Supporting Information for "Emulation of cloud microphysics in a climate model"

W. Andre Perkins[1], Noah D. Brenowitz[2],

Christopher S. Bretherton[1], Jacqueline M. Nugent[3]

[1]Allen Institute for Artificial Intelligence

[2]NVIDIA

[3]University of Washinton

Contents of this file

1. Figures S1 to S5
2. Tables S1 to S2

Corresponding author: W. A. Perkins, Climate Modeling, Allen Insitute for Artificial Intelligience, Seattle, WA, USA. (andrep@allenai.org)

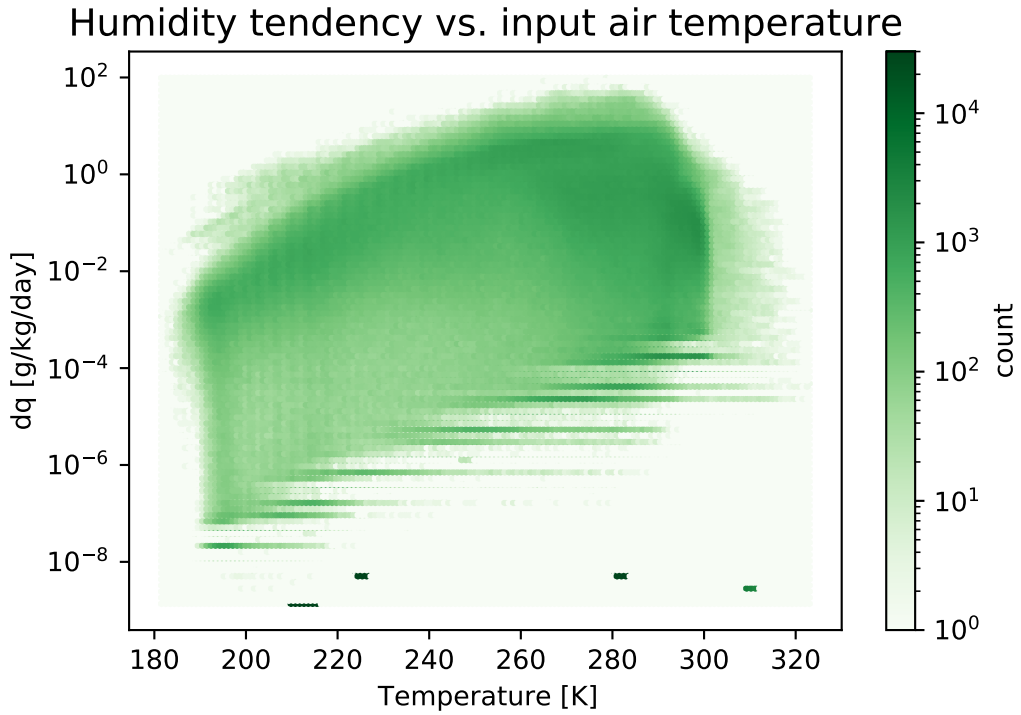

Figure S1. A 2D histogram comparing the input air temperature against the gscond $\Delta_g q$ tendency using 150,000 random columns from the training dataset.

