Summary of HSRL Subset Data File

14 August 2018

Table of Contents:

Overview:

This a draft document that will be continually updated but is intended to help with reading and using the HSRL data sets.

This document provides information on the HSRL HDF subset data files. The data are stored in HDF5 format and the files contain all the processed data products for each flight. The main purpose of this document is to provide a simple and brief overview of the variables that are contained in the file. This document does not contain any special conditions that may have been included for specific flights. It is also recommended that users of the HSRL data products contact the HSRL team before performing their research. This document will periodically be updated and the user should request the latest version available from the HSRL team. Contact Emails: Johnathan.W.Hair@nasa.gov; Richard.A.Ferrare@nasa.gov; Chris.A.Hostetler@nasa.gov; Amy.Jo.Scarino@nasa.gov .

The data is stored by ***profiles*** which has typically been hardware averaged to 0.5 seconds. The range sampling for the HSRL-1 data is currently set at ~1.25m range ***bins***. Two HDF5 files are created, one that retains the high bandwidth resolution (hereafter referred to as the **HBW** file) of 1.25m. The second low bandwidth resolution file (hereafter referred to as the **LBW** file) applies a digital filter of 30m to the 532nm data channels (the 1064nm data channels have an electronic filter of 30m) and data are interpolated to an oversampled ~15m vertical resolution. Therefore the signals have 30m resolution with a 15m sampling interval. In both files (HBW and LBW) the raw data is altitude aligned based on the aircraft GPS altitude and interpolated to standard sets of altitude range bins that begin at -300m below sea level and extend to 150m above the highest altitude during the flight. The altitude vector that the data is interpolated to is stored in the file at DataProducts /Altitude in each respective file.

All of the atmospheric data products are calculated from the 15m interpolated altitudes stored in the LBW file. However, the ocean data products are calculated from the ~1.25m range bin signals which have been re-registered to the retrieved surface bin to ~500m above sea level and stored in the HBW file. In the case of the ocean products, the vector DataProducts /Depth in the HBW files represents the vertical extent of the analyzed data products – depth has been corrected for the change in the index of refraction of water below the ocean surface. All averaging is done in a running (filtered) average which maintains the same number of profiles and interpolated range bins. Note that the averaged data products will not be statistically independent. All averaging intervals are recorded in the file (see table UserInputs).

The subset HDF5 files are a combination of the LBW atmospheric and HBW ocean files typically sub-sampled at 10 seconds. Atmospheric products are under DataProducts and ocean products are found in the folder OceanDataProducts. The ocean products are truncated at the ocean surface which is reflected in the accompanying OceanDataProducts/Depth vector.

Reduced and Subset HDF Files

The 20 channels (only 16 channels used) of raw data signal returns are gridded and stored along with the various data products at resolutions described previously in the LBW (atmospheric) and HBW (ocean) files. The engineering, navigation, gains, user input and state parameter data are saved in each file.

The HSRL analyzed data files are sub-sampled for distribution due to the large file size associated with the full HSRL HDF files (~12.5 Gb/hour of flight for the HBW ocean data files and ~2 Gb/hour of flight for the LBW atmospheric data files). The subset file contains all of the calculated variables as well as relevant metadata, however the engineering, gain and raw data are not included. The data are also decimated to further reduce the file size. The amount of decimation depends on the backscatter product's temporal average, *532_bs_time_avg*, and is chosen so that only one profile in each time average is included in the subset. For example if the raw HSRL data are sampled at 0.5 s resolution and a 10 s average is applied in the backscatter calculation, the file is decimated by a factor of 20. The atmospheric and ocean extinction product, which typically has a longer temporal average than the backscatter and depolarization products, is decimated by the same amount as the other products to preserve the array size. *Therefore, please note that the extinction product is over sampled in the subset file.* No sub-setting is performed along the height dimension of these products.

