

# Trajectory Subsetter Configuration

---

## VarInfo Configuration:

---

The following collections requires specific configuration to support trajectory subsetting, including both missing CF Convention attributes and some custom pseudo-CF convention attributes required specifically for the trajectory subsetter.

**GEDI\_L[12][AB] | GEDI\_L4A** # trajectory contents, not gridded.

- The /BEAMxxxx/shot\_number variables do not have coordinate references.
  - Add e.g.: coordinates : “delta\_time lat\_lowestmode lon\_lowestmode”
- Also - the shot\_number variables should be required.
  - set **required\_variable** or **ancillary\_variables** references.
- Set the Time-Epoch = ‘2018-01-01T00:00:00.000000’ (for all GEDI)

**GEDI\_L1[AB] | GEDI01\_[AB]** # Just the L1 cases

- For '/BEAM[\d]+/geolocation/' group - all variables in group.
  - Not all geolocation variables have e.g. latitude\_bin0 as a coordinate variable, e.g., altitude\_instrument has latitude\_instrument
  - Resetting the coordinates for subsetting purposes (overrides)
    - Name: ‘coordinates’
    - Value: ‘delta\_time latitude\_bin0 longitude\_bin0’
  - Then, for consistency, move the current coordinate values to create ‘ancillary\_variables’ references.
- This has the effect of ensuring these required support variables are included, but does not actually dictate the subsetting behavior. See Trajectory Subsetter configuration itself (below).
- Note the trajectory subsetter currently treats all variables in a group in the same way and does not handle separate coordinates per variable within the group. A

reimplementation will have to consider if this behavior - all variables in the group ("coincident") have the same subsetting results - but I believe it is the preferred behavior, despite the different coordinates values.

### **GEDI\_L2A | GEDI02\_A** # Just the L2 cases

- For '/BEAM[\\d]+/' and '/BEAM[\\d]+/geolocation/'
- Resetting the spatial coordinates for subsetting purposes
  - Name: 'coordinates'
  - Value: 'delta\_time latitude\_lowestmode longitude\_lowestmode'

### **GEDI L(1[AB] | 2A)** # Just the L1 & L2 cases, not L4

- Set the segment control coordinates.
- On "/BEAM[\\d]+/rxwaveform" and "/BEAM[\\d]+/txwaveform"
- Set coordinates to "(**rx|tx**)\_sample\_start\_index (**rx|tx**)\_sample\_count"
- And for these coordinate variables:
  - set "subset\_control\_type" attribute accordingly (start\_index, index\_count)

### **ICESat-2 ATL0[2-9]|ATL1[0123]** # ICESat-2 trajectory contents, not gridded

- "ProductEpochs": {
 

```
"ATL11": "2000-01-01T00:00:00.000000",
"ATL\\[\\\\\\\\d\\\\]{2}": "2005-01-01T00:00:00.000000",
"GEDI\\_L\\[124\\\\]\\[AB\\\\]": "2018-01-01T00:00:00.000000",
"GLAH\\[\\\\\\\\\\\\d\\\\]{2}": "2005-01-01T00:00:00.000000" }
```
- /gt[123][lr]/geolocation/. \*
  - Add podppd\_flag as a required variable
  - set **required\_variable** or **ancillary\_variables** references

### **ICESat-2 ATL03:**

- For "/gt[123][lr]/geophys\_corr/. \*"

- Resetting the spatial coordinates for subsetting purposes
  - Name: 'coordinates'
  - Value: '.../geolocation/delta\_time, .../geolocation/reference\_photon\_lat  
'.../geolocation/reference\_photon\_lon'

**This is an example of SuperGroup behavior** - these two groups (./geophys\_corr and ./geolocation) are coincident and should be processed as one. The ./geophys\_corr group does not have spatial coordinates (only delta\_time) and without spatial coordinates would not be spatially subset. But it should be subset as ./geolocation is being subset, with the subset-control coordinates as stated in that neighboring group .../geolocation - Verified with NSIDC and Science group.

