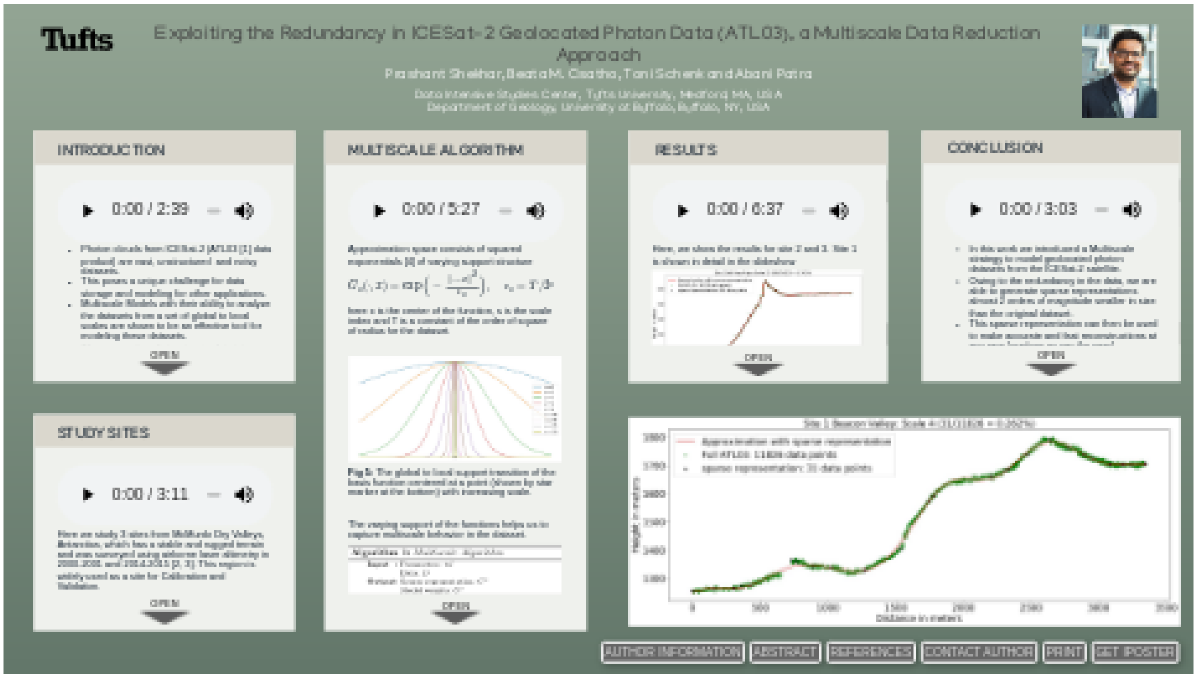
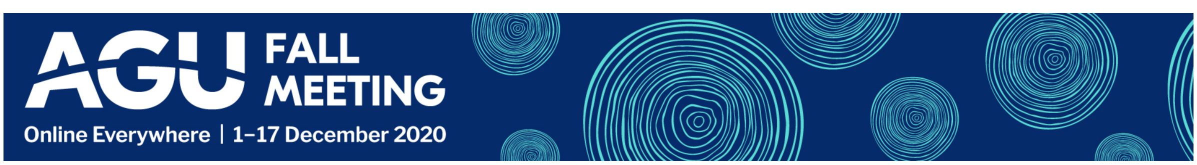
Exploiting the Redundancy in ICESat-2 Geolocated Photon Data (ATL03), a Multiscale Data Reduction Approach



Prashant Shekhar, Beata M. Csatho, Toni Schenk and Abani Patra

Data Intensive Studies Center, Tufts University, Medford, MA, USA
Department of Geology, University at Buffalo, Buffalo, NY, USA

PRESENTED AT:



INTRODUCTION

0:00 / 2:39

- Photon clouds from ICESat-2 (ATL03 [1] data product) are vast, unstructured and noisy datasets.
- This poses a unique challenge for data storage and modeling for other applications.
- Multiscale Models with their ability to analyze the datasets from a set of global to local scales are shown to be an effective tool for modeling these datasets.
- The segregation of structure in data into a set of scales, also provides an efficient tool for dealing with inherent noise.
- These models are also demonstrated as a means for effective data reduction in the sense

$$Big\ Data \implies Sparse\ Data + Sparse\ Model$$

STUDY SITES

0:30 / 3:11

Here we study 3 sites from McMurdo Dry Valleys, Antarctica, which has a stable and rugged terrain and was surveyed using airborne laser altimetry in 2000-2001 and 2014-2015 [2, 3]. This region is widely used as a site for Calibration and Validation.