The following tables use these notations for size:

nr = number of records in hdf5 file
plen = profile length for atmospheric data products
polen = profile length for ocean data products

State Parameter Variables: Input data from balloon sondes or Meteorological Models (GMAO)

State parameters used in the retrievals are input from either balloon sondes or meteorological models. Currently there are several model options and the filename is provided in the user input parameters. Most preliminary analyses use the balloon data for the entire flight. Post processed data typically use the GMAO state data interpolated to each altitude, latitude and longitude for each profile.

	Variable Name	*Type*	*Size*	*Units*	*Description*
State	Number_Density	double	[plen nr]	m^-3	Molecular Number Density
	O3	double	[plen nr]	kg/kg	GMAO ONLY: O3 mixing ratio profile
	Pressure	double	[plen nr]	atm	Pressure
	Relative_Humidity	double	[plen nr]	%	Relative Humidity in percent
	State_Type	uint16	[1 1]	flag	1=sonde, 2=model (GMAO)
	Temperature	double	[plen nr]	K	Temperature
	U	double	[plen nr]	m/s	GMAO ONLY: U (east) component of wind
	V	double	[plen nr]	m/s	GMAO ONLY: V (north) component of wind

Aircraft position information from Nav_Data

	Variable Name	Type	Size	Units	Description
Nav_Data	HeadingAccuracy	double	[1 nr]	deg	TBD
	HeadingFlag	double	[1 nr]	none	TBD
	IMUFlag	double	[1 nr]	none	TBD
	PitchAccuracy	double	[1 nr]	deg	TBD
	RollAccuracy	double	[1 nr]	deg	TBD
	TrueHeading	double	[1 nr]	deg	Aircraft TRUE heading
	TrueVehicleTrack	double	[1 nr]	deg	Aircraft track heading
	gps_alt	double	[1 nr]	m	Altitude of aircraft
	gps_date	double	[1 nr]	none	Date
	gps_fixquality	double	[1 nr]	none	Quality flag from Applanix GPS data stream
	gps_geoid_alt	double	[1 nr]	m	Altitude of geoid
	gps_gnd_speed_kmph	double	[1 nr]	kmph	Ground speed
	gps_gnd_speed_knts	double	[1 nr]	knots	Ground speed
	gps_heading	double	[1 nr]	none	Aircraft heading
	gps_horz_dilution	double	[1 nr]	none	TBD
	gps_lat	double	[1 nr]	deg	Latitude
	gps_lon	double	[1 nr]	deg	Longitude
	gps_num_satellites	double	[1 nr]	none	Number of satellites acquired from GPS
	gps_time	double	[1 nr]	hrs	UTC hours
	imu_pitch	double	[1 nr]	deg	Aircraft pitch angle
	imu_roll	double	[1 nr]	deg	Aircraft roll angle
	imu_x_vel	double	[1 nr]	m/s	TBD
	imu_y_vel	double	[1 nr]	m/s	TBD
	imu_z_vel	double	[1 nr]	m/s	TBD

Science Data Products: Analyzed data products

The backscatter ratios are defined to be the ratio of the backscatter coefficient to the Cabannes portion of the molecular backscatter coefficient. The total molecular backscatter coefficient is 2.5 % greater due to the rotational Raman signal as noted by She, 2001:

She, C. 2001, "Spectral Structure of Laser Light Scattering Revisited: Bandwidths of Nonresonant Scattering Lidars." Applied Optics LP, vol. 40, Issue 27, pp.4875-4884.

