Summary of HALO Subset Data File

22 July 2020

Table of Contents:

Overview:

This a draft document that will be continually updated but is intended to help with reading and using the HALO data sets for ACT-America 2019.

This document provides information on the HALO 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 HALO data products contact the HALO team before performing their research. Contact Emails: Amin.R.Nehrir@nasa.gov; James.E.Collins@nasa.gov; Susan.A.Kooi@nasa.gov; Rory.A.Barton-Grimley@nasa.gov.

HALO is an airborne multi-function Differential Absorption Lidar (DIAL) and High Spectral Resolution Lidar (HSRL). HALO has multiple configurations intended to support different NASA airborne science focus areas. For ACT-America (summer 2019), HALO was configured in the methane DIAL/HSRL configuration to measure profiles of aerosol optical properties, distributions of mixed layer heights (MLH), as well as columns of tropospheric methane. The archive only contains the HSRL aerosol and MLH information as the methane retrievals are still under development. The HALO HSRL and backscatter aerosol channels operate at 532 nm and 1064 nm, respectively. The ACT-America summer 2019 campaign took place over eastern and central North America based from Shreveport, La., Lincoln, Ne., and NASA Wallops Flight Facility located on the eastern shore of Virginia.

The data are stored by ***profiles*** which are hardware averaged to 0.5 seconds (2Hz) native resolution. The vertical sampling resolution for the HALO raw data is 1.25m. For aerosol profiling these high vertical resolution bins are filtered and binned to 15 m to create a lower resolution and higher signal to noise ratio raw signal used for to calculate the aerosol optical properties described below. Aerosol backscatter and depolarization products which include spectral depolarization ratios, angstrom exponents, and dust mixing ratio are averaged 10 s horizontally with a vertical resolution of 15m, while aerosol extinction products are averaged 60 s horizontally and ~300 m vertically. Aerosol optical thickness (AOT) is reported in 3 data sets. *532_AOT_lo* data are based on 532 extinction data and have the same horizontal and vertical resolution. *532_AOT_hi* and *532_AOT_hi_col* have a 10 s horizontal resolution but 300 m vertical averaging. *532_AOT_hi* is a single value for the AOT to the ground, while *532_AOT_hi_col* data are profiles of AOT. The polarization and HSRL gain ratios are calculated as described in Hair et al., 2008. Operational retrievals also provide mixing ratio of non-spherical-to-spherical backscatter (Sugimoto and Lee, 2006), aerosol type (Burton et al., 2012) and aerosol mixed-layer height (Scarino et al., 2014). The raw data are quality controlled by applying a cloud screening mask to remove attenuated signals below clouds. For the mixing layer height product, which is the principal lidar observable for ACT-America, the retrievals are quality controlled beyond the methods described in Scarino et al. 2014 by applying a user defined and time dependent threshold on the wavelet transform. To increase the precision of the mixing layer height retrievals, a ±3 point gliding or moving window (60 second window) is applied to 10 second resolution data. All data products are archived in an H5 file format with 10 second horizontal resolution and ranged reported data using an altitude array that is the same for all these data at ~15 m regardless of vertical averaging.

Parameter	Wavelength (nm)	Approximate Precision	Horizontal Resolution	Vertical Resolution
Aerosol Backscatter	532/1064	0.2 Mm-1sr-1	10 s	15 m
Aerosol Extinction	532	0.01 km -1	60 s	300 m
Depolarization	532/1064	0.01	10 s	15 m
Aerosol Optical Thickness	532	0.01	60s/10 s	N/A
Boundary Layer Height	532	Varies with conditions	60 s	15 m
Aerosol Type (e.g., marine, dust, smoke)	N/A	Qualitative	60 s	300 m

All of the aerosol data products are calculated from the 15m interpolated altitudes. All averaging intervals (time and range) are recorded in the HDF5 file (see table UserInputs). The archived subset HDF5 files are the atmospheric files typically sub-sampled at 10 seconds as described in the next section. Atmospheric products are located in the HDF5 file under DataProducts.