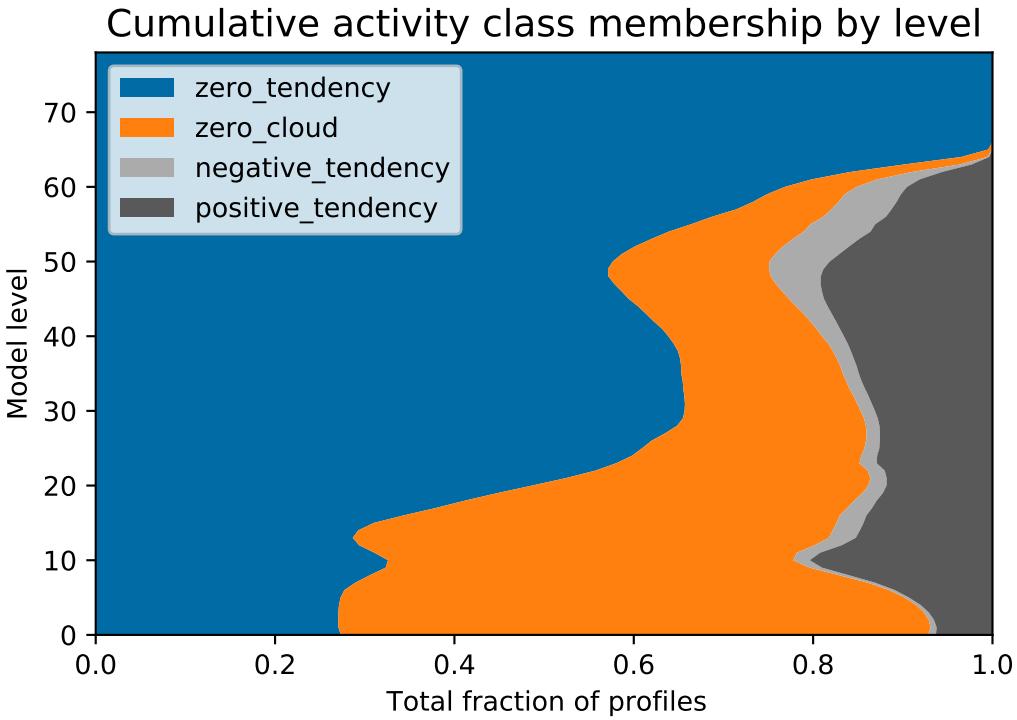

Figure S2. Class membership fraction by level for the gscond activity classifier. Class fractions are calculated over 150,000 random columns from the training dataset.