Data Products - Atmospheric

	Variable Name	Type	Size	Units	Description
DataProducts	1064_aer_dep	double	[plen nr]	ratio	Retrieved 1064nm aerosol depolarization ratio. This parameter is calculated by taking the ratio of the perpendicular (cross polarized) and parallel (co-polarized) *aerosol* backscatter coefficients. $$1064\_aerdep = \frac{\beta^{\perp}_{aer.}}{\beta^{\parallel}_{aer.}}$$
	1064_bsc	double	[plen nr]	km^-1sr^-1	Retrieved aerosol volume backscatter coefficient at 1064nm, using the Fernald solution and the 532 β_a in a clean region. $$1064\_bsc = \beta_{aer} = \beta^{\parallel}_{aer} + \beta^{\perp}_{aer}$$
	1064_bsc_Sa	double	[plen nr]	km^-1sr^-1	1064 nm aerosol backscatter ratio used for Sa calculation (averaged similar to extinction coeff.)
	1064_bsc_cloud_screened	double	[plen nr]	km^-1sr^-1	Aerosol volume backscatter coefficient at 1064nm with cloud mask applied
	1064_bsr	double	[plen nr]	ratio	Retrieved ratio of aerosol backscatter coefficient to the molecular backscatter coefficient at 1064nm, using the Fernald solution and the 532 β_a in a clean region. $$1064\_bsr = \frac{\beta_{aer}}{\beta_{mol}} = \frac{\beta^{\parallel}_{aer} + \beta^{\perp}_{aer}}{\beta^{\parallel}_{mol} + \beta^{\perp}_{mol}}$$
	1064_bsr_cloud_screened	double	[plen nr]	ratio	1064 backscatter ratio with cloud mask applied
	1064_dep	double	[plen nr]	ratio	Retrieved 532nm total (volume) depolarization ratio. This parameter includes both molecular and aerosol scattering and is calculated by taking the ratio of the perpendicular (cross polarized) and parallel *total* scattering channels. The parallel and perpendicular signals are corrected for optical and electronic gains via an internal calibration made during flight. $$1064\_dep = \frac{\beta^{\perp}_{mol.} + \beta^{\perp}_{aer.}}{\beta^{\parallel}_{mol.} + \beta^{\parallel}_{aer.}}$$
	1064_ext	double	[plen nr]	km^-1	Retrieved 1064 aerosol extinction coefficient using scaled Sa values from the 532 channel. In some campaigns this may only use a constant value of Sa for the retrievals.
	1064_total_attn_bsc	double	[plen nr]	km^-1sr^-1	1064 attenuated backscatter coefficient. The data are normalized to the calculated aerosol backscatter + molecular in 1.0km to 1.5km from the aircraft altitude. $$1064\_total\_attn\_bsc = \frac{\beta^{\parallel}_{aer} + \beta^{\perp}_{aer}}{\beta_{norm}}$$

532_AOT_hi	double	[1 nr]	none	532nm Aerosol optical thickness determined from the molecular channel near the aircraft and near the surface. The lower point uses a fit over a small range of range bins to determine the signal level near the surface and the fit is extrapolated to the ground.
532_AOT_hi_col	double	[plen nr]	none	532nm Aerosol optical thickness as a function of altitude as determined from the molecular signal. The AOT value in the lowest altitude bin is filled with AOT_hi.
532_AOT_lo	double	[1 nr]	none	532nm Aerosol optical thickness determined from the extinction profile and also using the backscatter and Sa value near the surface.
532_Sa	double	[plen nr]	sr	Retrieved extinction-to-backscatter ratio at 532nm. The product is averaged to the same horizontal and vertical average as the extinction coefficient.
532_aer_dep	double	[plen nr]	ratio	Retrieved 532nm aerosol depolarization ratio. This parameter is calculated by taking the ratio of the perpendicular (cross polarized) and parallel *aerosol* backscatter coefficients. $$532\_dep = \frac{\beta^{\perp}_{aer.}}{\beta^{\parallel}_{aer.}}$$
532_bsc	double	[plen nr]	km^-1sr^-1	Aerosol volume backscatter coefficient at 532nm $$532\_bsc = \beta_{aer} = \beta^{\parallel}_{aer} + \beta^{\perp}_{aer}$$
532_bsc_Sa	double	[plen nr]	km^-1sr^-1	Aerosol volume backscatter coefficient at 532nm. This backscatter coefficient is different from above due to the fact that it is averaged to the same horizontal and vertical average as the extinction coefficient and is used to calculate extinction-to-backscatter ratio.
532_bsc_cloud_screened	double	[plen nr]	km^-1sr^-1	Aerosol volume backscatter coefficient at 532nm with cloud mask applied.
532_bsr	double	[plen nr]	ratio	Ratio of Aerosol backscatter coefficient to the molecular backscatter coefficient at 532nm. The total scattering ratio is calculated by adding one. $$532\_bsr = \frac{\beta_{aer}}{\beta_{mol}} = \frac{\beta^{\parallel}_{aer} + \beta^{\perp}_{aer}}{\beta^{\parallel}_{mol} + \beta^{\perp}_{mol}}$$
532_bsr_cloud_screened	double	[plen nr]	ratio	Backscatter ratio with cloud mask applied.

