MATLAB procedures for retrieving particle density, retrieving dry complex refractive index, retrieving hygroscopic growth factor, and fitting multimodal size distributions from measured data

The following files are included in this archive: (1) cal\_rho.m, (2) MieKernal.m, (3) S\_e.m, (4) m\_g\_ret.m, and (5) fit\_SD.m. Each procedure includes a description of the purpose of the procedure, a description of the required inputs and outputs of each procedure, and the required MATLAB toolboxes. NOTE: Procedure 1, 2, 3, and 4 require the Mie forward model MATLAB functions (C. Mätzler May, 2002). These procedures were designed for MATLAB 2021a and some require the machine learning, statistics and/or parallel computing MATLAB toolboxes.

The cal\_rho.m code is used to compute the density (i.e., rho) of the particles measured within the overlapping measurement region of the aerodynamic particle sizer (APS) and the laser aerosol spectrometer (LAS). The following single dimension vectors are required as inputs for cal\_rho.m code:

1) no\_las: The lognormal size resolved number concentrations measured by the LAS that overlap the native APS size range.

2) D\_las: The LAS mid-point diameters that overlap the native APS size range.

3) no\_aps: The lognormal size resolved number concentration measured by the APS.

4) D\_aps: The APS mid-point diameters.

The output of cal\_rho.m code is a double value of the retrieved rho.

The MieKernal.m code is used to generate a dry scattering and absorption efficiency lookup table (i.e., Qf). This code requires the set of Mie forward model MATLAB functions (C. Mätzler May, 2002) which was retrieved from https://omlc.org/software/mie/ on January 1st, 2019. Each of the required inputs for the MieKernal.m code are described as follows:

1) D: The mid-point diameters for each size resolved number concentration (i.e., n\_o) channel.

-> dimensions: [number of n\_o channels,1]

-> Units: nm

2) mR: Ascending column vector of desired search range of the real portion of the m\_d,t (i.e., n\_d,t).

-> dimensions: [1,number of mR increments]

-> EXAPMLE: 1.33:0.01:1.70

-> Units: unitless

3) mI: Ascending column vector of desired search range of the imaginary portion of the m\_d,t (i.e., k\_d,t).

-> dimensions: [1,number of mI increments]

-> EXAPMLE: 0:0.01:0.1

-> Units: unitless

4) Lmda\_s: Wavelengths that correspond to the 3 dry scattering coefficient measurement channels.

-> dimensions: [1,3]

-> Units: nm

5) Lmda\_a: Wavelengths that correspond to the 3 dry absorption coefficient measurement channels.

-> dimensions: [1,3]

-> Units: nm

The output of the MieKernal.m code is Qf, which is a multidimensional matrix that is described as follows:

-> dimensions: [number of mR increments, number of mI increments, of n\_o channels,6]

-> Units: unitless

The S\_e.m code is used to generate the covariance matrix (covM) of 6 optical coefficient measurements that are correlated by the correlation factor (i.e., cr). Each of the required inputs for the S\_e.m code are described as follows:

1) y: 6 measured optical coefficients

-> dimensions: [1, 6]

-> Units: # Mm-1

2) Acur: 3 Scattering and 3 absorption measurement accuracy values used

to calculate Chi'.

-> dimensions: [1, 6]

-> FORMAT: [accuracy of scattering channels 1 to 3, accuracy of absorption channels 1 to 3]

-> EXAPMLE: [0.1, 0.1, 0.1, 0.05, 0.05, 0.05]

-> Units: unitless

3) Sig1: 3 Scattering and 3 absorption measurement precision values divided by the square root of the data resolution (i.e., number of native optical measurements averaged to create data set).

-> dimensions: [1, 6]

-> FORMAT: [precision of scattering channels 1 to 3, precision of absorption channels 1 to 3]

-> Units: Mm-1

4) cFct: Optical measurement correlation factor.

-> dimensions: [1, 1]

-> EXAPMLE: 0.3

-> Units: unitless

The output of for the S\_e.m code covM that is described as follows:

-> dimensions: [6, 6]

-> Units: mM-1

The m\_g\_ret.m code is used to retrieve both total dry complex refractive index (i.e., m\_d,t = n\_d,t + k\_d,t) and hygroscopicity (i.e., kappa). Calculated total ambient extinction coefficient at 550 nm (i.e., Cwe) and retrieved total growth factor (i.e., g) at the given relative humidity (RH) are also provided as part of the of the final output (i.e., op). This code requires the set of Mie forward model MATLAB functions (C. Mätzler May, 2002) which was retrieved from https://omlc.org/software/mie/ on January 1st, 2019 and is available with this work. The "S\_e.m" function is used to generate the covariance matrix and "MieKernal.m" must be used to generate the dry scattering and absorption efficiency lookup table (i.e., Qf) for the data set. Each of the required inputs for the m\_g\_ret.m code are described as follows:

1) n\_o: Measured lognormal size resolved number concentration.

-> dimensions: [length of data set, number of measurement channels]

-> Units: # cm-3

2) D: The n\_o measurement channels' mid-point diameters.

-> dimensions: [number of no measurement channels,1]

-> Units: nm

3) mR: Ascending column vector of desired search range of the real portion of the m\_d,t (i.e., n\_d,t).