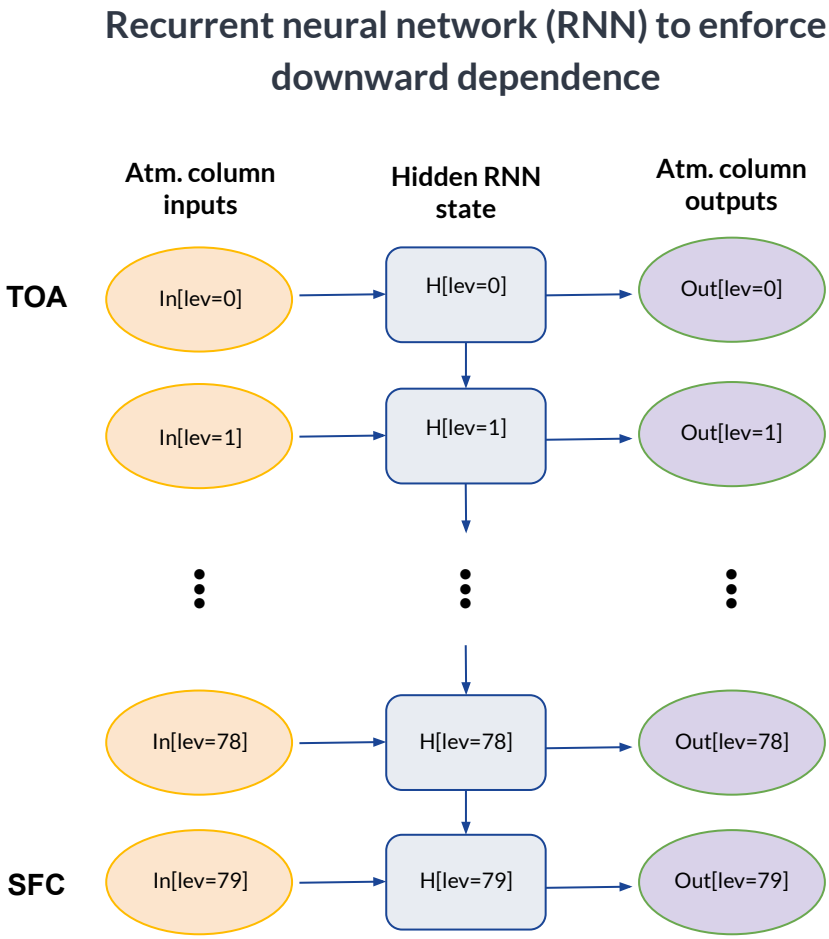

Figure S3. A schematic of information flow through the RNN network, which enforces a downward dependence in the output state starting at the top of atmosphere (TOA). All inputs for a given level level are fed into the two-layer hidden state as the model recurses downward in the atmosphere. For each level (recursive step), the model translates to outputs via a linear readout layer.

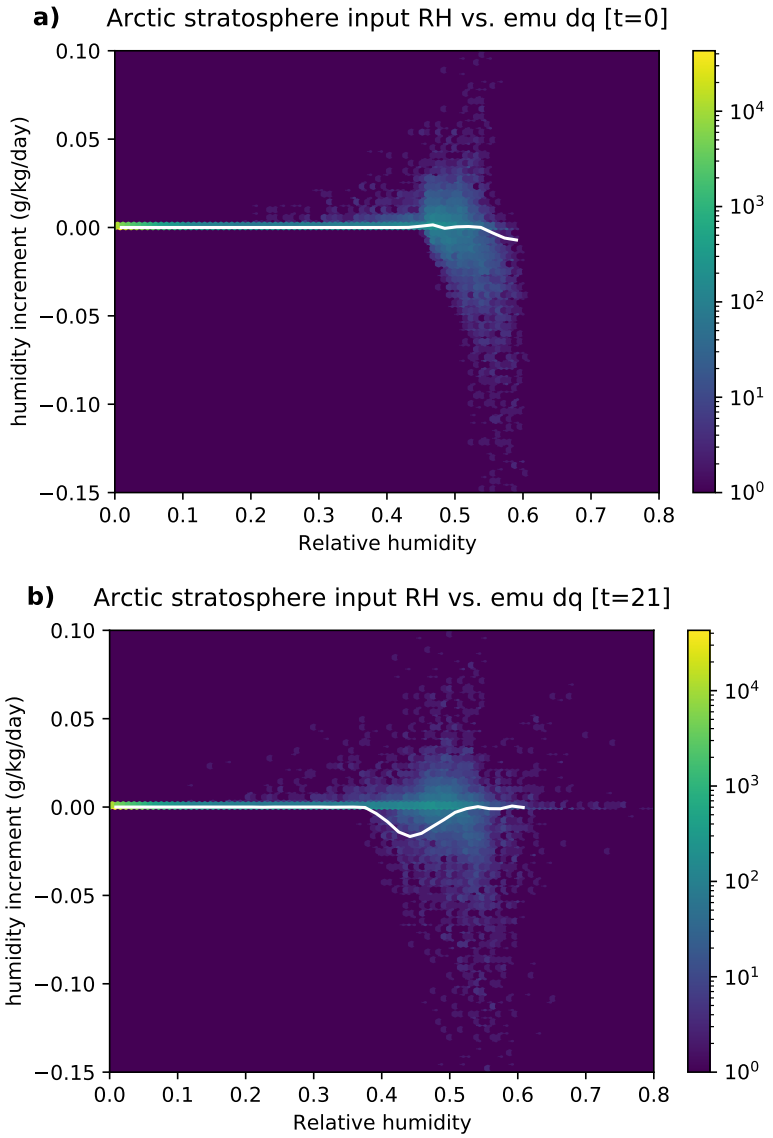

Figure S4. A 2D histogram of the gscond emulator humidity increment vs. input relative humidity at (a) t=0 (first model timestep) and (b) t=21 (5.25 hours later). The humidity increment and relative humidity values are gathered from the selected timestep of all 12 monthly initializations. The solid white line depicts the average bias of the emulator increment compared against diagnostic Fortran increment binned by relative humidity for bins with >0.05% of the total samples.

