

1 **Regimes of cloud vertical structure from active observations**

2 **Dongmin Lee,^{1,2} Lazaros Oreopoulos², Nayeong Cho,^{2,3}**

3 ¹ Morgan State University, Baltimore, Maryland

4 ² Earth Sciences Division, NASA's Goddard Space Flight Center, Greenbelt, Maryland

5 ³ University of Maryland Baltimore County, Baltimore, Maryland

6 Corresponding author: Lazaros Oreopoulos (Lazaros.Oreopoulos@nasa.gov)

7 **Key Points:**

- 8 • Mixtures of cloud vertical structures from active observations and their large-scale model
9 counterparts are amenable to classification.
- 10 • The classification enables the creation of a gridded dataset from data that were originally
11 sparse in active observations.
- 12 • This dataset can be used to understand contributions to Earth energetics by various cloud
13 systems in both observations and models.

14 **Abstract**

15 We introduce a new type of cloud class, which we call “active cloud regime” (ACR), owing to
16 its provenance from active (lidar and cloud radar) spaceborne cloud observations. ACRs intend
17 to provide a climatological description based on cloud vertical structure (CVS) of the most
18 prevalent monthly CVS mixtures encountered at large spatial scales of ~400 km. ACRs are thus
19 a way to create a gridded dataset of a vertically-resolved cloud mask that can facilitate joint
20 analysis with other gridded datasets. The detailed 2D cloud mask comes from the 2B-
21 CLDCLASS-LIDAR CloudSat dataset fusing CALIPSO (lidar) and CloudSat (cloud radar) cloud
22 detections. We show that the global classification of cloudiness under the ACR framework
23 provides valuable insights on how the world’s dominant cloud systems regulate the two major
24 components of atmospheric energetics, precipitation and radiative cooling. NASA’s GEOS
25 model allows us to demonstrate the feasibility of applying the ACR concept in Earth System
26 Models that have the capability to produce subgrid cloudiness obeying pre-specified vertical
27 overlap rules. Comparison of observed and simulated ACRs provides thus another means to
28 assess the realism of modeled clouds.