Reduced and Subset HDF Files

The 64 channels (not all of the channels are used and depends on configuration) of raw data signal returns are gridded and stored along with the various data products at resolutions described previously. The engineering, navigation, gains, user input and state parameter data are also saved in each file.

The HALO analyzed data files are sub-sampled for distribution due to the large file size associated with the full HALO HDF5 files (~30.5 Gb/hour of flight for raw data files and ~2.5 Gb/hour of flight for the 15 m averaged data files). The subset file contains all of the calculated variables as well as atmospheric state, aircraft data, and 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 HALO 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 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	Temperature	double	[plen nr]	K	Temperature
	Pressure	double	[plen nr]	atm	Pressure
	Number_Density	double	[plen nr]	m^-3	Molecular Number Density
	Relative_Humidity	double	[plen nr]	%	Relative Humidity in percent
	State_Type	uint16	[1 1]	flag	1=sonde, 2=model (GMAO)

Aircraft position information from Nav_Data/GPS (Note that this can vary from one aircraft to another)

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

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	Altitude	double	[plen 1]	m	Altitude vector for data products
	532_ext	double	[plen nr]	km^-1	Retrieved 532nm aerosol extinction coefficient. Values closer than 1500m 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_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. $532\_bsr = \frac{\beta_{aer}}{\beta_{mol}} = \frac{\beta_{aer}^{\parallel} + \beta_{aer}^{\perp}}{\beta_{mol}^{\parallel} + \beta_{mol}^{\perp}}$
	532_bsr_cloud_screened	double	[plen nr]	ratio	Backscatter ratio with cloud mask applied.
	532_bsc	double	[plen nr]	km^-1sr^-1	Aerosol volume backscatter coefficient at 532nm $532\_bsc = \beta_{aer} = \beta_{aer}^{\parallel} + \beta_{aer}^{\perp}$
	532_bsc_cloud_screened	double	[plen nr]	km^-1sr^-1	Aerosol volume backscatter coefficient at 532nm with cloud mask applied.
	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. $532\_total\_attn\_bsc = \frac{\beta_{aer}^{\parallel} + \beta_{aer}^{\perp}}{\beta_{norm}}$
	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_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_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_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.
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_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_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_cloud_screened	double	[plen nr]	km^-1sr^-1	Aerosol volume backscatter coefficient at 1064nm with cloud mask applied
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. $1064\_total\_attn\_bsc = \frac{\beta^{\parallel}_{aer} + \beta^{\perp}_{aer}}{\beta_{norm}}$

	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. $$1064\_dep = \frac{\beta_{mol.}^{\perp} + \beta_{aer.}^{\perp}}{\beta_{mol.}^{\parallel} + \beta_{aer.}^{\parallel}}$$
	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. $$1064\_aerdep = \frac{\beta_{aer.}^{\perp}}{\beta_{aer.}^{\parallel}}$$
	1064_bsc_Sa	double	[plen nr]	km^-1sr^-1	1064 nm aerosol backscatter ratio used for Sa calculation (averaged similar to extinction coefficient)
	WVD_1064_532	double	[plen nr]	ratio	Retrieved wavelength dependence based on the aerosol backscatter coefficient at 1064 and 532nm. $$WVD\_1064\_532 = -\ln\left(\frac{\beta_{aer}^{1064nm}}{\beta_{aer}^{1532nm}}\right) / \ln\left(\frac{1064nm}{532nm}\right)$$
	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_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.
	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.
	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.

	Dust_Mixing_Ratio	double	[plen nr]	sr	Dust mixing ratio, X, calculated from Eq 7 of Sugimoto and Lee, Appl Opt 2006: $X = \left(\frac{(1+\delta_{d532}) \cdot \delta_{532}}{\delta_{d532} \cdot (1+\delta_{532})} \right)$ $\delta_{d532} = 0.35$
	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) $\gamma_d = \ln\left(\frac{A}{B}\right) / \ln(2)$
	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): $\gamma_s = \ln\left(\frac{1-X}{R_\beta - XB/A}\right) / \ln(2)$
	MixedLayerHeight	double	[1 nr]	m, ASL	Mixed layer height determined from 532 backscatter data that has been cloud screened using wavelet method described in Scarino et al., 2014 and using a time dependent threshold value. See MLH calculation section.