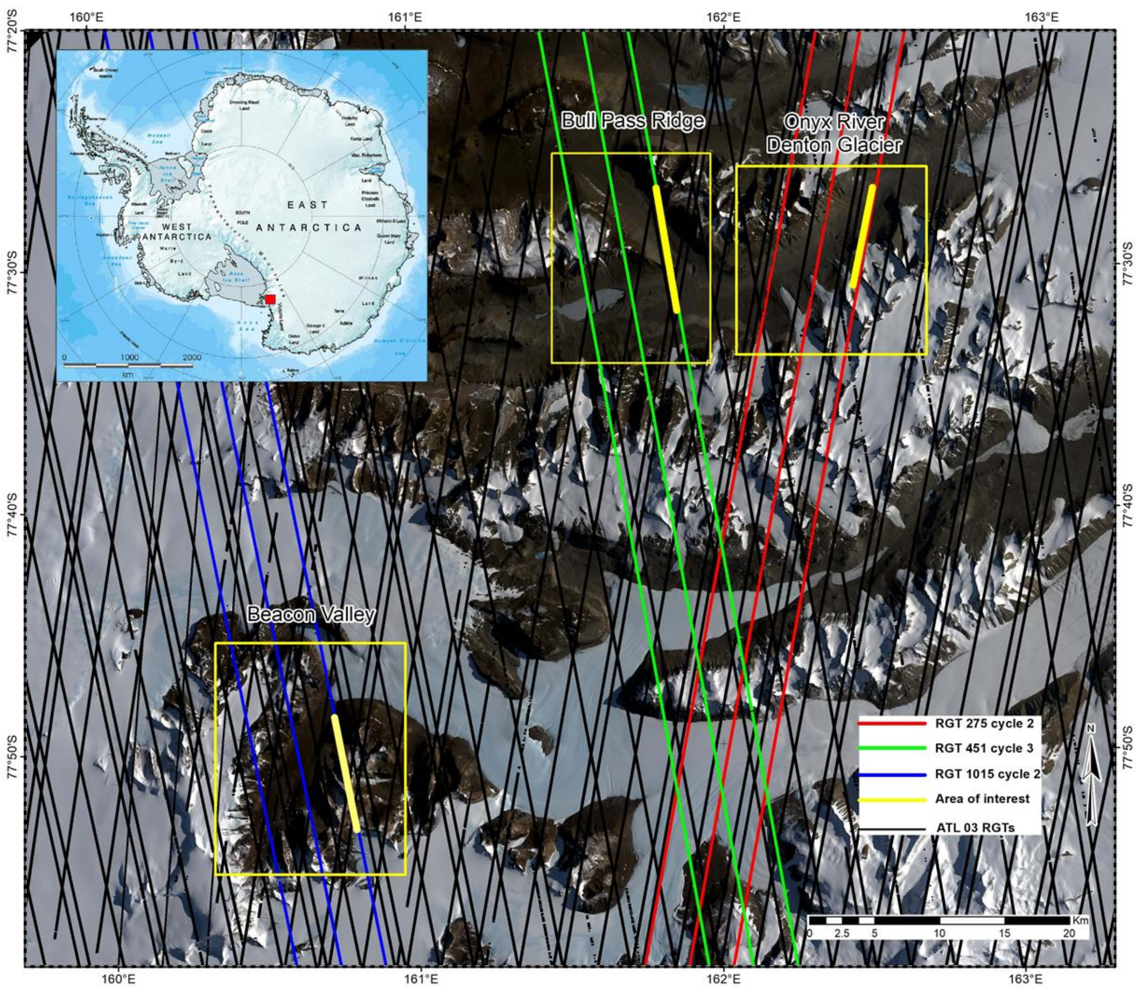


Fig 1: Selected transects (thick yellow line) in the Dry Valleys, East Antarctica, shown on Landsat-8 2020/02/03, composite bands 4,3,2 satellite image. Black lines are ICESat-2 repeat ground tracks and colored lines show selected ground-tracks.