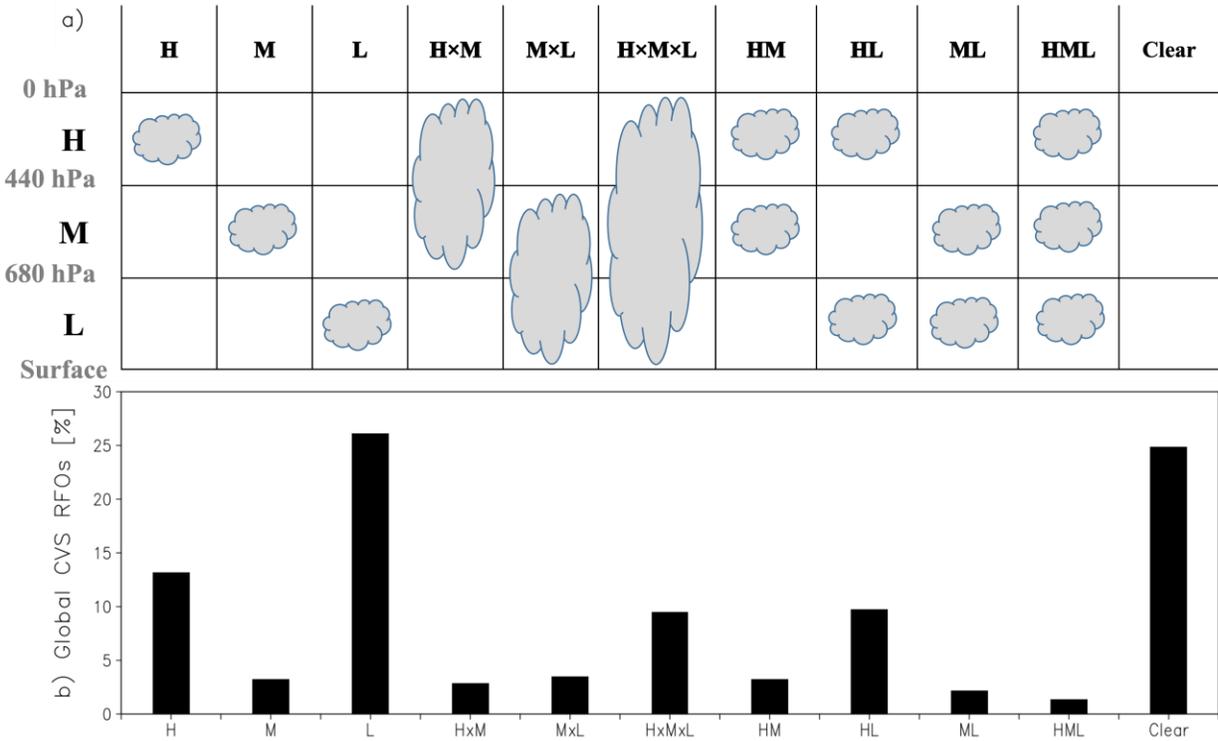
29 **Plain Language Summary**

30 So-called “active” space-based observations such as those from backscatter lidars and cloud
31 radars offer unique views of the vertical structure of clouds that enhance our understanding about
32 many aspects of the planet’s cloudiness and can be used to thoroughly assess clouds simulated
33 by climate models. In this work we show that one way to compensate for the inherent sparseness
34 of such observations is to group alike sparse products using standard classification schemes. The
35 specific implementation of our study classifies previously derived instantaneous ~ 2 km cloud
36 vertical structures into classes we call “active cloud regimes” defined at ~400 km domains and

37 resolved at monthly time scales. We show that this new cloud product enables easier probing of
38 atmospheric energetics and evaluation of cloudiness simulated by large-scale models.

39 **1. Introduction**

40 Cloud views from active sensors such as the Cloud-Aerosol Lidar with Orthogonal
41 Polarization (CALIOP) lidar (Winker et al. 2009) of the Cloud-Aerosol Lidar and Infrared
42 Pathfinder Satellite Observation (CALIPSO) satellite and CloudSat’s Cloud and Precipitation
43 radar (CPR) (Stephens et al. 2002) provide information about certain aspects of clouds that is
44 otherwise difficult to obtain from passive shortwave and thermal infrared imagers such as the
45 Moderate Resolution Imaging Spectroradiometer (MODIS) and the Visible Infrared Imaging
46 Radiometer Suite (VIIRS). Specifically, under a range of conditions, active sensors can more
47 reliably detect cloud thermodynamic phase, the presence of multiple cloud layers, and cloud
48 vertical structure, including the location of cloud top and base and the geometrical thickness of
49 distinct cloud layers. In previous work we have taken advantage of this capability to generate
50 distinct classes of cloud mask profiles we called Cloud Vertical Structures (CVS, Oreopoulos et
51 al. 2017), as in Fig. 1. We also showed that Earth System Models equipped with a subcolumn
52 cloud generator that incorporates vertical overlap rules can in principle also produce these major
53 vertical cloud configurations inferred from individual observational “rays” of active sensors (Lee
54 et al. 2020), and can thus have their cloudiness assessed with regard to the realism of their own
55 cloud vertical configurations.



56

57 **Figure 1.** (a) The 10 CVS classes of Oreopoulos et al. (2017) representing co-occurrence
 58 combinations of high, middle, and low clouds (as defined by the International Satellite Cloud
 59 Climatology Project ISCCP) according to the CloudSat 2B-CLDCLASS-LIDAR (2BCL)
 60 product; (b) CVS global relative frequency of occurrence (RFO) from 2007-2010 observations,
 61 shown in histogram form.