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
	range_interp	double	[plen nr]	m	Range of interpolated lidar signal at same grid as data products
	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
	offset_angle	double	[1 nr]	none	Cosine of the angle of the mount plus pitch and roll

Maximum and minimum altitudes of calculated products

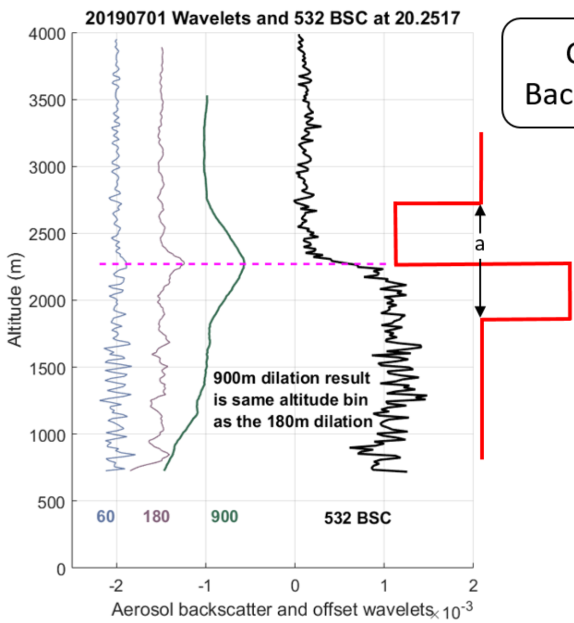
The maximum and minimum altitudes for the calculated variables are listed in Table 1

Table 1

Variable Name	*Minimum Altitude*	*Maximum Altitude*
532_ext	210 m above ground, 300 m above cloud	700 m below aircraft
532_bsr	90 m above ground, 100 m above cloud	500 m below aircraft
532_bsc	90 m above ground, 100 m above cloud	500 m below aircraft
532_total_attn_bsc	90 m above ground, 100 m above cloud	500 m below aircraft
532_dep	90 m above ground, 100 m above cloud	500 m below aircraft
532_aer_dep	90 m above ground, 100 m above cloud	500 m below aircraft
1064_ext	300 m above ground, 240 m above cloud	500 m below aircraft
1064_bsr	90 m above ground, 100 m above cloud	500 m below aircraft
1064_bsc	90 m above ground, 100 m above cloud	500 m below aircraft
1064_total_attn_bsc	90 m above ground, 100 m above cloud	500 m below aircraft
1064_dep	90 m above ground, 100 m above cloud	500 m below aircraft
1064_aer_dep	90 m above ground, 100 m above cloud	500 m below aircraft
Sa_532	210 m above ground, 300 m above cloud	700 m below aircraft
WVD_1064_532	90 m above ground, cloud	500 m below aircraft
AOT_lo	Same as *532_ext*	1.5 km below aircraft
AOT_hi	Extrapolated to ground	1.5 km below aircraft
AOT_hi_col	Extrapolated to ground	1.5 km below aircraft

