Large-scale processing of satellite LIDAR data and deep learning using satellite imaging data

● **Responsible Representative**

KIMURA Toshiyoshi, Director, Sensor System Research Group, Research and Development Directorate

● **Contact Information**

Rei Mitsuhashi, Yoshito Sawada(mitsuhashi.rei@jaxa.jp)

● **Members**

Rei Mitsuhashi, Yoshito Sawada

● **Abstract**

For the ISS LIDAR mission (MOLI:Multi-footprint Observation Lidar and Imager), we are considering the specified Project Use of JSS3 for the ground processing system due to the characteristics of MOLI project. In this project, we will study the generation process from L0 product data to L1 product in JSS3 and the processing by multi-GPU in JSS3 because deep learning is used for L3 and L4 product generation.

● **Reasons and benefits of using JAXA Supercomputer System**

For the MOLI mission, data from the LIDAR and imager on board the ISS JEM-EF will be stored on HDDs inside the ISS and that HDDs will be transported to the ground when the astronauts return to Earth. Therefore, MOLI require a facility that can process a large amount of observation data in a specified period of time than a regular data processing system. In addition, MOLI products(L3, L4) to be distributed by the G-Portal are generated by using deep learning, which requires large-scale deep learning of a large amount of data. For these reasons, generating MOLI mission products using JSS3 will be very beneficial and we expect to achieve world-class results in deep learning of satellite data using multiple GPUs.

● **Achievements of the Year**

We conducted an airborne experiment in FY2016 to simulate the observation of MOLI mission, and implemented the L1 and L2 products algorithm and processed using that observation data on JSS3 (Fig.1). The parallel computing on JSS3 is expected to process large amounts of Earth obserbasion data at high speed., and we are going to estimate the number of occupied nodes and processing time for creating L1 and L2 products appropriately.

In addition, a prototype of the MOLI research product was developed on the GPU node of JSS3, and a deep learning process was conducted to estimate the forest canopy height from multi-channel image data from the airborne experiments(Fig. 2). we also conduct same prosess using ICESat-2 and GCOM-C/SGLI (Fig. 3). Since

number of MOLI data is expected more than 5 billion LIDAR data, using multiple GPUs of V100 enables us to build a model that cannot be achieved on a desktop PC.
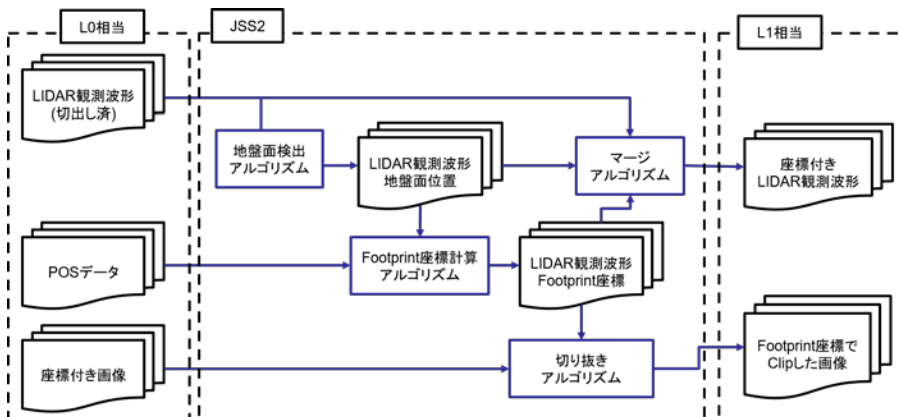