62 Our previous work defined observed CVS at the fundamental spatial scale of the
 63 associated CloudSat-CALIPSO cloud product, 2B-CLDCLASS-LIDAR (henceforth “2BCL”,
 64 Sassen et al. 2008), i.e., ~ 2 km. Conveniently, radiative fluxes at the boundaries of the
 65 atmospheric column, but also profiles of such fluxes throughout the column, associated with
 66 each CVS kind can be obtained from CloudSat’s 2B-FLHXR-LIDAR (2BFL) product
 67 (Henderson et al. 2013). With appropriate averaging, this enables assessments of the overall
 68 long-term radiative impact of each CVS at either global or zonal scales (Lee et al. 2020).
 69 L’Ecuyer et al. (2019) performed a similar climatological radiative assessment, but based on

70 2BCL’s “cloud type” classification. In these works, the gridded radiation composites obtained
71 from radiative fluxes at ~ 2km horizontal resolution are specific to only the particular cloud
72 class, be it CVS (Lee et al. 2020) or cloud type (L’Ecuyer et al. 2019); individual grid-averaged
73 radiative fluxes cannot be associated with grid-averaged cloudiness since the latter is ill-defined.
74 Effectively, one has to choose between either gridded radiative fluxes corresponding to
75 unidentified mixtures of cloud classes, or construct a separate gridded radiation dataset for each
76 cloud class. In either approach, the relationship between regional radiation and regional
77 cloudiness cannot be resolved.

78 Given this state of affairs, this paper introduces the concept of a cloud class conveying
79 vertical morphology that can serve as a gridded entity. We call this concept Active Cloud
80 Regime (ACR). Such a gridded (“Level 3”) cloud dataset, while by construction is much more
81 coarsely resolved than its spatially sparse Level 2 progenitor, can be more straightforwardly
82 combined with existing or tailored user-generated gridded products for studies aimed at gaining
83 insight into the climatic role of distinct cloud vertical morphologies. An example of such an
84 application in this work is the breakdown by ACR of the components of atmospheric diabatic
85 heating which is accomplished by compositing independent monthly gridded datasets as a
86 function of the mixtures of CVS kinds represented by ACRs.

87 In the remainder of the paper, we document the derivation of ACRs from observations
88 and the cloud fields of a global model, and then show how they can be used to decompose and
89 interpret various components of the Earth’s energy budget. Performing such a decomposition on
90 both observed and modeled ACRs, enables us to assess the realism of the model’s cloudiness and
91 distribution of energetics.

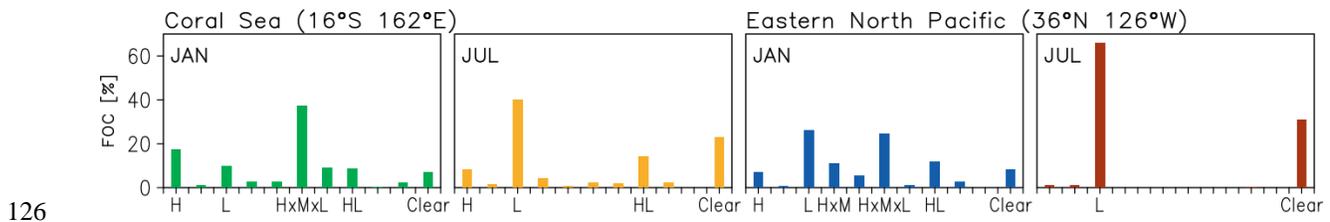
92 **2. Dataset and derivation of Active Cloud Regimes**

93 The fundamental concept applied to create a gridded cloud vertical structure dataset is
94 classifying how various CVS (Fig. 1, see also Oreopoulos et al. 2017) mix over a chosen
95 spatiotemporal scale. The relatively simple CVS classification of Fig. 1 introduced by
96 Oreopoulos et al. (2017) adopted the International Satellite Cloud Climatology Project (ISCCP)
97 convention of distinguishing clouds based on the altitude of their radiative top into high (*H*),
98 middle (*M*), and low (*L*) categories using 680 and 440 hPa pressure levels as the delineating
99 boundaries (Rossow and Schiffer 1999). The graphical depiction of Fig. 1 hopefully makes self-
100 explanatory the nature of the different CVS kinds. In the CVS framework, the clouds
101 encountered in each ~ 2 km profile of active observations are assigned to either one, two or all
102 three of the ISCCP broad layers in which the troposphere is segmented; when they appear
103 simultaneously in more than one, clouds can occupy adjacent layers either a contiguously
104 (indicated by \times) or non-contiguously.