Note - Other cases of supergroup configuration settings below are GEDI overrides of existing coordinate settings. These are configured as dependent and referenced groups being one and the same, which might seem unnecessary, but this is the only way to set coordinates explicitly (overrides) in the subsetter itself.

```
"SuperGroups": {
  "ATL03": { "/gt\\[\\\\w\\]+/geophys\\_corr/": \[ "/gt\\[\\\\w\\]+/geolo
  "GEDI\\_L\\[24\\]A": { "/BEAM\\[\\\\d\\]+/": \[ "/BEAM\\[\\\\d\\]+/" \]
  "GEDI\\_L2B": { "/BEAM\\[\\\\d\\]+/": \[ "/BEAM\\[\\\\d\\]+/geoloca
  "GEDI\\_L1\\[AB\\]": { "/BEAM\\[\\\\d\\]+/": \[ "/BEAM\\[\\\\d\\]+/geoloca
  "GLAH01": { "/Data\\_40HZ": \[ "/Data\\_1HZ/Geolocation" \] }
  # GLAH01 is a special case where the 1\\_Hz geolocation
  # has to be interpolated to the 40\\_HZ data
```

I.e., a “SuperGroup” is one or more groups that are coindexed and should be subsetted in the same way, with an explicit (overriding) coordinate set.

By default, the Trajectory Subsetter finds groups and attempts to find corresponding lat, lon and time variables for subset control. The subsetter configuration has a listing of corresponding variable names for lat, lon and time that are used in this process.

## ICESat-2 ATL03

- Set the segment control coordinates.
  - On `"/gt[123][lr]/heights/.*`
  - Set coordinates to `".../geolocation/ph_index_beg  
.../geolocation/segment_ph_cnt"`
  - And for these coordinate variables:
    - set `"subset_control_type"` attribute accordingly.

## ICESat-2 ATL08

- Set the segment control coordinates.
  - On `"/gt[123][lr]/signal_photons/.*`
  - Set coordinates to `".../land_segments/ph_ndx_beg  
.../land_segments/n_seg_ph"`
  - And for these coordinate variables:
    - set `"subset_control_type"` attribute accordingly.

## ICESat-2 ATL10

- Set the segment control coordinates # with fwd & rvs segment references.
  - ??? Missing from `TrajectorySubsetter_varinfo_config.json` ???  
but present in `VarInfoConfig15.yml`
- (Full details below...)

## Configuration for Subsetter binary:

---

(See ProductEpochs and SuperGroups above. Also for ATL10 configuration it helps to refer to: [ATL10 Data Architecture and Required Processing](#))

```
"SuperGroupCoordinates": {
  "ATL03": { "/gt[\w]+/geolocation/": [
    "reference_photon_lat", "reference_photon_lon", "delta_time
  "GEDI_L[24]A": { "/BEAM[\d]+/": [
    "lat_lowestmode", "lon_lowestmode", "delta_time" ] },
```

```

"GEDI_L2B": { "/BEAM[\d]+/geolocation/": [
    "lat_lowestmode", "lon_lowestmode", "delta_time" ] },
"GEDI_L1[AB]": { "/BEAM[\d]+/geolocation/": [
    "latitude_bin0", "longitude_bin0", "delta_time" ] },
"GLAH01": { "/Data_1HZ/Geolocation": [ "d1_pred_lat", "d1_pred_lon"

# (revised terminology from current subsetter):
# (SegmentGroups      was PhotonSegmentGroups)
# (SegmentGroup       (no change)
# (SegmentedGroup     was PhotonGroup)
# (SegmentIndexBegin  was PhotonIndexBegin)
# (SegmentIndexCount  was SegmentPhotonCount)
# (SegmentLatitude    was PhotonLatitude)
# (SegmentLongitude   was PhotonLongitude)

"SegmentGroups": {
  "ATL03": {
    "SegmentGroup": "/gt[\w]+/geolocation/",
    "SegmentedGroup": "/gt[\w]+/heights/",
    "SegmentIndexBegin": "ph_index_beg",
    "SegmentIndexCount": "segment_ph_cnt",
    "SegmentLatitude": "lat_ph",
    "SegmentLongitude": "lon_ph"
  },
  "ATL08": {
    "SegmentGroup": "/gt[\w]+/land_segments/",
    "SegmentedGroup": "/gt[\w]+/signal_photons/",
    "SegmentIndexBegin": "ph_ndx_beg",
    "SegmentIndexCount": "n_seg_ph"
  },
  "ATL10": {
    "SegmentGroup": "/freeboard_swath_segment/",
    # (FreeboardSwathSegmentGroup)
    "SegmentGroup": "/gt[\w]+/freeboard_beam_segment/",
    # (FreeboardBeamSegmentGroup)
    "SparseSegmentGroup": "/gt[\w]+/leads/",
    # (LeadsGroup, Uncertain if sparse-segment-group
    # is handled any differently from segment-group)