SITE 1

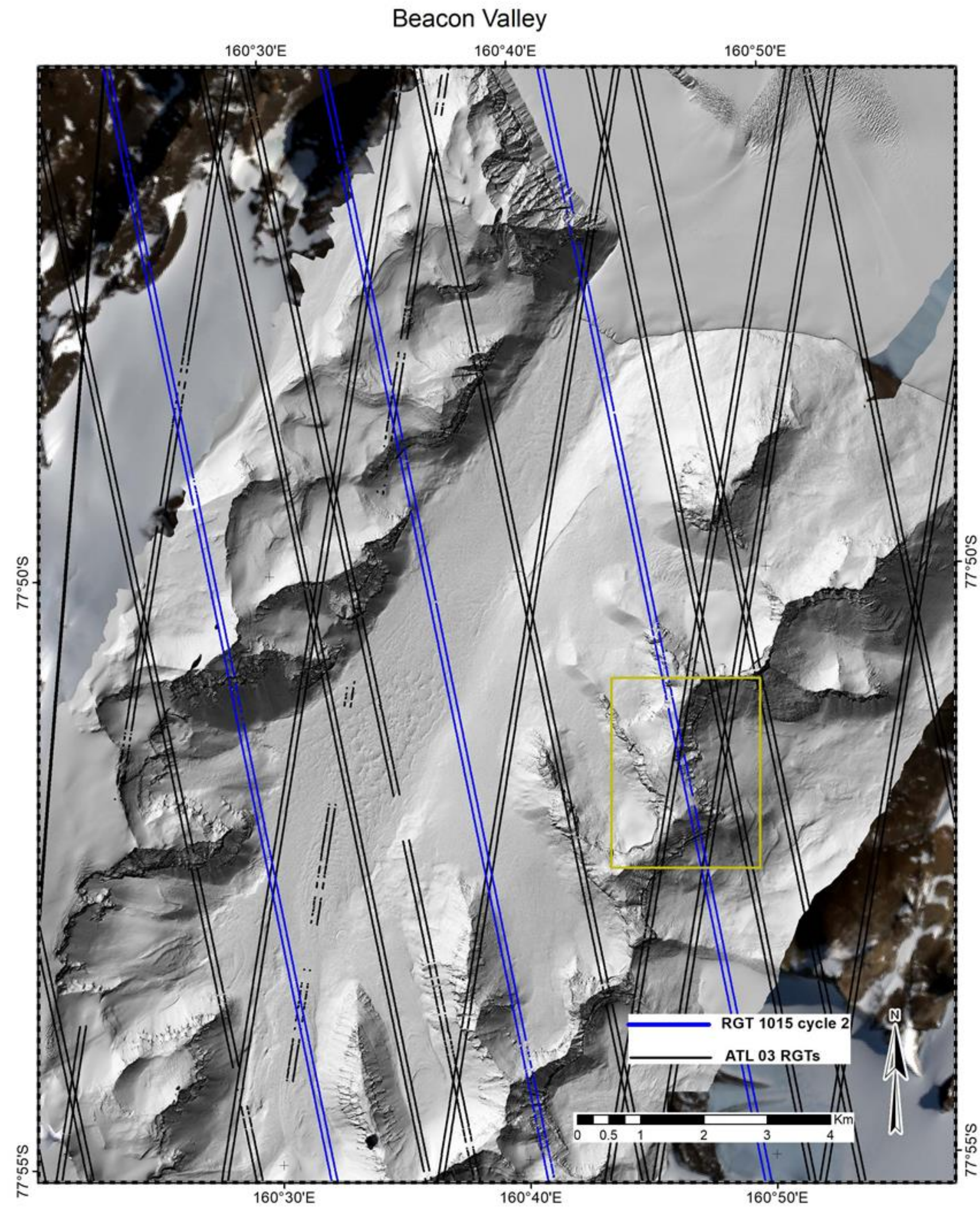


Fig 2: Beacon Heights: ICESat-2 ground track (RGT 1015, cycle 2, beam 5, March 4, 2019, descending) crossing the main ridge of Beacon Heights south of Beacon Valley (77º 51’ 30’’ S, 160º 46’ 17’’ E). Beacon Heights is characterized by steep cliffs and rugged topography and therefore poses a difficult challenge for data reduction.