532_dep	double	[plen nr]	ratio	Retrieved 532nm total (volume) depolarization ratio. This parameter includes both molecular and aerosol scattering and is calculated by taking the ratio of the perpendicular (cross polarized) and parallel (co-polarized) *total* scattering channels. The parallel and perpendicular signals are corrected for optical and electronic gains via an internal calibration made during flight. $$532\_dep = \frac{\beta^{\perp}_{mol.} + \beta^{\perp}_{aer.}}{\beta^{\parallel}_{mol.} + \beta^{\parallel}_{aer.}}$$
532_ext	double	[plen nr]	km^-1	Retrieved 532nm aerosol extinction coefficient. Values closer than 2500m to the aircraft are not calculated due to overlap. $$532\_ext = -\frac{1}{2}\frac{\partial}{\partial r}\ln\left\{\frac{P_{mol}r^2}{F\beta_{mol}}\right\} - a_m$$
532_total_attn_bsc	double	[plen nr]	km^-1sr^-1	Attenuated backscatter coefficient. The data are normalized to the calculated aerosol backscatter + molecular in 1.0km to 1.5km from the aircraft altitude. $$532\_total\_attn\_bsc = \frac{\beta^{\parallel}_{aer} + \beta^{\perp}_{aer}}{\beta_{norm}}$$
Aerosol_ID	double	[plen nr]	none	Aerosol type per S. Burton's typing algorithms: 1=ice, 2=dusty mix, 3=maritime, 4=urban, 5=smoke, 6=fresh smoke, 7=polluted maritime, 8=pure dust.
Altitude	double	[plen 1]	m	Altitude vector for data products
Angstrom_Dust	double	[plen nr]	sr	Angstrom coefficient of dust particles calculated from Eq 17 of Sugimoto and Lee, Appl Opt 2006 (with A and B defined as in Eq. 16) $$\gamma_d = \ln\left(\frac{A}{B}\right) \Big/ \ln(2)$$
Angstrom_Spherical	double	[plen nr]	sr	Angstrom coefficient of spherical particles calculated from Eq 18 of Sugimoto and Lee, Appl Opt 2006 (using values defined in Eq.18 and Eq. 13): $$\gamma_s = \ln\left(\frac{1-X}{R_\beta - XB/A}\right) \Big/ \ln(2)$$
Dust_Mixing_Ratio	double	[plen nr]	sr	Dust mixing ratio, X, calculated from Eq 7 of Sugimoto and Lee, Appl Opt 2006: $$X = \left(\frac{(1+\delta_{d532})\cdot\delta_{532}}{\delta_{d532}\cdot(1+\delta_{532})}\right) \qquad \delta_{d532} = 0.35$$

	WVD_1064_532	double	[plen nr]	ratio	Retrieved wavelength dependence based on the aerosol backscatter coefficient at 1064 and 532nm. $WVD\_1064\_532 = -\ln\left(\frac{\beta_{aer}^{1064nm}}{\beta_{aer}^{1532nm}}\right) / \ln\left(\frac{1064nm}{532nm}\right)$
	cloud_top_height	double	[1 nr]	km	Cloud identification flag and cloud height information
	mask_low	double	[plen nr]	none	Mask for all data products based on the molecular backscatter signal. A signal threshold over background light is used to determine when there is no signal to produce retrieved products.

