*README: EAGLES Liquid Cloud Testbed Large Eddy Simulation Library (v2)*

This dataset contains data used for “Evaluation of Autoconversion Representation in E3SM.v2 using an Ensemble of Large-Eddy Simulations of Low-Level Warm Clouds” by M. Ovchinnikov, P.-L. Ma, C. M. Kaul, K. G. Pressel, M. Huang, J. Shpund, and S. Tang

The data was generated using the PINACLES large eddy simulation (LES) model that is described in Pressel and Sakaguchi (2021; doi: 10.2172/1869291) and Dhandapani et al. (2023; doi: 10.22541/essoar.170365352.28240067/v1). The spectral bin microphysics scheme used for this work is described in Shpund et al. (2019; doi: 10.1029/2019JD030576).

**Case studies**

The LES cases consist of the following “base” cases:

* ATEX : a trade-wind cumulus case with a relatively strong inversion, developed as an LES intercomparison case in Stevens et al. (2001; doi: 10.1175/1520-0469(2001)058<1870:SOTWCU>2.0.CO;2)
* DYCOMS [RF02]: marine stratocumulus case developed as an LES intercomparison case in Ackerman et al. (2009; doi: 10.1175/2008MWR2582.1)
* ENA 18 July: based on observed conditions at the ARM Eastern North Atlantic site in the Azores
* ENA 19 Jan: based on observed conditions at the ARM Eastern North Atlantic site in the Azores
* ENA 25 Jan: based on observed conditions at the ARM Eastern North Atlantic site in the Azores
* SGP 30 Aug a: based on observed conditions at the ARM Southern Great Plains site and using LASSO (Gustafson et al., 2020; doi: 10.1175/BAMS-D-19-0065.1) forcing
* SGP 30 Aug b: as for SGP 30 Aug a, but aerosol distribution estimated from the “b” set of flight legs

Along with each of the base cases, two additional simulations are performed using the same meteorological initial state, large scale forcing, and surface fluxes but doubling (“2x”) and halving (“0.5x”) the number concentration of aerosol.

The case studies are listed in the spreadsheet “PINACLES\_Autoconversion\_Simulation\_Table.xlsx”. Input files for the simulations are packaged with the simulation output data bundles. The JSON input files for each simulation can be consulted to obtain details of their setups. The typical domain size is 25.6 km x 25.6 km extents in horizontal directions and 2.5 km to 6 km vertical extent. All ENA and DYCOMS cases use 50 m horizontal/10 m vertical grid spacings, and all SGP cases use 40 m horizontal/40 m vertical grid spacing. Time durations of the simulations range from 6 to 12 hours. Additional input files are also provided that contain initial conditions and forcing for meteorological fields and initial conditions of the aerosol size distributions for each simulation.

**Data files**

Two types of files are included in this data archive for each case study:

* ***Fields*** files are instantaneous snapshots of 3-d volumes of data including prognostic and diagnostic variables. Within each case study directory, there is a “fields” subdirectory that contains hourly output (except for the ena\_19jan\_0.5x case, which has some missing outputs). These files are in HDF5 format and are named by the time since simulation start: *xx*d-*xx*h-*x*xs.h5 (for ATEX and DYCOMS simulations) or *xx*d-*xx*h-*x*xs-*xx*ms.h5 (for all other simulations). Variables used from these files for the manuscript’s analyses include:
	+ qc\_autoconv: cloud droplet mixing ratio immediately before evaluation of collision-coalescence in SBM [kg kg-1]
	+ nc\_autoconv : cloud droplet number concentration immediately before evaluation of collision-coalescence in SBM [# kg-1]
	+ qr\_autoconv: rain mixing ratio immediately before evaluation of collision-coalescence in SBM [kg kg-1]
	+ nr\_autoconv: rain drop number concentration immediately before evaluation of collision-coalescence in SBM [# kg-1]
	+ auto\_cldmsink\_b: sink of cloud droplet mass due to autoconversion [kg kg-1 s-1]
	+ auto\_cldnsink\_b: sink of cloud droplet number due to autoconversion [# kg-1 s-1]
* ***Stats*** files provide horizontal domain average profiles, minimum/maximum profiles, timeseries of quantities such as liquid water path, surface heat fluxes, etc. These files are in NetCDF4 format with outputs every 60 s. A single stats.nc file is provided for each case study. Variables used from these files for the manuscript’s analyses include:
	+ qc\_autoconv: mean profile of cloud droplet mixing ratio immediately before evaluation of collision-coalescence in SBM [kg kg-1]
	+ qc\_autoconv\_squared: mean profile of the square of qc\_autoconv
	+ qc\_autoconv\_cloud: in-cloud mean profile of cloud droplet mixing ratio immediately before evaluation of collision-coalescence in SBM [kg kg-1] (in-cloud conditions are defined by grid cell qc > 1e-20 kg kg-1)
	+ nc\_autoconv\_cloud :in-cloud mean profile of cloud droplet number concentration immediately before evaluation of collision-coalescence in SBM [# kg-1]
	+ qr\_autoconv\_cloud: in-cloud mean profile of rain mixing ratio immediately before evaluation of collision-coalescence in SBM [kg kg-1]
	+ nr\_autoconv\_cloud: in-cloud mean profile of rain drop number concentration immediately before evaluation of collision-coalescence in SBM [# kg-1]
	+ auto\_cldmsink\_b\_cloud:in-cloud mean profile of the sink of cloud droplet mass due to autoconversion [kg kg-1 s-1]
	+ auto\_cldnsink\_b: in-cloud mean profile of the sink of cloud droplet number due to autoconversion [# kg-1 s-1]
	+ rho0: reference state density

NB1: the horizontal velocity components *u* and *v* are decremented by a Galilean transformation velocity provided in the file “transformation\_velocities.txt” for each case study. To recover *u* and *v* in a fixed reference frame, add u0/v0 to u/v output variables in the fields and stats files.

NB2: Data is archived beginning at time 0. However, a minimum of two hours should be allowed as model spin-up time and data users are advised to assess if a longer spinup time should be allowed for the specific analyses planned.

NB3: Later in the simulations, cloud tops in the SGP cases sometimes protrude into the damping layer that extends over the top 500 m of the computational domain. Data from times after this protrusion should be used with caution, if at all, depending on the analysis to be performed.

Data point of contact: Colleen Kaul (colleen.kaul@pnnl.gov)