-> dimensions: [1, number of mR increments]

-> EXAPMLE: 1.33:0.01:1.70

-> Units: unitless

4) mI: Ascending column vector of desired search range of the imaginary portion of the m\_d,t (i.e., k\_d,t).

-> dimensions: [1, number of mI increments]

-> EXAPMLE: 0:0.01:0.1

-> Units: unitless

5) Cs: Dry scattering coefficients measured in 3 wavelengths.

-> dimensions: [length of data set, 3]

-> Units: Mm-1

6) Ca: Dry absorption coefficients measured in 3 wavelengths.

-> dimensions: [length of data set, 3]

-> Units: Mm-1

7) Cws: Humidified scattering coefficients measured in 1 wavelength.

-> dimensions: [length of data set, 1]

-> Units: Mm-1

8) Lmda\_s: Wavelengths that correspond to the 3 dry scattering coefficient measurement channels.

-> dimensions: [1, 3]

-> Units: nm

9) Lmda\_a: Wavelengths that correspond to the 3 dry absorption coefficient measurement channels.

-> dimensions: [1, 3]

-> Units: nm

10) Lmda\_w\_s: Wavelength that corresponds to the humidified scattering coefficient measurement channels.

-> dimensions: [1, 1]

-> Units: nm

11) RH: Measured relative humidity.

-> dimensions: [length of data set,1]

-> Units: nm

12) Qf: Dry scattering and absorption efficiency lookup table for the 3 scattering and 3 absorption wavelengths measured.

-> dimensions: [number of mR increments, number of mI increments, number of no measurement channels,6]

-> Units: unitless

13) Acur: 3 Scattering and 3 absorption measurement accuracy values used to calculate Chi'.

-> dimensions: [1, 6]

-> FORMAT: [accuracy of scattering channels 1 to 3, accuracy of absorption channels 1 to 3]

-> EXAPMLE: [0.1, 0.1, 0.1, 0.05, 0.05, 0.05]

-> Units: unitless

14) Uncr: 3 Scattering and 3 absorption measurement precision values

used to calculate Chi'.

-> dimensions: [1, 6]

-> FORMAT: [precision of scattering channels 1 to 3, precision of absorption channels 1 to 3]

-> Units: Mm-1

15) npt: Resolution of data set (i.e. number of native optical measurements averaged to create data set). Used to calculate Chi'.

-> dimensions: [1, 1]

-> EXAPMLE: (10 second)/(1 second)

-> Units: unitless

16) cFct: Optical measurement correlation factor used to calculate Chi'.

-> dimensions: [1, 1]

-> EXAPMLE: 0.3

-> Units: unitless

The output of the m\_g\_ret.m code is op, which is a MATLAB structure that is described as follows:

1) op.DryRetrievals: Table containing the average of at least 10 valid n\_d,t solutions, average of at least 10 valid n\_d,t solutions, absolute relative error in 3 scattering channels, absolute error in absorption channels, and Chi' threshold corresponding to where 10 or more solutions are valid.

-> dimensions: [length of data set,9]

-> Units: [unitless, unitless, unitless, unitless, unitless, Mm-1, Mm-1, Mm-1, unitless]

2) op.WetRetrievals: Table containing the first valid kappa solution, the first valid g solution, and the Cwe at this solution.

-> dimensions: [length of data set,3]

-> Units: [unitless, unitless, Mm-1]

The fit\_SD.m code iteratively fits a measured total size resolved lognormal aerosol concentration (i.e., n\_o) to a multimodal lognormal model and outputs the fitted modal number concentration (i.e., Nj), geometric standard deviation (i.e., sgj), and geometric mean diameter (i.e., Dgj). where. This code was created using the outline and guidelines provided by Hussein et al. (2005). Each of the required inputs for the fit\_SD.m code are described as follows:

1) n\_o: Measured lognormal size resolved number concentration.

-> dimensions: [1, number of measurement channels]

-> Units: # cm-3

2) N1: The number concentration of particles with diameters between 7 and 20 nm.

-> dimensions: [1,1]

-> Units: # cm-3

3) N2: The number concentration of particles with diameters between 35 and 80 nm.

-> dimensions: [1,1]

-> Units: # cm-3

4) N3: The number concentration of particles with diameters between 90 and 500 nm.

-> dimensions: [1,1]

-> Units: # cm-3

5) Nc: The number concentration of particles with diameters greater than 900 nm.

-> dimensions: [1,1]

-> Units: # cm-3

6) D: The n\_o measurement channels' mid-point diameters.

-> dimensions: [number of no measurement channels,1]

-> Units: nm

The output of the fit\_SD.m code is op, which is a MATLAB table containing 12 columns for each of the potential modal parameters (e.g., Nj, sgj, and Dgj) and the corresponding goodness of fit parameters (i.e., R and p-value). Where j = 1:4 for the nucleation, Aitken, Accumulation, and coarse modes. NOTE: This algorithm will determine the optimal number of modes and some modes may be removed during the fitting, which will result in NAN values being place for the fit parameters of the mode that is removed.

-> dimensions: [1,12]

-> Units: [# cm-3, unitless, nm, # cm-3, unitless, nm, # cm-3, unitless, nm, # cm-3, unitless, nm, unitless, unitless]