105 When CVS mix over a give spatiotemporal scale, the mixtures are expressed in terms of CVS
106 frequency of occurrence (FOC) histograms. To construct these histograms we updated the CVS
107 dataset previously used in Oreopoulos et al. (2017) with CVS from the newer 2BCL Release 5
108 (R05) dataset. As in that paper, we used four years of data from the 2007–2010 period belonging
109 to the “golden” epoch of CloudSat–CALIPSO observations (i.e., before transition to daytime-
110 only operations). We have chosen, for reasons explained below, 4° and one month as the
111 underlying spatial and temporal scales on which CVS occurrences are accumulated into FOC
112 histograms. While arbitrary, these scales were deemed suitable for the purposes of our
113 classification as discussed below.

114 With the CVS mixtures expressed as histograms of CVS FOC at the aforementioned scales, their
115 subsequent classification was accomplished by clustering via a *k*-means algorithm. The concept
116 is fundamentally the same as the one used to classify into cloud regimes the passive imager

117 observations of cloud occurrences in specific cloud top pressure (CTP) and cloud optical
 118 thickness (COT) combinations (“cloud types”) (e.g., Cho et al. 2021). In essence, instead of
 119 classifying a 42-element (combinations of 6 COT and 7 CTP bins) MODIS array of near-
 120 instantaneous cloud type occurrence in a 1° grid, here we classify an 11-element CVS FOC
 121 histogram (array) assembled from CVS occurrences over a month in a 4° domain. Examples of
 122 such histograms for two regions and two months are shown in Fig. 2. Note the large seasonal
 123 variability of CVS mixtures in both example regions, the Coral Sea and NE Pacific. The
 124 objective of the *k*-means clustering is to group FOC histograms like those in the four examples of
 125 Fig. 2 with alike histograms.

