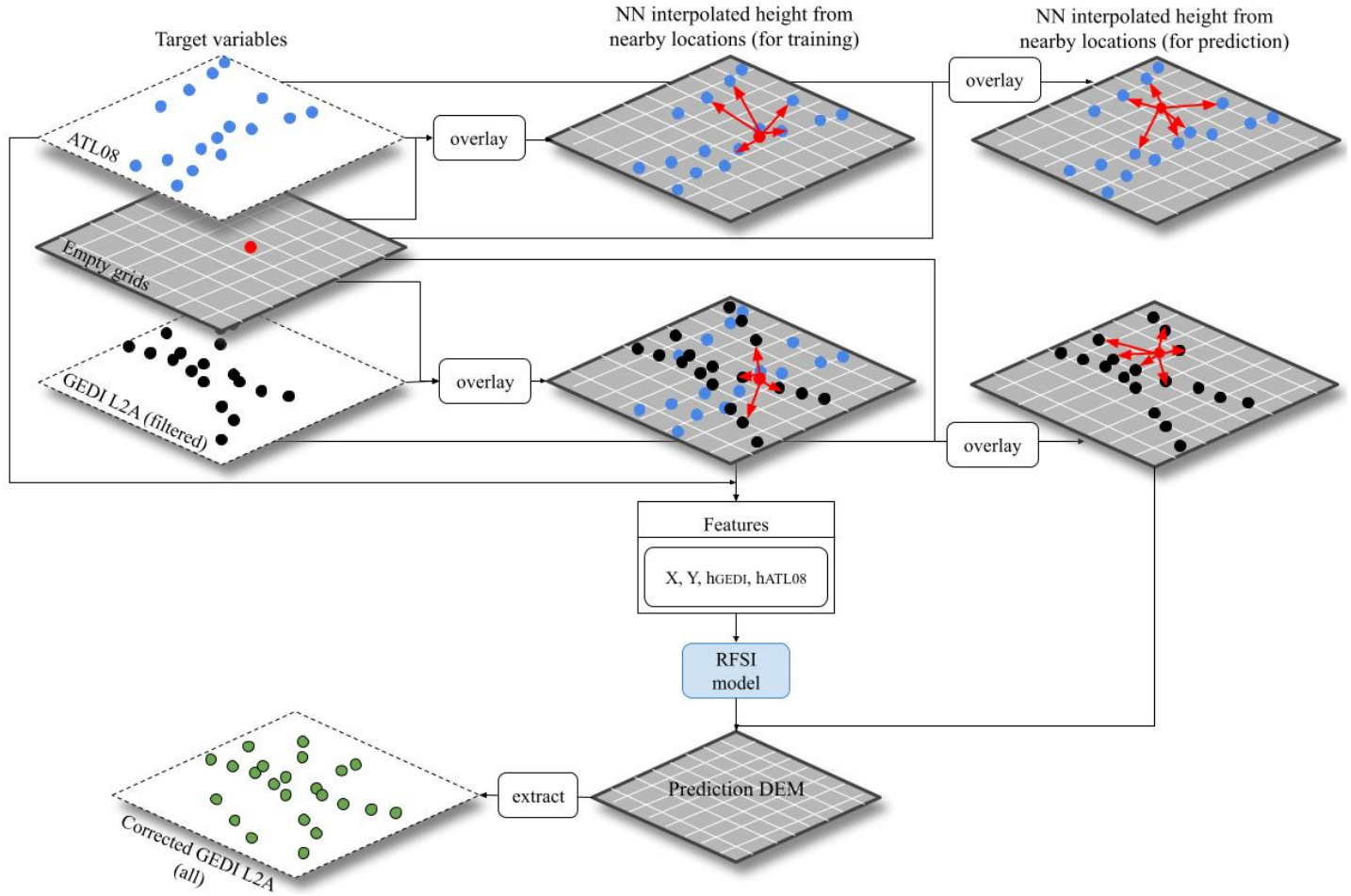
Schematic Representation of the Method

Inputs:

- ICESat-2 ATL08 centers of segments
- Filtered GEDI L2A points (~67% and 75% of total)