SITE 2

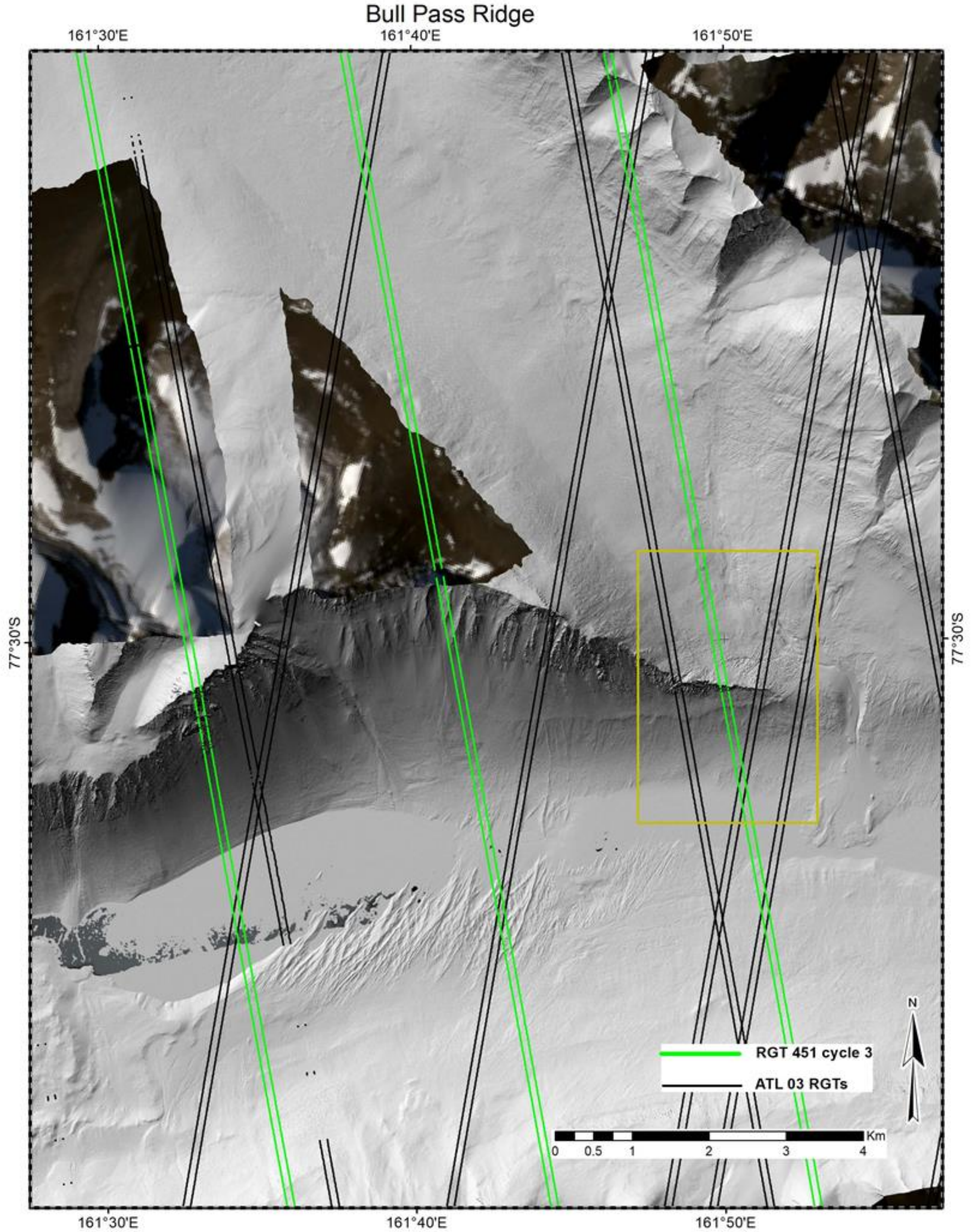


Fig 3: Bull Pass Ridge: ICESat-2 ground track (RGT 451, cycle 3, beam 5, April 27, 2019, descending) crossing the Bull Pass Ridge, separating Wright Valley and Bull Pass (77º 30' 22'' S, 151º 49' 58" E). This site was selected to test the proposed approach to capture sharp ridges in the topography