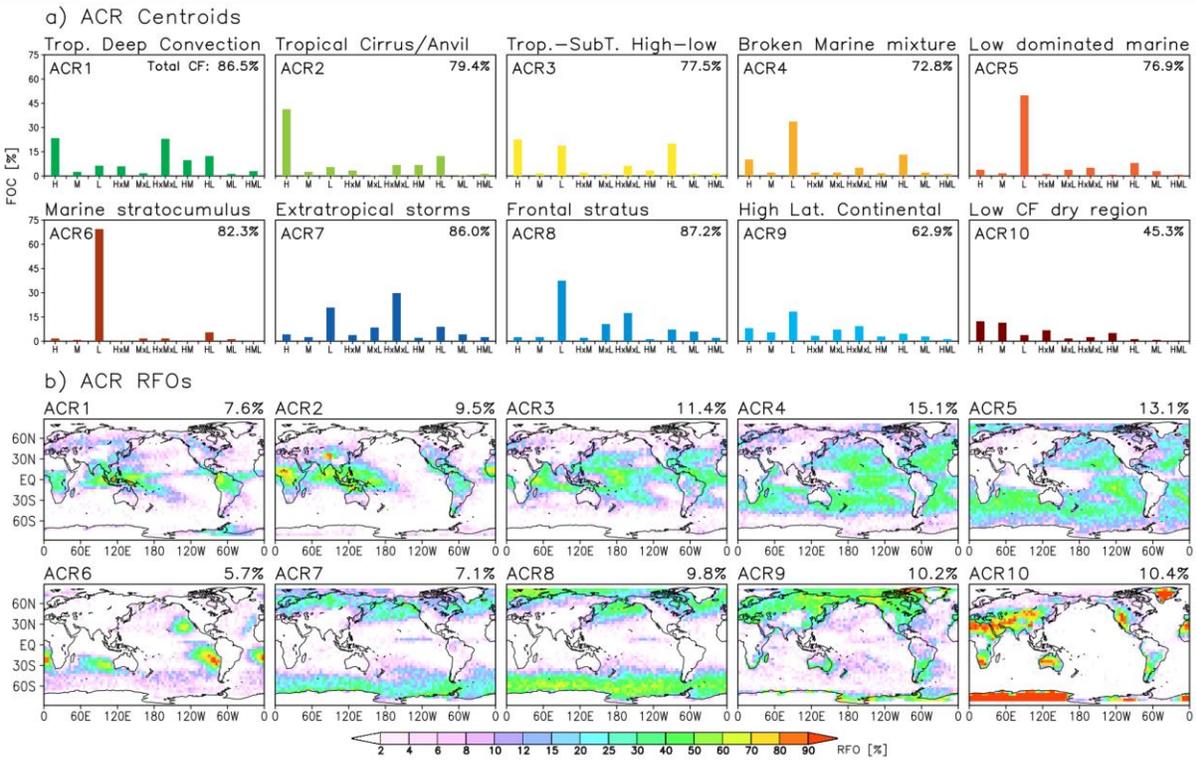


126
 127 **Figure 2.** Four CVS FOC histogram examples from two select 4° regions and two different
 128 months (January and July). Large seasonal variations in CVS populations can be seen.

129 The 4° monthly spatiotemporal scale typically contains a large enough number of CVS samples
 130 ($\sim 10^3$) to populate the histograms subjected to clustering. FOC histograms from even larger CVS
 131 populations could be produced by expanding the spatiotemporal scales, but this would bring
 132 drawbacks: first, cloud mixing would start becoming excessive to the detriment of the
 133 distinctiveness of the CVS mixture; second, the number of histograms furnished to the clustering
 134 algorithm would progressively become smaller, endangering algorithm convergence. With our
 135 current choice of spatiotemporal scales (four degrees monthly), geographical coverage (longitude
 136 from -180° to 180° and latitude from -82 to 82°) and length of dataset (48 months), 177120
 137 $(90 \times 41 \times 48)$ gridcells and corresponding histograms are available for the clustering process. The

138 bottom line is that with a fixed number of CVS samples, a judicial allocation needs to be made
139 between CVS populations used to build FOC histograms and numbers of such histograms. Our
140 trial and error with different spatial and temporal scales led to the conclusion that 4° and one
141 month is indeed the Goldilocks choice of both robust, well-sampled, CVS FOC histograms, and
142 also sufficient number of histograms for k -means clustering.

143 While objective methods of selecting the optimal number of clusters exist, they do not always
144 yield physical results, so it is not uncommon common in k -means clustering to resort to some
145 trial and error in order to settle on an appropriate number of clusters. In our case, after
146 experimenting with different numbers of clusters K , we settled on $K=10$ as the best number of
147 clusters to represent the ACRs. We found that fewer clusters muddled the geographical
148 occurrence patterns and yielded less distinct peaks in the histogram centroids indicating too
149 much mixing of dissimilar CVSs in the population. The additional separation provided by a
150 larger number of clusters, on the other hand, was not as impactful as hoped for clusters that
151 appeared to need splitting, and also made interpretation and analysis more unwieldy. One
152 appealing aspect of our $K=10$ choice is the absence of extreme disparities in the resulting global
153 RFOs of the ACRs (Fig. 3b): global ACR RFOs range between 5.7 and 15.1 indicating a rather
154 balanced assignment of global cloudiness, with no ACR being overly rare or dominant.



155

156 **Figure 3.** (a) Top two rows: observed ACR centroids, i.e., mean FOC histograms of the 10 CVS
 157 (non-clear) classes corresponding to the clusters obtained from the k -means algorithm for $K=10$
 158 applied on monthly 4° data. Global cloud fraction values in % (100 minus FOC of clear CVS)
 159 are also shown in each panel. (b) Bottom two rows: geographical distribution of ACR RFO after
 160 assignment of 4° monthly FOC histograms to the closest ACR centroid. Global RFO values are
 161 shown at the top of each panel.

162 **3. Observed ACRs**

163 The two rows of Fig. 3a show the centroids of the derived ACRs, which are simply the
 164 mean histograms of all the monthly histograms of 4° grid cells that belong to the same cluster.
 165 The two rows of Figure 3b show the maps of relative RFO over the 4-year period; these were
 166 derived by assigning an ACR to each 4° gridcell based on minimum Euclidean distance between

167 its monthly histogram and the histogram representing the ACR centroid. The ACR centroids and
168 their occurrence maps are key pieces of information for interpreting their makeup and nature.