```

```

"SegmentedGroup": "/freeboard_swath_segment/gt[\\w]+/swath_free|
    # (SwathFreeboardGroup)
"SegmentedGroup": "/gt[\\w]+/freeboard_beam_segment/beam_freeboa
    # (BeamFreeboardGroup)
"SegmentedGroup": "/gt[\\w]+/freeboard_beam_segment/height_segme|
    # (HeightsGroup)
"SegmentedGroup": "/gt[\\w]+/freeboard_beam_segment/geophysical/
    # (GeophysicalGroup)
"SegmentIndexBegin": "fbswath_lead_ndx_gt[\\w]+", # (SwathIndex)
"SegmentIndexCount": "fbswath_lead_n_gt[\\w]+", # (SwathCount)
"SegmentIndexBegin": "fbswath_ndx", # (SwathHeightIndex)
"SegmentIndexCount": "fbswath_n", # (SwathHeightCount)
"SegmentIndexBegin": "beam_lead_ndx", # (BeamIndex)
"SegmentIndexCount": "beam_lead_n", # (BeamCount)
"SegmentIndexBegin": "ssh_ndx", # (LeadsIndex)
"SegmentIndexCount": "ssh_n", # (LeadsCount)
"SegmentIndexBegin": "beam_refsur_ndx|beam_refsurf_ndx",
    # (BeamFreeboardIndex)
"SegmentIndexCount": "beam_refsur_n", # (BeamFreeboardCount)
"SegmentLatitude": "latitude", # (PhotonLatitude)
"SegmentLongitude": "longitude", # (PhotonLongitude)
"HeightSegmentSSHFlag": "height_segment_ssh_flag"
},
"GEDI_L1[AB]": {
    "SegmentGroup": "/BEAM[\\d]+/",
    "SegmentedDataset": "/BEAM[\\d]+/[rt]xwaveform", # (PhotonDataset)
    "SegmentIndexBegin": "[rt]x_sample_start_index", # (PhotonIndex)
    "SegmentIndexCount": "[rt]x_sample_count" # (SegmentPhotonCount)
},
"GEDI_L2B": {
    "SegmentGroup": "/BEAM[\\d]+/",
    "SegmentedDataset": "/BEAM[\\d]+/pgap_theta_z", # (PhotonDataset)
    "SegmentIndexBegin": "rx_sample_start_index", # (PhotonIndex)
    "SegmentIndexCount": "rx_sample_count" # (SegmentPhotonCount)
}

```

- Note - Reverse Segment References (\_id) are not identified in the subsetter configuration, but do require special handling (recomputed index values after

subsetting). This requires the code to identify the relevant group paths and the `_id` variable without configuration settings.