SITE 3

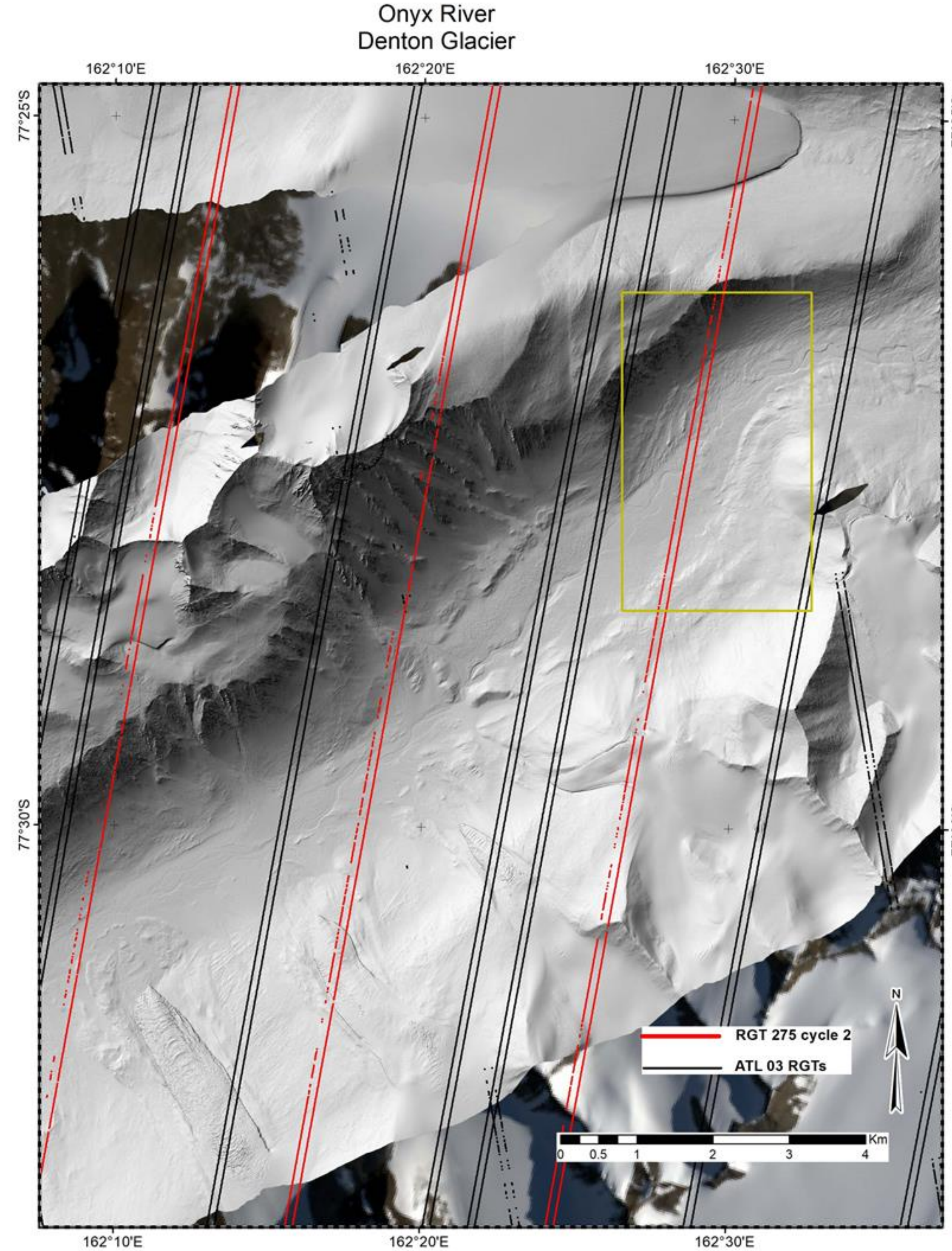


Fig 4: Onyx River: ICESat-2 ground track (RGT 275, cycle 2, beam 1, January 15, 2019, ascending) crossing the eastern part of Wright Valley with Onyx River in the center (77º 28’ 6’’ S, 162º 25’ 34” E). This site was selected to test the proposed approach on relatively smooth topography.

MULTISCALE ALGORITHM

0:00 / 5:27

Approximation space consists of squared exponentials [4] of varying support structure

$$G_s(\cdot, x) = \exp\left(-\frac{||\cdot - x||^2}{\epsilon_s}\right), \quad \epsilon_s = T/2^s$$

here x is the center of the function, s is the scale index and T is a constant of the order of square of radius for the dataset.

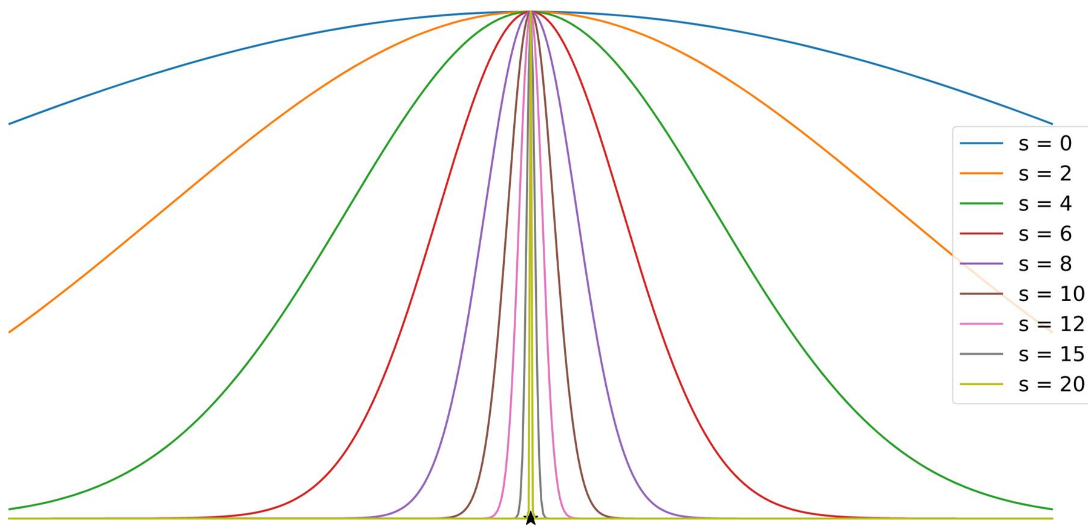


Fig 5: The global to local support transition of the basis function centered at a point (shown by star marker at the bottom) with increasing scale.