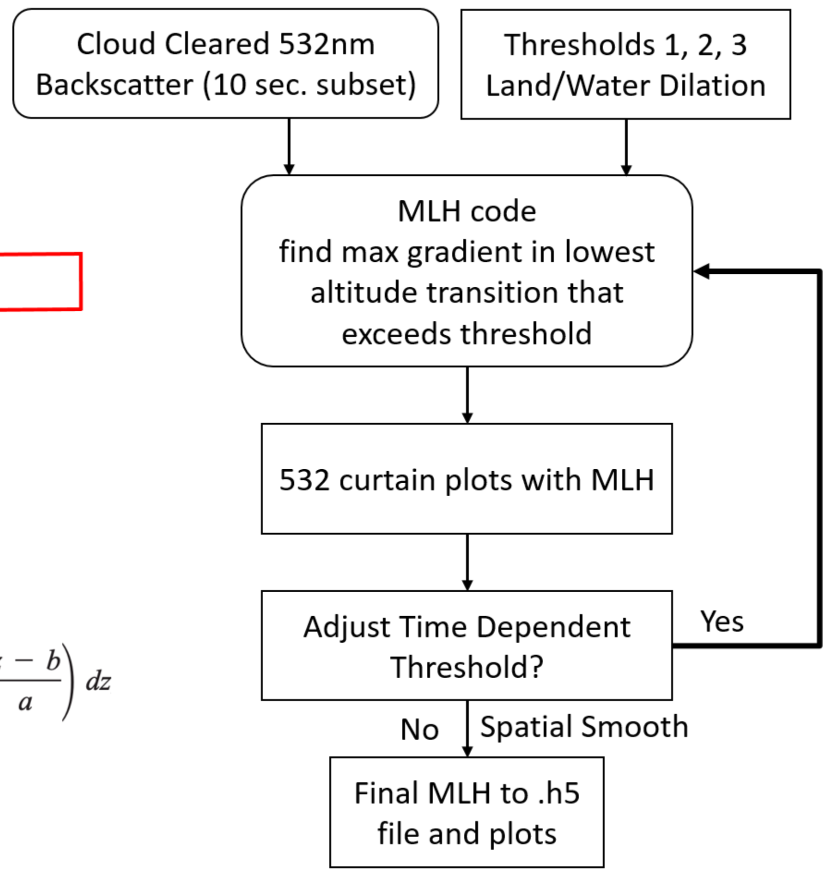
Mixed Layer Height Calculation Procedure and Example

Mixed Layer Height Processing Flow Diagram



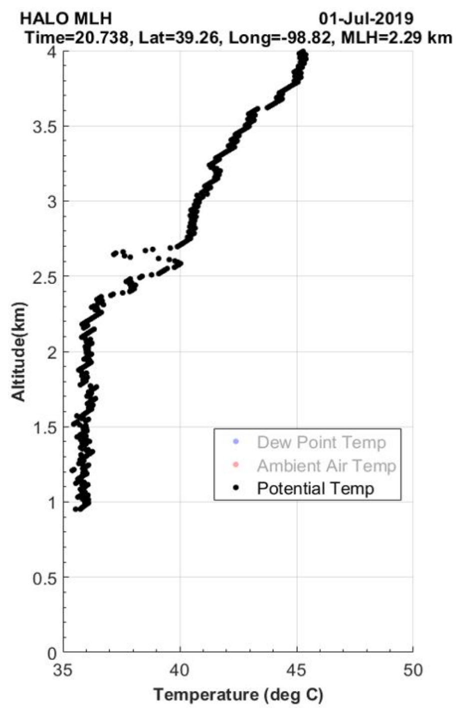
$$W_f(a,\ b) = \frac{1}{a}\int_{z_b}^{z_t} f(z)h\left(\frac{z-b}{a}\right)dz$$