### ATL10 Configuration (varinfo.yml):

- In a reimplementaion I would set the coordinates attribute and not use `subset_control_variables`.
  - Where coordinates exist (overrides), I would move the coordinate value to the `ancillary_variables` attribute to ensure these coordinate variables are handled as required support variables
- Applicability:
- ```
Variable_Pattern: '../gt[123][lr]/leads'
# these datasets have two reference datasets, either pair can be used
# for subsetting purposes, but both must be recomputed after subsetting
# Thus two are listed for subset control, but all four are listed
# as segment control variables
```
- Attributes:
- ```
- Name: 'subset_control_variables'
  Value: '/freeboard_swath_segment/fbswath_lead_ndx_gt[123][lr]
         /freeboard_swath_segment/fbswath_lead_n_gt[123][lr]
        # or:
        #   '/gt[123][lr]/freeboard_beam_segment/beam_lead_ndx
        #   '/gt[123][lr]/freeboard_beam_segment/beam_lead_n'
        # Note - ground-track ids (gt[123][lr]) have to match parent ground-track
```
- Applicability:
- ```
Variable_Pattern: '/freeboard_swath_segment/fbswath_lead_ndx_gt[123][lr]
```
- Attributes:
- ```
- Name: 'segment_control_variable_type'
  Value: 'segment_index_beg'
```
- Applicability:
- ```
Variable_Pattern: '/freeboard_swath_segment/fbswath_lead_n_gt[123][lr]
```
- Attributes:
- ```
- Name: 'segment_control_variable_type'
  Value: 'segment_index_cnt'
```
- Applicability:
- ```
Variable_Pattern: '/gt[123][lr]/freeboard_beam_segment/beam_lead_ndx'
```

```
Attributes:
  - Name: 'segment_control_variable_type'
    Value: 'segment_index_beg'
- Applicability:
  Variable_Pattern: '/gt[123][lr]/freeboard_beam_segment/beam_lead_n'
Attributes:
  - Name: 'segment_control_variable_type'
    Value: 'segment_index_cnt'

- Applicability:
  Variable_Pattern: '/gt[123][lr]/freeboard_beam_segment/[^/]*'
  # A reverse segment reference case, excluding nested elements
Attributes:
  - Name: 'subset_control_variables'
    Value: 'fbswath_ndx'
- Applicability:
  Variable_Pattern: '/gt[123][lr]/freeboard_beam_segment/fbswath_ndx'
Attributes:
  - Name: 'segment_control_variable_type'
    Value: 'rvs_segment_index'
  - Name: 'rvs_segment_coordinates'
    Value: '/freeboard_swath_segment/delta_time
           /freeboard_swath_segment/latitude /freeboard_swath_segment/'

- Applicability:
  Variable_Pattern: '/freeboard_swath_segment/gt[\\w]+/swath_freeboard/'
Attributes:
  - Name: 'subset_control_variables'
    Value: 'fbswath_ndx'
- Applicability:
  Variable_Pattern: '/freeboard_swath_segment/gt[\\w]+/swath_freeboard/'
Attributes:
  - Name: 'segment_control_variable_type'
    Value: 'rvs_segment_index'
  - Name: 'rvs_segment_coordinates'
    Value: '/freeboard_swath_segment/delta_time
           /freeboard_swath_segment/latitude /freeboard_swath_segment/'
```



```

- Applicability:
  Variable_Pattern: '../gt[123][lr]/leads/ssh_ndx'
  Attributes:
    - Name: 'segment_control_variable_type'
      Value: 'sparse_segment_index_beg'
- Applicability:
  Variable_Pattern: '../gt[123][lr]/leads/ssh_n'
  Attributes:
    - Name: 'segment_control_variable_type'
      Value: 'sparse_segment_index_cnt'

- Applicability:
  Variable_Pattern: '/gt[123][lr]+/freeboard_beam_segment/beam_freeboard'
  Attributes:
    - Name: 'subset_control_variables'
      Value: 'beam_refsur_ndx'
- Applicability:
  Variable_Pattern: '/gt[123][lr]+/freeboard_beam_segment/beam_freeboard'
  Attributes:
    - Name: 'segment_control_variable_type'
      Value: 'rvs_segment_index'
    - Name: 'rvs_segment_coordinates'
      Value: '/gt[123][lr]+/freeboard_beam_segment/delta_time
              /gt[123][lr]+/freeboard_beam_segment/latitude
              /gt[123][lr]+/freeboard_beam_segment/longitude'
# '../gt[123][lr]/leads/ssh_ndx' and '../gt[123][lr]/leads/ssh_n'
# already have forward segment_control_variable_type attribute settings al

```