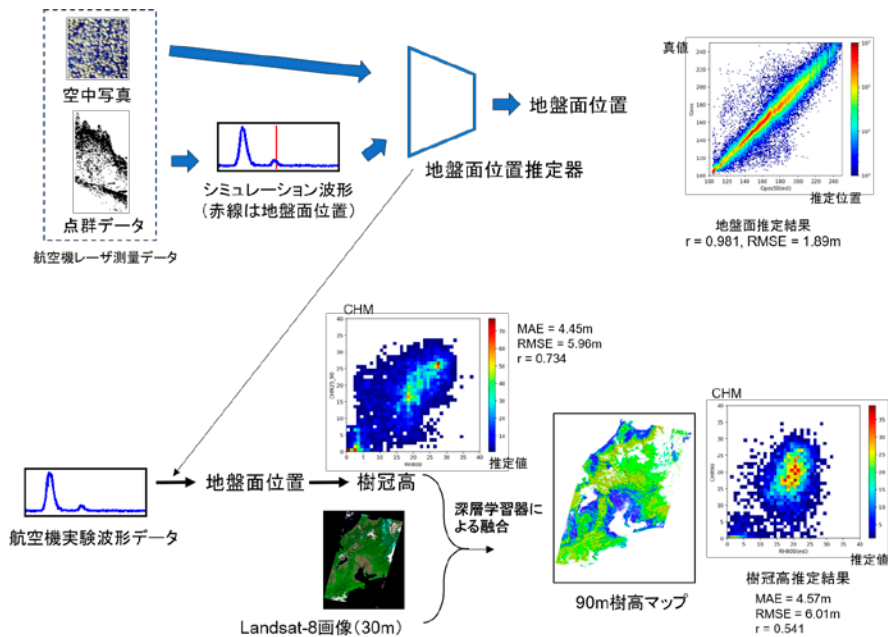
Fig. 1: Flow diagram of L1 prototype algorithm

Fig. 2: L3 and L4 prototype algorithm flow based on airborne experiment data
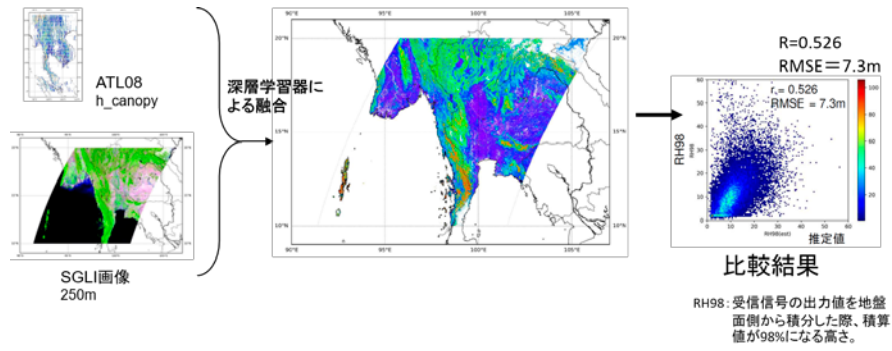
Fig. 3: Tree canopy height map by ICESat-2 and GCOM-C/SGLI with deep learning.

● **Publications**

- Poster Presentations

AGU2020 Fall Meeting 'New ground estimation method applied to GEDI waveforms in Japanese forests'

● **Usage of JSS**

● **Computational Information**

Process Parallelization Methods	MPI
Thread Parallelization Methods	OpenMP
Number of Processes	2 - 4
Elapsed Time per Case	12 Hour(s)

● **Resources Used(JSS2)**

Fraction of Usage in Total Resources[1](%): 0.03

Details

Computational Resources		
System Name	Amount of Core Time (core x hours)	Fraction of Usage[2](%)
SORA-MA	0.00	0.00
SORA-PP	50,401.93	0.40
SORA-LM	0.00	0.00
SORA-TPP	0.00	0.00

File System Resources		
File System Name	Storage Assigned (GiB)	Fraction of Usage[2](%)
/home	9.54	0.01
/data	9,813.31	0.19
/ltmp	1,953.13	0.17

Archiver Resources		
Archiver Name	Storage Used (TiB)	Fraction of Usage[2](%)
J-SPACE	0.00	0.00

[1]: Fraction of Usage in Total Resources: Weighted average of three resource types (Computing, File System, and Archiver).

[2]: Fraction of Usage：Percentage of usage relative to each resource used in one year.

- **Resources Used(JSS3)**

Fraction of Usage in Total Resources[*1](%): 0.06

Details

Computational Resources		
System Name	Amount of Core Time (core x hours)	Fraction of Usage[*2](%)
TOKI-SORA	0.00	0.00
TOKI-RURI	81,466.66	0.47
TOKI-TRURI	0.00	0.00

File System Resources		
File System Name	Storage Assigned (GiB)	Fraction of Usage[*2](%)
/home	9.54	0.01
/data	9,813.31	0.16
/ssd	95.37	0.05

Archiver Resources		
Archiver Name	Storage Used (TiB)	Fraction of Usage[*2](%)
J-SPACE	0.00	0.00

[*1]: Fraction of Usage in Total Resources: Weighted average of three resource types (Computing, File System, and Archiver).

[*2]: Fraction of Usage：Percentage of usage relative to each resource used in one year.