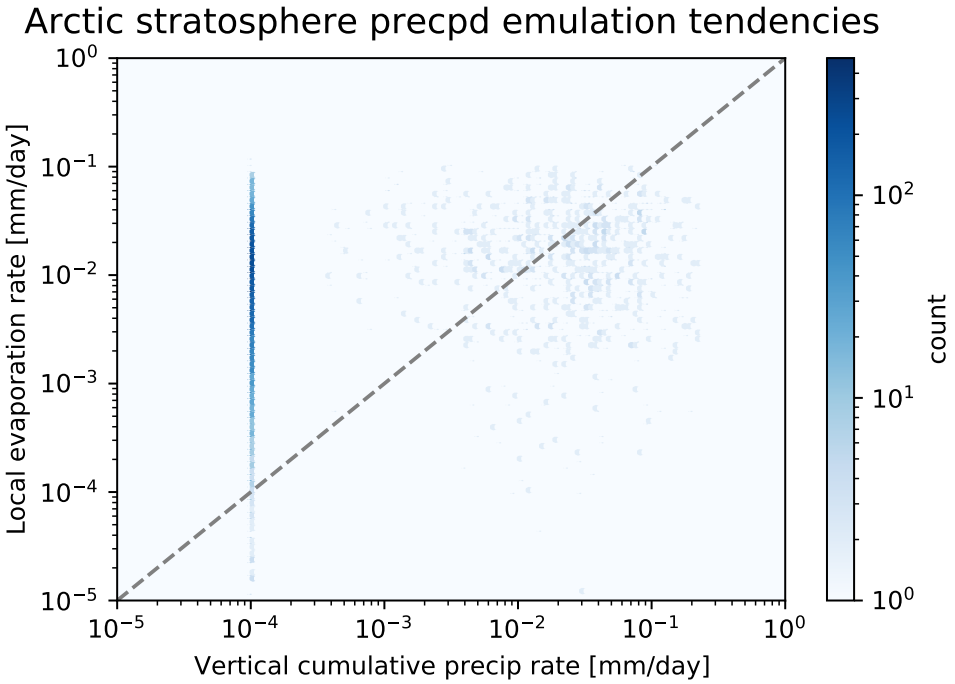

Figure S5. A 2D histogram comparing the vertically cumulative precipitation flux (at and above a given gridpoint) with the local evaporation in the Arctic stratosphere for the offline precpd emulator predictions. The grey dashed line depicts the 1-to-1 ratio where all precipitation falling through a given level is evaporated. Any points above and to the left of this line signify non-conservation, where more liquid is evaporated than available from the precipitation. A value of 1e-4 is added to precipitation in order to visualize all values (including zero precipitation) on the log-log scale.