Davis et al., 2000
Brooks 2003
Scarino et al. 2014

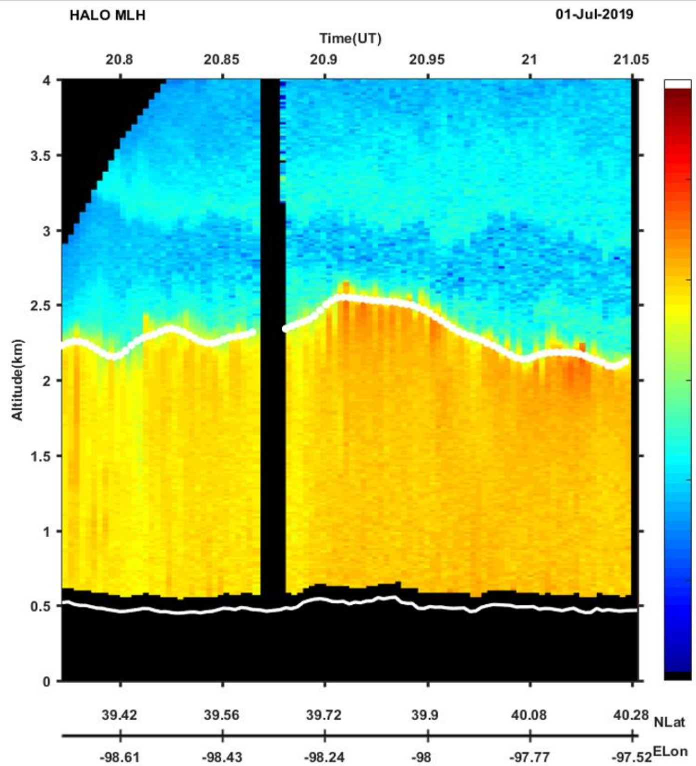


ANALYSIS PROCEDURE

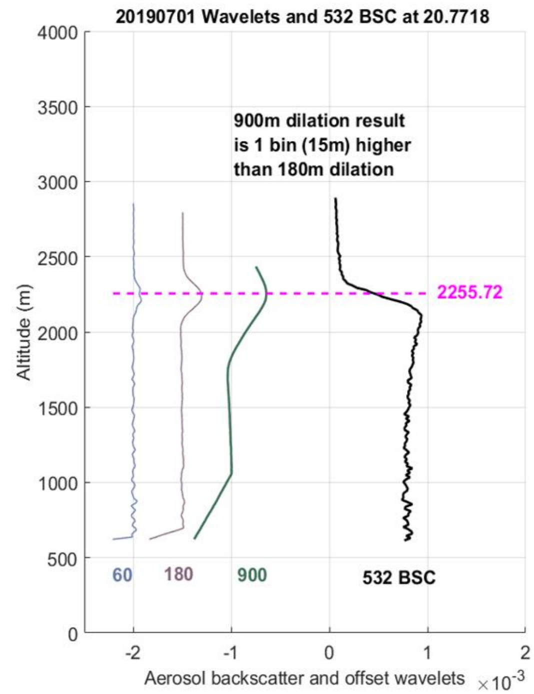
The 10 second smoothed 532 nm cloud and low signal screened backscatter data from the nadir oriented HALO lidar is analyzed using Haar wavelet covariance transform. An example is shown in the figure on the far left, and profiles similar to the left side of figure are generated. For HALO analysis, profiles are generated using the dilation value (a) of 900 m (or 360 m over water) and evaluated to find lowest altitude peak value that exceeds a preset threshold value. The threshold value can be adjusted during the flight to handle varying atmospheric conditions. The final step prior to archiving the data involves running the calculated MLH through a +/- 3 data point gliding or moving window (+/- 30 s).



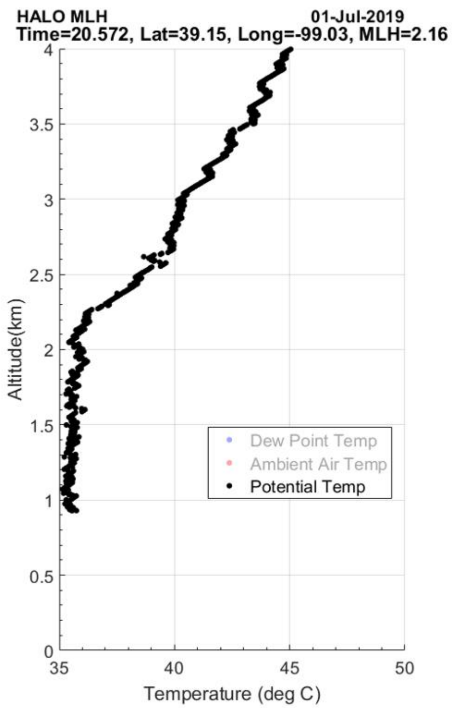
In progress ascent starting just before HALO MLH measurement. In-situ potential temperature suggests MLH ~2.2 – 2.4 km.