The varying support of the functions helps us to capture multiscale behavior in the dataset.

Algorithm 1: *Multiscale Algorithm*

Input : Parameters: tol
Data: D
Output: Sparse representation: C^s
Model weights: Θ^s

Initialize: $s = 0, f^s = 0, r = y$
while $||y - f^s||_2 > tol$ **do**
    $B_s, \Theta_s, C_s \leftarrow Basis\_selection(D, s, r)$
    $B^s, \Theta^s, C^s \leftarrow Update\_model(B^s, \Theta^s, C^s, B_s, \Theta_s, C_s)$
    $f_s \leftarrow B_s\Theta_s$
    $r \leftarrow r - f_s$
    $f^s \leftarrow f^s + f_s$
    $s \leftarrow s + 1$
Return $[C^s, \Theta^s]$

Given a dataset D = (X,y) and starting with scale 0, the iterations continue till the error of reconstruction is acceptable (within tol)

1. We generate scale specific basis functions B_s weights Θ_s and sparse representation C_s using a greedy basis selection [5] approach.
2. We then append the scale specific quantities to global basis B^s, global weight Θ^s and global sparse representation C^s.
3. Before moving onto the next scale, we update the current approximation f^s, the residual r.

The sparse representation C^s is composed of the coordinate at which the basis function is centered.

With the sparse representation, any future predictions can be made.

RESULTS

0:00 / 6:37

Here, we show the results for site 2 and 3. Site 1 is shown in detail in the slideshow

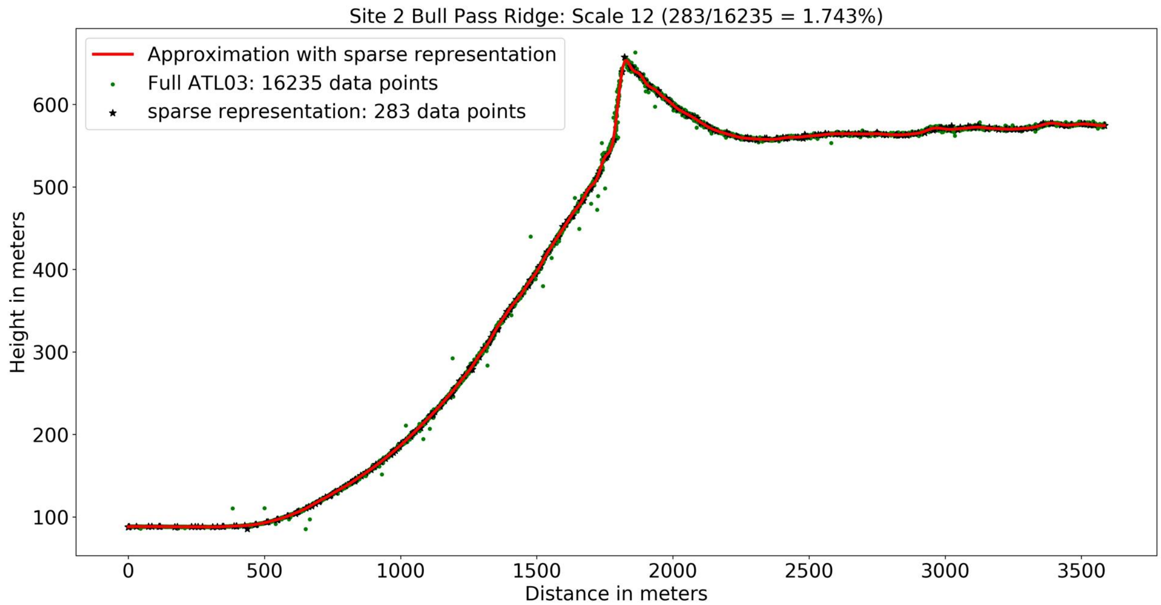


Fig 6: Site 2(Bull Pass Ridge) with Track 451 Cycle 3. With 1.7% of the full dataset, we obtain the reconstruction shown in red. Notice the capability to capture sharp changes (ridges) which was made possible due to rapidly decaying basis functions at higher scales.