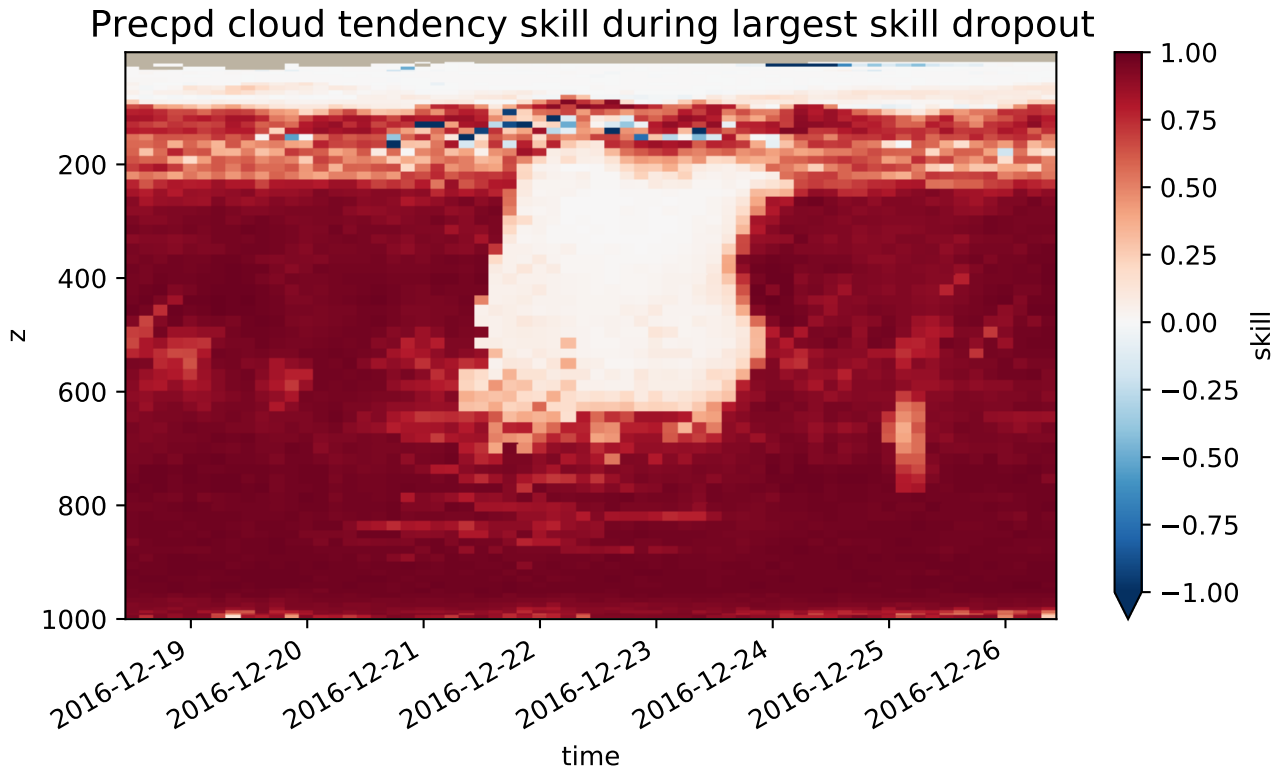

Figure S6. A time–height plot of precpd emulator skill focusing on the days surrounding a large skill dropout occurring on Dec. 22nd of the gscond top-5 masking 1-year simulation (see Figure B1).

Field	Description
delp	pressure thickness of the atmospheric layer
air_pressure	pressure at center of atmospheric layer
surface_air_pressure	air pressure at the surface (lowest model interface)
air_temperature_input	air temperature input into the ZC scheme
specific_humidity_input	specific humidity input into the ZC scheme
cloud_water_mixing_ratio_input	cloud water mixing ratio input to the ZC scheme
specific_humidity_after_gscond	specific humidity after the current timestep call to the gscond subroutine
air_temperature_after_gscond	air temperature after the current timestep call to the gscond subroutine
cloud_water_mixing_ratio_after_gscond	cloud water mixing ratio after the current timestep call to the gscond subroutine
air_temperature_after_precpd	air temperatureafter the current timestep call to the precpd subroutine
specific_humidity_after_precpd	specific humidity after the current timestep call to the precpd subroutine
cloud_water_mixing_ratio_after_precpd	cloud water mixing ratio after the current timestep call to the precpd subroutine
total_precipitation	surface precipitation rate after the current timestep call to the precpd subroutine
air_temperature_after_last_gscond	air temperature after the previous timestep call to gscond
specific_humidity_after_last_gscond	specific humidity after the previous timestep call to gscond
surface_air_pressure_after_last_gscond	surface air pressure after the previous timestep call to gscond

Table S1. A list of ZC microphysics fields pushed from the Fortran state to the call_py_fort

Python environment to save for training and to use for inference at runtime.

	observed %	predicted %	precision	recall	accuracy
positive_tendency	0.11	0.11	0.95	0.96	0.99
zero_tendency	0.63	0.65	0.95	0.99	0.96
zero_cloud	0.24	0.21	0.98	0.86	0.96
negative_tendency	0.02	0.02	0.90	0.90	0.99

Table S2. Metrics for the gscond activity classifier calculated on 150,000 random sample columns from the test dataset.