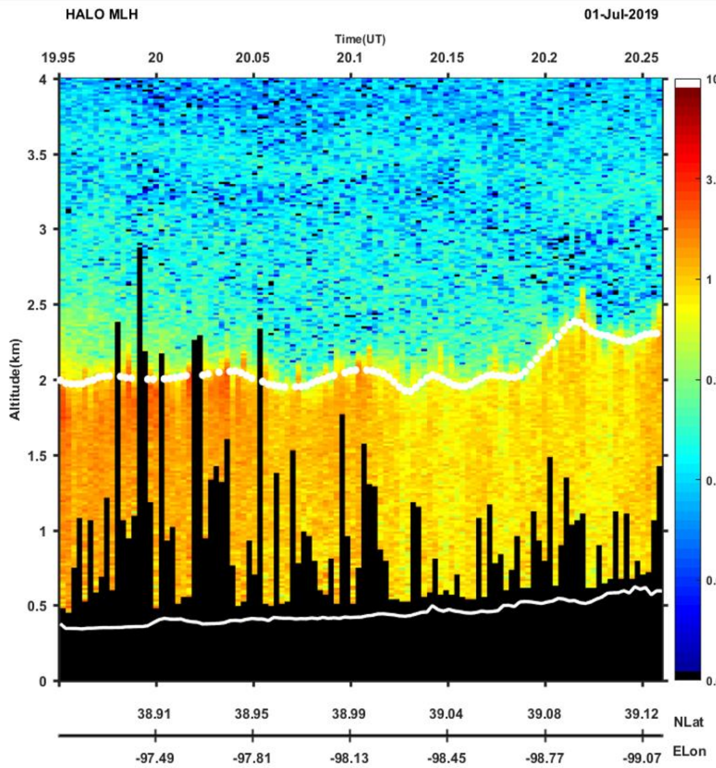
HALO lidar curtain showing 532 backscatter (color bar), smoothed calculated MLH (white dots), ground height (white line at bottom). Left edge is comparison point with in-situ profile in the left hand panel.



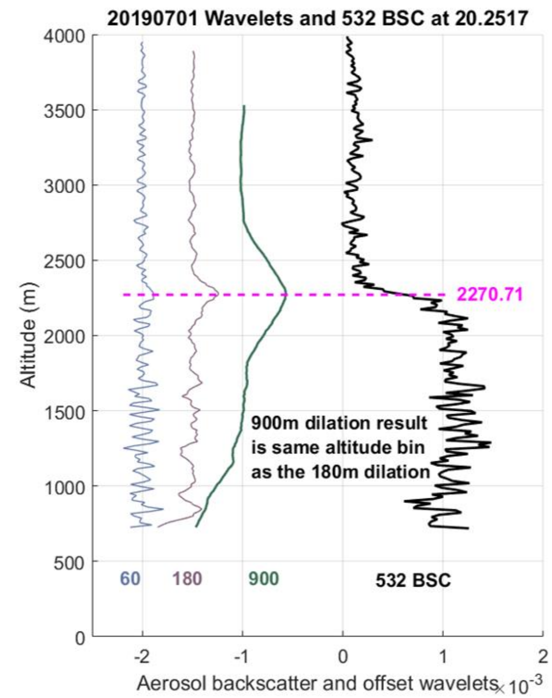
Details of HALO MLH retrieval at 20.77 (close to in-situ location in left hand panel) that shows 60m, 180m and 900m dilation profiles along with 532 BSC. Magenta line shows MLH of 2.26km using 900m dilation. Analysis details per Scarino et al., 2014.



Spiral descent starting just after HALO MLH measurement. In-situ potential temperature suggests MLH ~2.1 – 2.3 km.