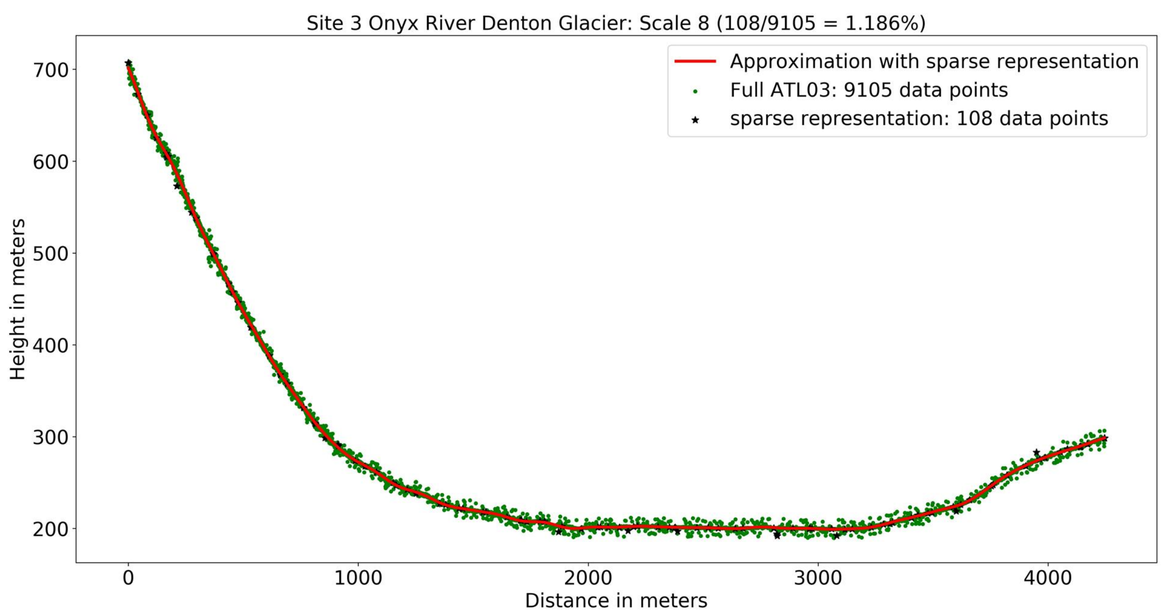


Fig 7: Site 3 (Onyx river) with Track 275 Cycle 2. Being a relatively smooth terrain, Multiscale algorithm is able to capature the structure with even smaller sparse representation.

Overall, with less than 2% of the datapoints for all 3 sites, the Multiscale Algorithm is able to capture the structure of the data and provide reconstructions to user specified tolerances.

CONCLUSION

0:00 / 3:03

- In this work we introduced a Muiltiscale strategy to model geolocated photon datasets from the ICESat-2 satellite.
- Owing to the redundancy in the data, we are able to generate sparse representations almost 2 orders of magnitude smaller in size than the original dataset.
- This sparse representation can then be used to make accurate and fast reconstructions at any new locations as per the need.
- With the availability of explicit basis functions, we can also estimate the rate of change with the distance along the track.
- Our approach provides a continuous, analytical and precise representation of the surface, suitable for various applications, such as change detection, geomorphologic mapping, geodetic control.
- Future work includes modifying the strategy to incorporate other objectives like quantile losses besides the default mean squared error.

References

1. Neumann, Thomas A., et al. "The Ice, Cloud, and Land Elevation Satellite–2 Mission: A global geolocated photon product derived from the advanced topographic laser altimeter system." Remote sensing of environment 233 (2019): 111325

2. Csatho, B., Schenk, T., Krabill, W., Wilson, T., Lyons, W., McKenzie, G., Hallam, C., Manizade, S., Paulsen, T.W., 2005. Airborne laser scanning for high−resolution mapping of Antarctica. Eos Trans. AGU, 86(25), 237-238.

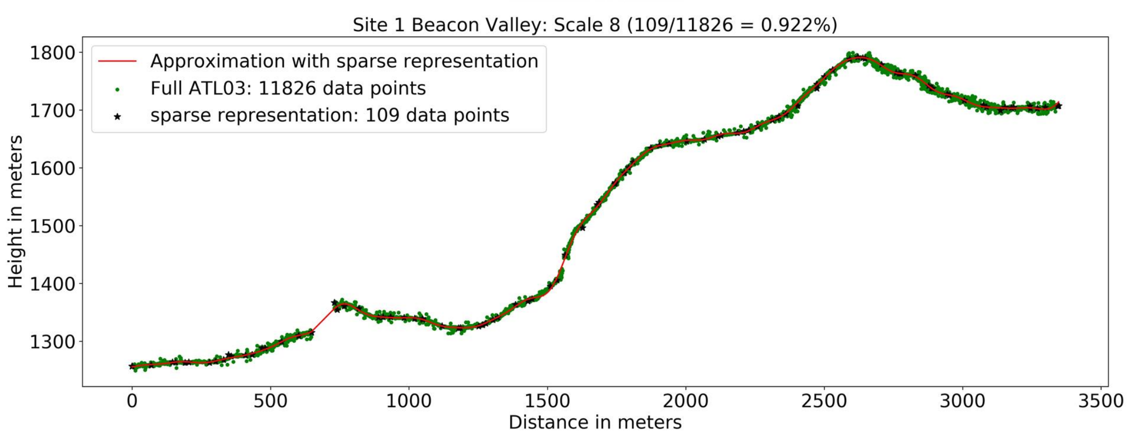
3. Fountain, A.G., Fernandez-Diaz, J. C. , Obryk, M., Levy, J., Gooseff, M., Van Horn, D. J., Morin, P., Shrestha, R., 2017. High-resolution elevation mapping of the McMurdo Dry Valleys, Antarctica, and surrounding regions. Earth System Science Data 9, 435–443.