Data Products - Ocean

Note that polen is a subset of the high resolution plen and spans from -224m (below surface) to ~500m (above surface.) The HPD designation in the variable names refers to hybrid photodiode detectors used for these measurements. Signals sensed concurrently by photomultiplier tube detectors were also recorded and data products from these detectors are available in the full set HBW files but not in the subset files.

	Variable Name	Type	Size	Units	Description
OceanDataProducts	Depth	double	[polen 1]	m	Altitude of ocean data products includes index of refraction correction below surface for water
	HPD_Kd_slope	double	[1 nr]	m^-1	532 diffuse lidar attenuation coefficient-based derivative of molecular signal from HPD
	HPD_ocean_aer_dep	double	[polen nr]	ratio	532 particulate depolarization ratios derived from HPD $532\_aerdep = \frac{\beta^{\perp}_{par}}{\beta^{\parallel}_{par}}$
	HPD_ocean_bbp	double	[polen nr]	m^-1	532 particulate hemispherical backscatter coefficients derived from HPD $b_{bp}=\chi_p\ (2\pi)\ \ \beta_{aer}$ Value for χ_p currently uses 0.5
	HPD_ocean_bsc	double	[polen nr]	m^-1sr^-1	532 particulate backscatter coefficients derived from HPD $ocean\_bsc = \beta_{par} = \left(\beta^{\parallel}_{par} + \beta^{\perp}_{par}\right)$
	HPD_ocean_bsr	double	[polen nr]	ratio	532 particulate backscatter scattering ratios derived from HPD $ocean\_bsr = \frac{\beta_{par}}{\beta_{water}} = \frac{\beta^{\parallel}_{par} + \beta^{\perp}_{par}}{\beta^{\parallel}_{water} + \beta^{\perp}_{water}}$
	HPD_ocean_dep	double	[polen nr]	ratio	532 total depolarization ratios derived from HPD $ocean\_dep = \frac{\beta^{\perp}_{water} + \beta^{\perp}_{par}}{\beta^{\parallel}_{water} + \beta^{\parallel}_{par}}$
	HPD_ocean_ext	double	[polen nr]	m^-1	532 diffuse lidar attenuation coefficient using HPD channel $ocean\_ext = -\frac{1}{2}\frac{\partial}{\partial r}\ln\left\{\frac{P_{mol}r^2}{F\beta_{water}}\right\}$ β_{water} = 2.45e-4 m^-1sr^-1 = molecular 180 backscatter at 532 nm
	HPD_ocean_mask_low	double	[polen nr]	none	Low signal mask index for the ocean returns based on HPD Molecular Signal