169 • ACR1 comprises a mixture of CVS where isolated high clouds ($CVS=H$) and multi-layer
170 vertically-developed cloud profiles that include high clouds ($CVS=H\times M$, $H\times M\times L$, HM)
171 occur at rates above climatology (compare with Fig. 1b). This ACR is more dominant in the
172 tropics than anywhere else, so is interpreted as corresponding largely to tropical deep
173 convection.

174 • ACR2 is also mostly of tropical nature and correlates geographically well with ACR1, but
175 differs from it in that $CVS=H$ dwarfs all other CVSs. ACR2 thus seems to contain much of
176 the cirrus and anvil outflows of tropical convection. Some association with orography can
177 also be seen (since the H designation is based on pressure level rather than height above
178 ground).

179 • ACR3 is largely confined within the tropics and subtropics, but has more isolated low
180 ($CVS=L$) clouds than the other two tropical ACRs, and in addition encompasses plentiful
181 isolated high clouds, which along with the largest among all ACRs presence of $CVS=HL$,
182 indicates also frequent co-existence of high and low clouds.

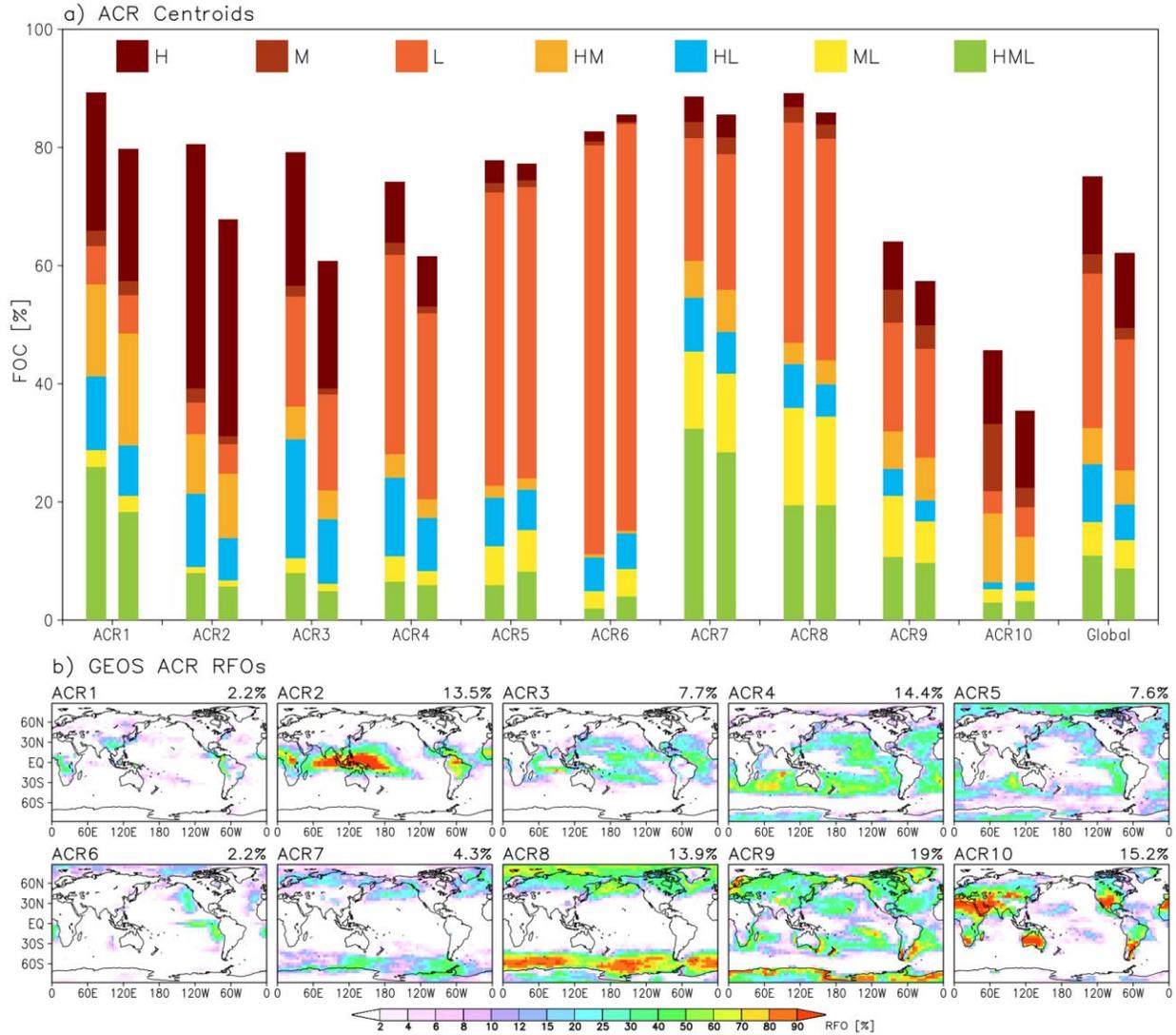
183 • ACR4 is a mixture of CVS that occupies most oceanic areas with the exception of the
184 northernmost and southernmost latitudes, and the tropical western Pacific. While $CVS=L$ is
185 the most populous CVS kind, substantial amounts of high cloud are also present either in
186 isolation or in co-occurrence with clouds at other levels. The average cloud fraction of $\sim 73\%$
187 indicates substantial occurrence of broken cloud conditions.