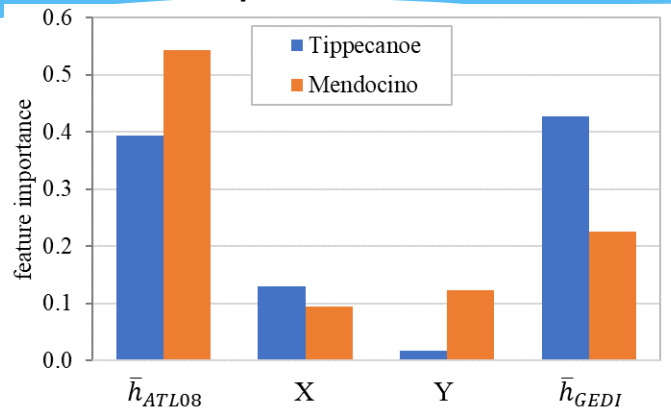
Outputs:

- 30m DEM & corrected GEDI height
- 90 m DEM & corrected GEDI height



Quality of the RFSI Trained Model (1)

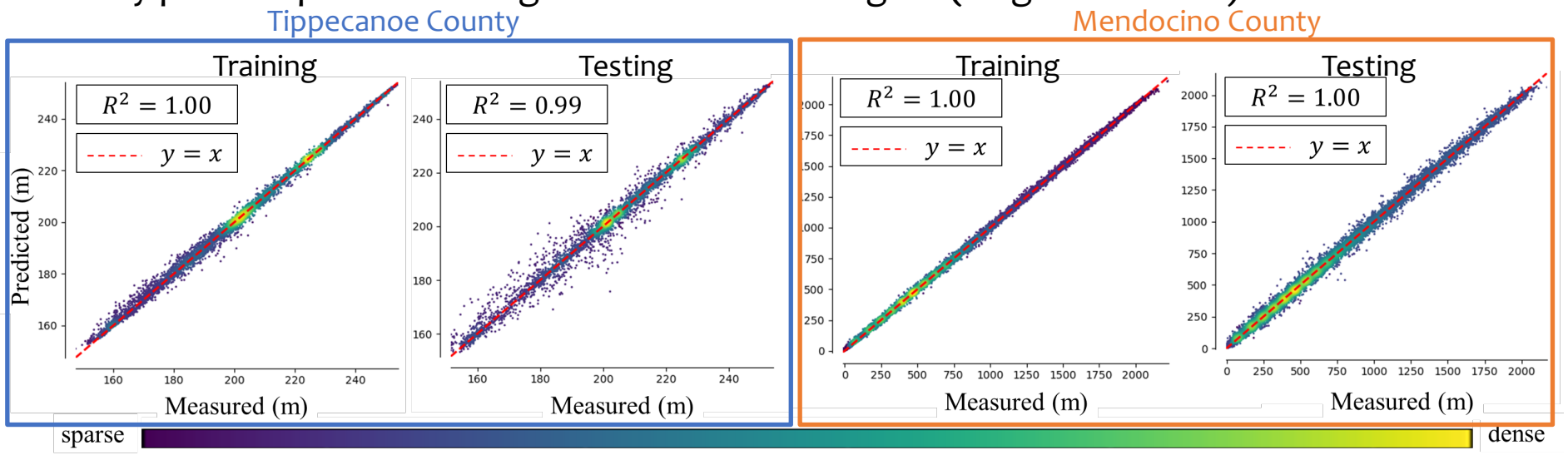
- Feature importance



- Number of data samples for model training

	Tippecanoe	Mendocino
# of data (training dataset)	21,272	121,660
# of data (testing dataset)	7,091	40,554
Total	28,363 (~22 pts/km^2)	162,214 (~16 pts/km^2)

- Scatter density plots of predicted height vs. measured heights (target variables)



Quality of the RFSI Trained Model (2)