4. Shekhar, Prashant, and Abani Patra. "Hierarchical approximations for data reduction and learning at multiple scales." Foundations of Data Science 2.2 (2020): 123.

5. Elad, Michael. Sparse and redundant representations: from theory to applications in signal and image processing. Springer Science & Business Media, 2010.

Acknowledgement:

We thank Ivan Parmuzin, UB, for data preprocessing and map compilation. ATL-03 ICESat-2 data are from the National Snow and Ice Data Center and Landsat-8 2020/02/03, composite bands 4,3,2 image is from Earth Explorer (https://earthexplorer.usgs.gov/). Funding for Beata Csatho and Toni Schenk was from NASA Cryospheric Sciences Program in the support of the ICESat-2 mission under the following awards: 80NSSC18K1289 and NNX17AB33G.


Site 1 Beacon Valley: Scale 8 (109/11826 = 0.922%)
Approximation with sparse representation
Full ATL03: 11826 data points
sparse representation: 109 data points
Height in meters
1800
1700
1600
1500
1400
1300
0
500
1000
1500
2000
2500
3000
3500
Distance in meters

AUTHOR INFORMATION

- Prashant Shekhar and Abani Patra are with the Data Intensive Studies Center at Tufts University
- Beata M. Csatho and Toni Schenk are with the Department of Geology, University at Buffalo

ABSTRACT

ICESat-2, the first photon counting satellite, maps the earth's surface with unprecedented details and accuracy. However, the resulting photon clouds, distributed as the ATL-03 geocoded photon product, are vast, unstructured, and noisy datasets. The high density and four-dimensional nature of the photon dataset (location and time), coupled with different responses over different surfaces (e.g., ice, forest cover, water), pose a unique and challenging problem regarding surface detection's overall objective of intelligently reducing the data volume. Multiscale models uncover hidden structures in data due to their ability to analyze the underlying processes at multiple scales. Besides the traditional wisdom of using multiple scales for improving local and global approximations, in this work, we show their application as an intelligent sampling mechanism for redundant and noisy datasets. Our proposed approach's fundamental idea is the generation of data dependent and multiscale basis functions and corresponding representative sparse representations, which retain points essential for minimizing the error of reconstruction. Thereby, points associated with rapid spatial change are chosen, while those that are easily reconstructed using the smooth basis functions are discarded. As the final output, the algorithm provides an efficient sparse representation of the data that captures all relevant features for modeling and prediction with quantified uncertainty. Our presentation includes a detailed description of the algorithm and theory as applied to process the ATL-03 geolocated photon product of the ICESat-2 mission. We will demonstrate the efficiency of our approach by examples of different ice sheet surfaces, including heavily crevassed glaciers, that pose a challenge for currently used change detection methods.

REFERENCES

1. Neumann, Thomas A., et al. "The Ice, Cloud, and Land Elevation Satellite–2 Mission: A global geolocated photon product derived from the advanced topographic laser altimeter system." Remote sensing of environment 233 (2019): 111325

2. Csatho, B., Schenk, T., Krabill, W., Wilson, T., Lyons, W., McKenzie, G., Hallam, C., Manizade, S., Paulsen, T.W., 2005. Airborne laser scanning for high-resolution mapping of Antarctica. Eos Trans. AGU, 86(25), 237-238.

3. Fountain, A.G., Fernandez-Diaz, J. C. , Obryk, M., Levy, J., Gooseff, M., Van Horn, D. J., Morin, P., Shrestha, R., 2017. High-resolution elevation mapping of the McMurdo Dry Valleys, Antarctica, and surrounding regions. Earth System Science Data 9, 435–443.

4. Shekhar, Prashant, and Abani Patra. "Hierarchical approximations for data reduction and learning at multiple scales." Foundations of Data Science 2.2 (2020): 123.

5. Elad, Michael. Sparse and redundant representations: from theory to applications in signal and image processing. Springer Science & Business Media, 2010.