- 188 • ACR5 has similar geographical coverage as ACR4, but also extends further towards the
189 poles. But it also encompasses a larger fraction of $CVS=L$, contains fewer high clouds, and
190 its broken cloud conditions appear less prevalent.
- 191 • ACR6 seems to represent mostly marine stratocumulus clouds regimes. The dominance of
192 $CVS=L$ is unambiguous and average cloud fraction exceeds 80%.
- 193 • ACR7 is a high latitude marine cloud regime that approaches overcast conditions and is
194 apparently associated with cyclonic and frontal systems. It has more vertically-developed
195 clouds $H\times M\times L$ than even ACR1, indicating that precipitating systems are frequently
196 embedded in this ACR.
- 197 • ACR8 is similar to ACR7 in geographical coverage, but appears to correspond to weaker
198 systems with more $CVS=L$ clouds and fewer vertically developed systems $H\times M\times L$. Average
199 cloud fraction nevertheless remains very high.
- 200 • ACR9 stands for a mixture of CVS that is mainly continental in nature, is mostly
201 concentrated at high latitudes, is quite broken, and has no clear orographic association; it
202 does not appear to be closely linked to weather systems.
- 203 • ACR10 is a low cloud fraction ACR occurring over mountainous and dry areas (deserts,
204 Antarctica). A significant proportion of its $CVS=H$ and $CVS=M$ (largest proportion among
205 all ACRs) likely comes from clouds over areas of high relief.

206 **4. GEOS ACRs**

207 The atmospheric GEOS model (https://gmao.gsfc.nasa.gov/GEOS_systems/, Rienecker et
208 al., 2008) can also produce CVS as discussed in Lee et al. (2020) thanks to the implementation of
209 the Räisänen et al. (2004) stochastic subcolumn generator. In order to produce counterpart GEOS