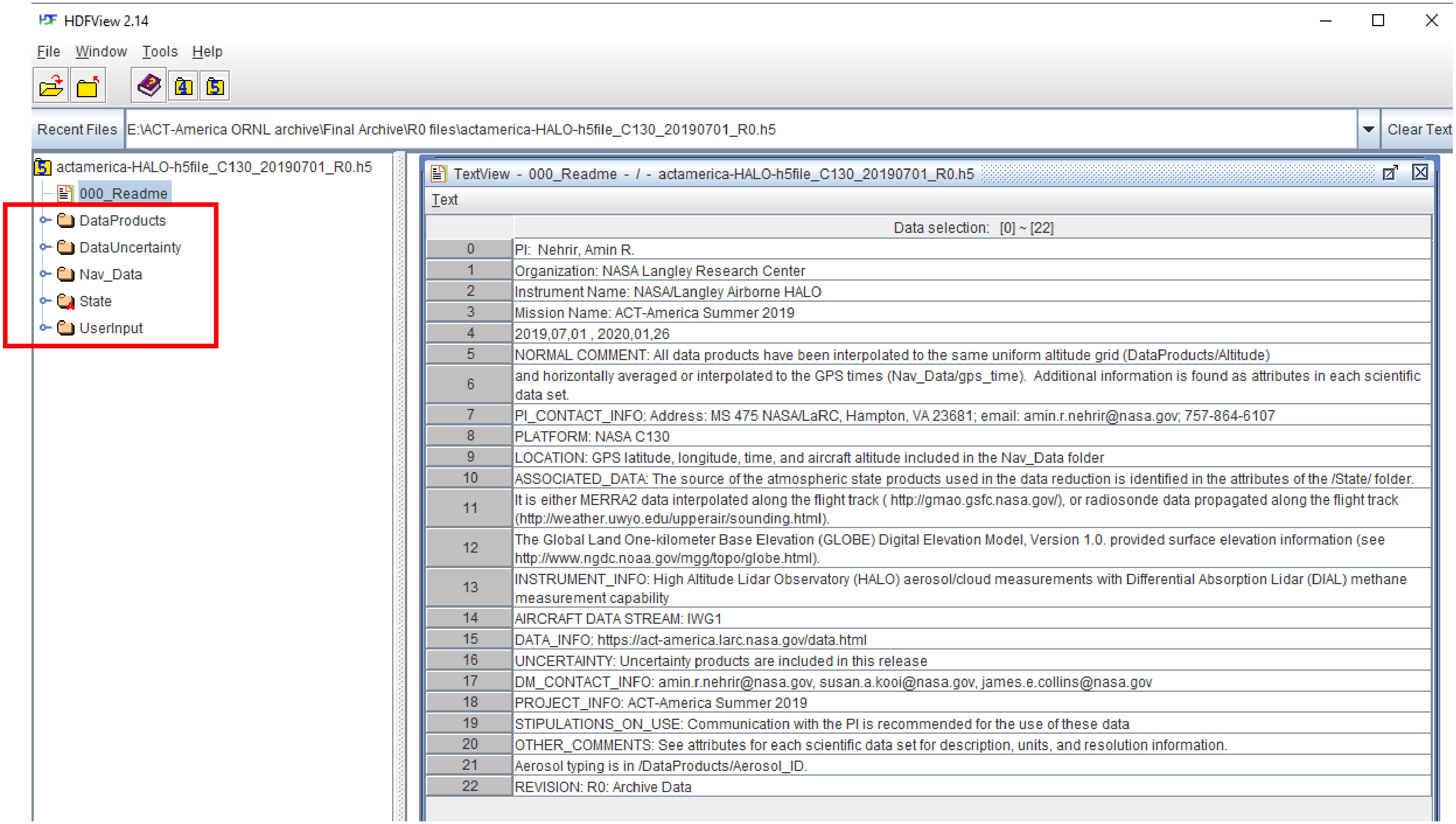


HALO lidar curtain showing 532 backscatter (color bar), calculated MLH (white dots), ground height (white line at bottom). Right edge is comparison point with in-situ profile in the left hand panel.



Details of HALO MLH retrieval at 20.25 (close to in-situ location in left hand panel) that shows 60m, 180m and 900m dilation profiles along with 532 BSC. Magenta line shows MLH of 2.27km using 900m dilation. Analysis details per Scarino et al., 2014.

HDFView image of file folders



HDFView of subset file and folders – 000_Readme section shown