User input parameters: input variables for data analysis

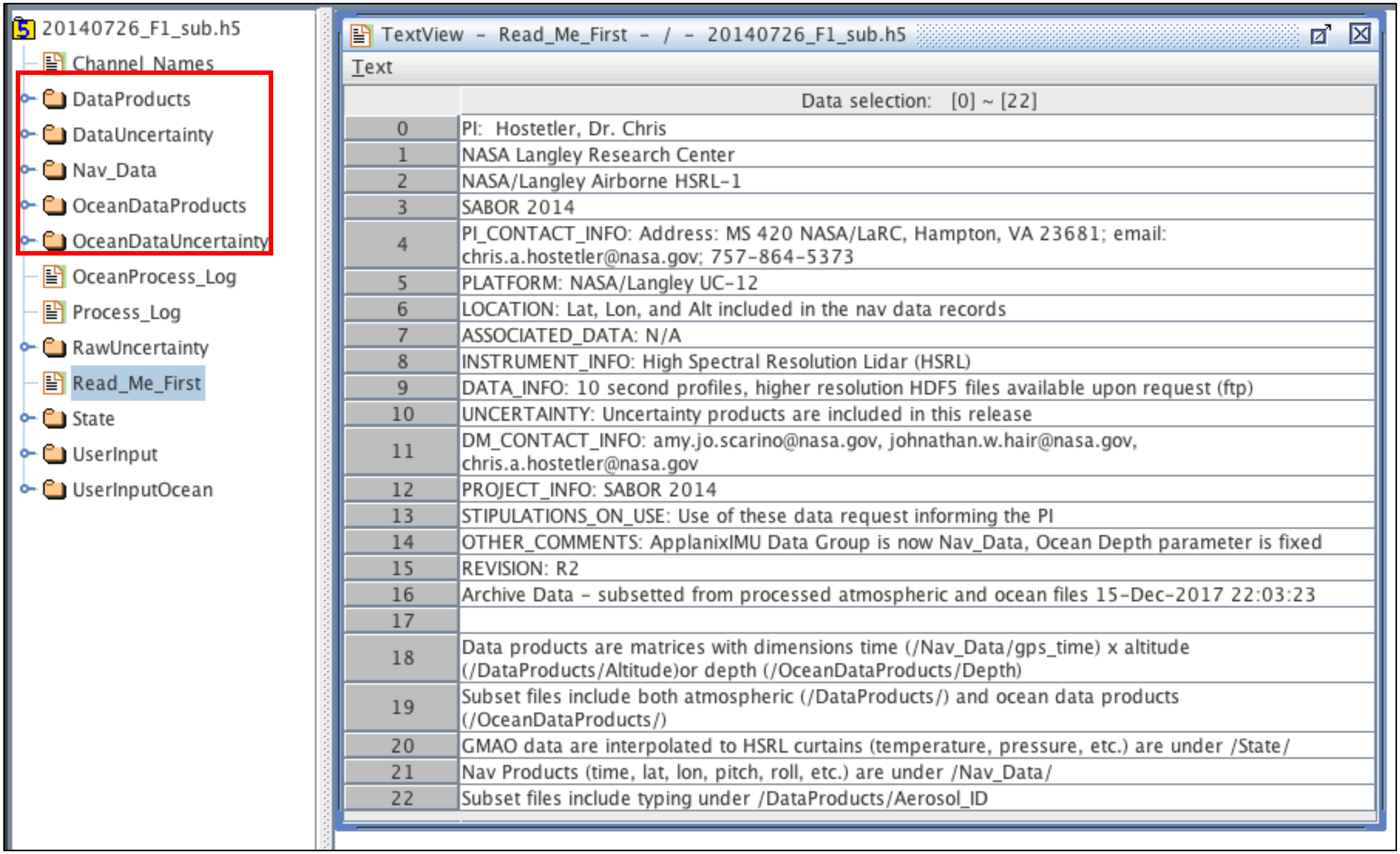
	Variable Name	Type	Size	Units	Description
UserInput	DEM_altitude	double	[1 nr]	m	Altitude of GLOBE digital elevation map including GEBCO ocean depths
	range_interp	double	[plen nr]	m	Range of interpolated lidar signal at same grid as data products
	filter_Brillouin_HPD	double	[1 1]	none	Iodine filter transmission coefficients for ocean Brillouin scattering in HPD channel
	ocean_backscatter_chi	double	[1 1]	none	Ocean backscatter chi
	ocean_water_backscatter	double	[1 1]	none	Ocean water backscatter (Morel)
	ocean_etalon_attenuation	double	[1 1]	none	Etalon attenuation
	ocean_raman_fraction	double	[1 1]	none	Raman attenuation fraction
	range_offset	double	[1 1]	m	Offset in high resolution bins between start of data and laser fire
	tilt_angle	double	[1 1]	degrees	Mounting pitch angle of the instrument (determined empirically)
	seed_lock_offset	double	[1 1]	GHz	Offset from seed lock line
	532_bs_time_avg	single	[1 1]	sec	Time average for 532nm backscatter coefficient
	532_bs_range_avg	single	[1 1]	m	Range average for 532nm backscatter coefficient
	532_ext_time_avg	single	[1 1]	sec	Time average for 532nm extinction coefficient
	532_ext_range_avg	single	[1 1]	m	Range average for 532nm extinction coefficient
	532_depol_time_avg	single	[1 1]	sec	Time average for 532nm total depolarization ratio
	532_depol_range_avg	single	[1 1]	m	Range average for 532nm total depolarization ratio
	1064_depol_time_avg	single	[1 1]	sec	Time average for 1064nm total depolarization ratio
	1064_depol_range_avg	single	[1 1]	m	Range average for 1064nm total depolarization ratio
	1064_bs_time_avg	single	[1 1]	sec	Time average for 1064nm backscatter coefficient
	1064_bs_range_avg	single	[1 1]	m	Range average for 1064nm backscatter coefficient
	1064_calibration	double	[3 nr]	none	Array of the 1064 calibration start altitude, stop altitude, and 532 bsr
	offset_angle	double	[1 nr]	none	Cosine of the angle of the mount plus pitch and roll