- Training and testing error

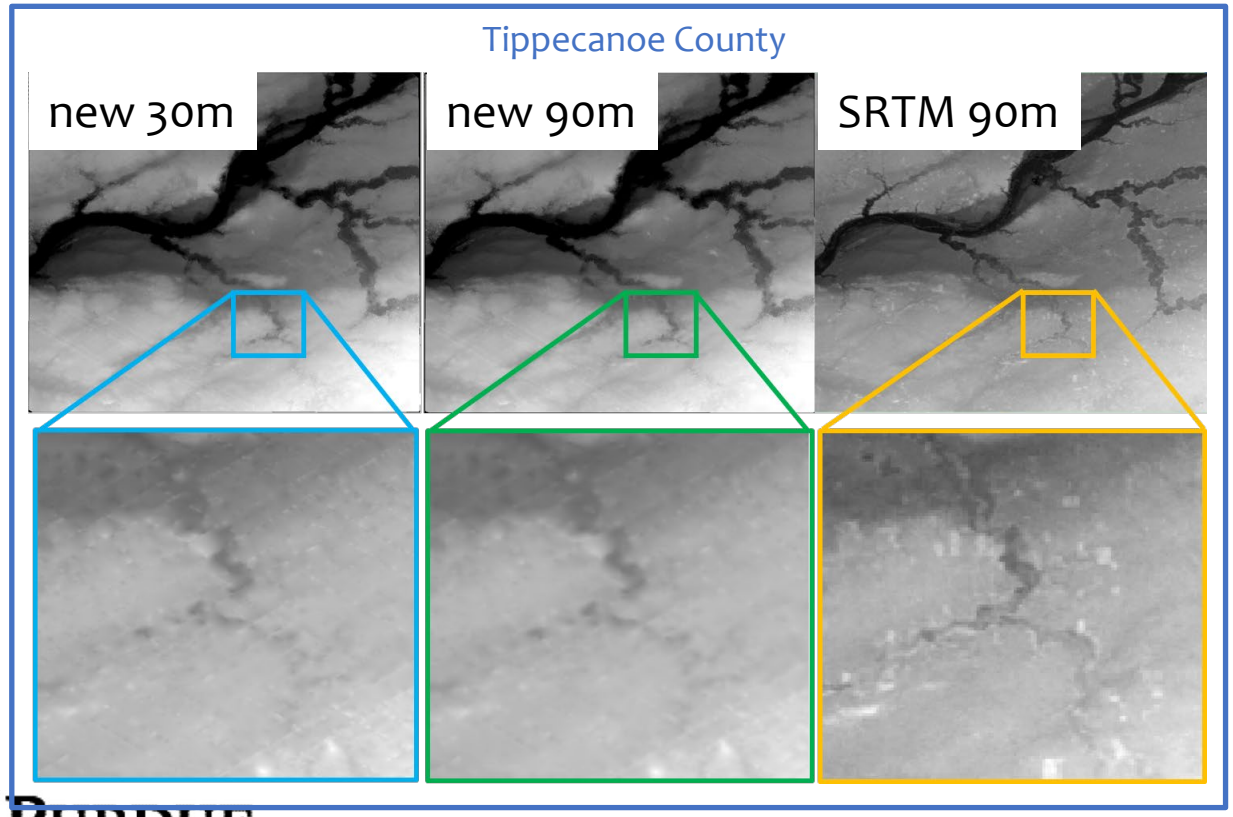
$$h_{prediction} - target\ variable$$

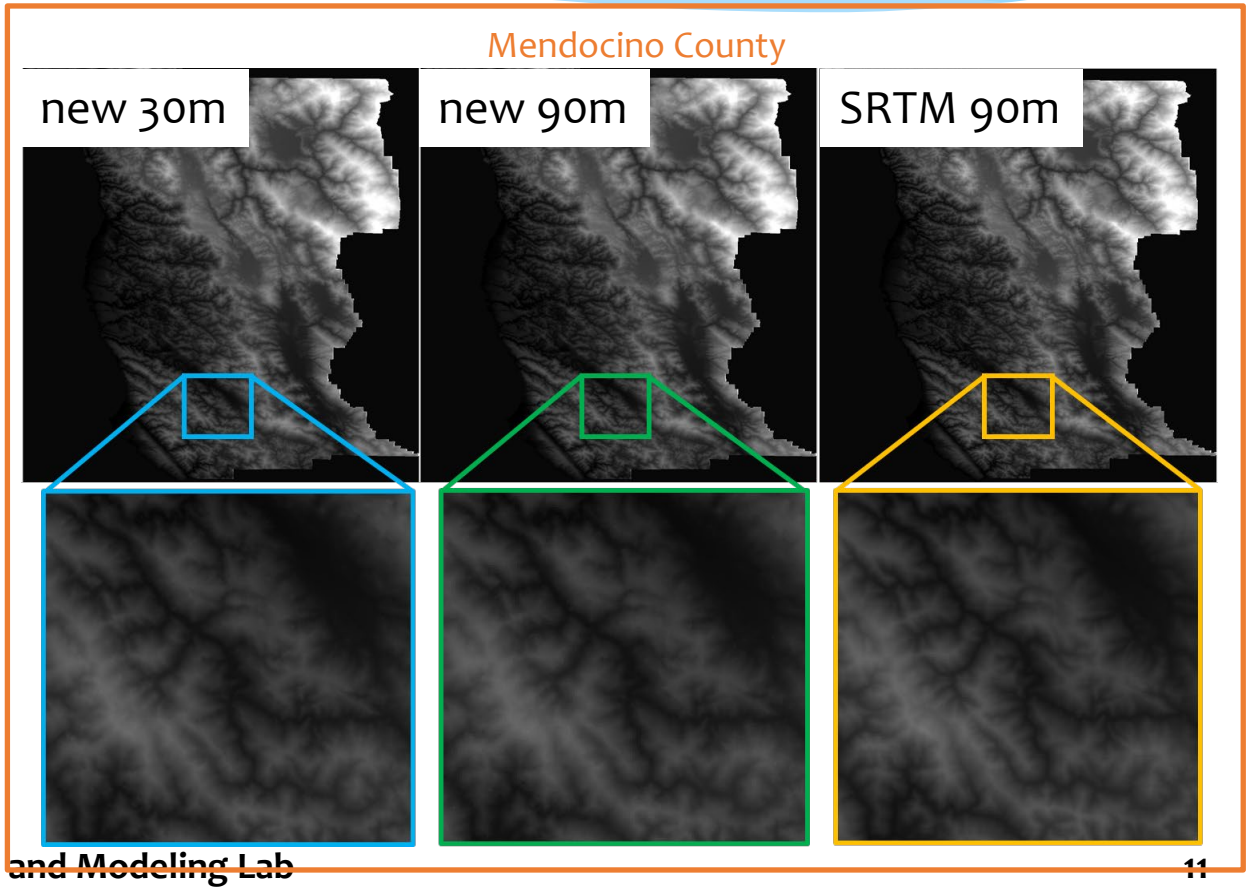
- Statistics

	Tippecanoe County			Mendocino County		
	median (m)	mean (m)	Std. dev (m)	median (m)	mean (m)	Std. dev (m)
training dataset	-0.0164	0.0036	0.8356	-0.0816	-0.0038	7.1862
testing dataset	-0.0498	-0.0392	2.1701	-0.2212	-0.0863	19.3260
total dataset	-0.0205	-0.0071	1.3042	-0.0969	-0.0244	11.4938

Quality of the Predicted DEM (1)

The created DEM is smoother than SRTM DEM (less noise)





Quality of the Predicted DEM (2)

- All DEMs are compared with 3DEP DEM
- The created DEM has smaller uncertainty than SRTM DEM

	Tippecanoe County			Mendocino County		
	median (m)	mean (m)	Std. dev. (m)	median (m)	mean (m)	Std. dev. (m)
SRTM 30m	1.1219	1.9927	3.2128	6.9626	8.2967	11.8224
New 30m DEM	-0.4654	-0.4499	2.8433	-2.7609	-2.5631	20.2978
SRTM 90m	1.1272	2.0109	3.5627	6.3917	7.8918	24.0673
New 90m DEM	-0.4730	-0.4574	2.8298	-2.9121	-3.1405	23.3834

Quality of the Corrected GEDI Height

- Large improvement on the terrain height for all GEDI footprints

	Tippecanoe County			Mendocino County		
	mean (m)	median (m)	Std. dev (m)	mean (m)	median (m)	Std. dev (m)
original height	424.2400	0.3213	853.2060	87.9613	-54.0217	804.5993
corrected height (extracted from new 30m DEM)	-0.4859	-0.5821	2.7545	-2.5061	-2.7557	17.2784
corrected height (extracted from new 90m DEM)	-0.4791	-0.5893	2.7300	-3.0689	-2.8320	22.8308

Summary and Future Work

1. Spatial features utilized by RFSI algorithm
2. Achieved a smaller biased DEMs (30m and 90m) by combining GEDI L2A and ICESat-2 ATL08 data
3. Extracted terrain height for all GEDI points shows a large improvement than the original height

4. Further reduction on the uncertainty of the new DEM
5. Explore other methods assisting the DEM interpolation

Thanks!