Maximum and minimum altitudes of calculated products

The maximum and minimum altitudes for the calculated variables are listed in Table 1

Table 1

Variable Name	*Minimum Altitude*	*Maximum Altitude*
532_ext	(filter window – 2 bins) above (ground + 2 bins).	2.5 km below aircraft
532_bsr	2 bins above ground	500 m below aircraft
532_bsc	2 bins above ground	500 m below aircraft
532_total_attn_bsc		1.5 km below aircraft
532_dep	2 bins above ground	500 m below aircraft
532_aer_dep	2 bins above ground	500 m below aircraft
1064_ext	2 bins above ground	750 m below aircraft
1064_bsr	2 bins above ground	750 m below aircraft
1064_bsc	2 bins above ground	750 m below aircraft
1064_total_attn_bsc		1.5 km below aircraft
1064_dep	2 bins above ground	500 m below aircraft
1064_aer_dep	2 bins above ground	500 m below aircraft
Sa_532	(filter window – 2 bins) above (ground + 2 bins).	2.5 km below aircraft
WVD_1064_532	2 bins above ground	500 m below aircraft
AOT_lo	2 bins above ground	2.5 km below aircraft
AOT_hi	Extrapolated to ground	2.5 km below aircraft
AOT_hi_col	Extrapolated to ground	2.5 km below aircraft



	Data selection: [0] ~ [22]
0	PI: Hostetler, Dr. Chris
1	NASA Langley Research Center
2	NASA/Langley Airborne HSRL-1
3	SABOR 2014
4	PI_CONTACT_INFO: Address: MS 420 NASA/LaRC, Hampton, VA 23681; email: chris.a.hostetler@nasa.gov; 757-864-5373
5	PLATFORM: NASA/Langley UC-12
6	LOCATION: Lat, Lon, and Alt included in the nav data records
7	ASSOCIATED_DATA: N/A
8	INSTRUMENT_INFO: High Spectral Resolution Lidar (HSRL)
9	DATA_INFO: 10 second profiles, higher resolution HDF5 files available upon request (ftp)
10	UNCERTAINTY: Uncertainty products are included in this release
11	DM_CONTACT_INFO: amy.jo.scarino@nasa.gov, johnathan.w.hair@nasa.gov, chris.a.hostetler@nasa.gov
12	PROJECT_INFO: SABOR 2014
13	STIPULATIONS_ON_USE: Use of these data request informing the PI
14	OTHER_COMMENTS: ApplanixIMU Data Group is now Nav_Data, Ocean Depth parameter is fixed
15	REVISION: R2
16	Archive Data - subsetted from processed atmospheric and ocean files 15-Dec-2017 22:03:23
17	
18	Data products are matrices with dimensions time (/Nav_Data/gps_time) x altitude (/DataProducts/Altitude)or depth (/OceanDataProducts/Depth)
19	Subset files include both atmospheric (/DataProducts/) and ocean data products (/OceanDataProducts/)
20	GMAO data are interpolated to HSRL curtains (temperature, pressure, etc.) are under /State/
21	Nav Products (time, lat, lon, pitch, roll, etc.) are under /Nav_Data/
22	Subset files include typing under /DataProducts/Aerosol_ID

HDFView of subset file and folders – Read_Me_First section shown