210 CVS, we run a prescribed sea surface temperature (SST) simulation at 1° resolution and over the
211 same period as the observations with a version of GEOS (“v10.21.0”) that has descended from
212 the model version used in Lee et al. (2020). As explained in Lee et al. (2020), because the model
213 cloud profiles can be a lot more complex than in observations, prohibiting a meaningful
214 application of the Oreopoulos et al. (2017) CVS classification algorithm, the model’s CVS kinds
215 were reduced from 10 to 7 by combining the multilayer categories comprising clouds in the
216 adjacent standard vertical layers (delineated by the 680 hPa and 440 hPa pressure levels). With
217 the distinction between contiguous and non-contiguous clouds across the standard layers
218 eliminated, the observed ACR centroids are then also expressed in terms of the reduced set of 7
219 CVS as seen in Fig. 4a. This figure, in contrast to Fig. 3, uses a stacked bar representation of
220 each ACR’s CVS composition. The model’s 7-element CVS FOC histograms over a month in 4°
221 regions were assigned to the closest (according to minimum Euclidean Distance – the squared
222 difference of FOC between observed an observed histogram and a centroid) of the 10 observed
223 (and simplified) ACR centroids. The means of the histograms assigned to each ACR resulting
224 from this procedure represent the GEOS ACR centroids; these are also shown in Fig. 4a as
225 stacked bars to the right of their observational counterparts.



226

227 **Figure 4.** (a) ACR centroids from observations (first bar of pair) and GEOS (second bar of pair)
 228 from monthly FOC histograms in 4° regions resolved into the reduced set of 7 CVS classes (see
 229 text for explanation). A different visualization than Fig. 3 is used here where each centroid is
 230 represented by a stacked bar containing the CVS FOC. (b) RFO maps of GEOS ACRs after CVS
 231 FOC assignment according to minimum Euclidean Distance. Global RFO values are shown at
 232 the top of each panel.

233 One can see that the observed and modeled ACR centroids exhibit considerable similarity. This
 234 is largely by design because of the “forced” assignment of modeled CVS FOC histograms to the

235 closest, in terms of Euclidean distance, observed ACR centroid. Even when the resemblance of
236 modeled histograms assigned to an observed ACR is imperfect, it is still better than that for other
237 ACRs, making the modeled ACR centroid (mean of modeled histograms) a reasonably close
238 (albeit flawed) kin of its observed counterpart. While individual occurrences of simulated ACRs
239 represent clouds similar to those in observed members of the same real-world ACR, the overall
240 frequency and locations at which alike simulated CVS histograms tend to occur differs from
241 observations, ultimately resulting in substantial discrepancies in geographical characteristics and
242 global occurrence statistics. This can be readily noticed when comparing the RFO maps of
243 GEOS ACRs in Fig. 4b with the counterpart maps of observed ACRs shown in Fig. 3b. Focusing
244 on differences in global and geographical RFO between observations and model of similar ACRs
245 allows for a much more straightforward interpretation of cloudiness differences than performing
246 clustering on modeled CVS FOC histograms; such an approach would have likely led to very
247 different ACRs, and would have made comparisons quite cumbersome.

248 When comparing global cloud fractions (sum of FOCs of all CVS other than the clear CVS), the
249 observed and modeled ACRs that differ the most, in spite of our “forced” assignment, are ACR1,
250 ACR2, ACR3, ACR4, and ACR10. The global cloud fraction differences arise in turn from
251 differences in the global FOC of certain constituent CVS kinds. The GEOS version of ACR1 is
252 overall less extensive, even though it has proportionally far more isolated CVS = *H* clouds. This
253 is because it also has far fewer vertically deep CVS = *HML* clouds (recall that this category now
254 includes CVS = *H×M×L* which are far more numerous in observations than original *HML*), and
255 also fewer CVS = *HL* clouds. Observed isolated CVS = *H* clouds are frequently encountered in
256 CVS mixtures that belong to ACR2, but *H* clouds in the GEOS version of ACR2 are less
257 common. Underestimates of CVS = *HL* are notable in modeled CVS FOC histograms assigned to
258 ACR4, but this issue seems to be related to the broader deficiency of too few low clouds in

259 GEOS (see ACR4's discrepancy for $CVS=L$). Even when the two ACR versions have similar
260 cloud fractions, this is the result of compensations of inconsistent contributions from different
261 CVSs. ACR5 is an example of such a situation: the model has more $CVS = HML$ clouds, but this
262 is compensated by fewer $CVS = H$ and HL clouds.

263 When ACR global RFOs and (normalized) regional frequencies are examined, even more
264 discrepancies between model and observations emerge (contrast the maps of Figs. 3 and 4 and
265 the accompanying global RFO values of each panel). The global RFO is substantially different
266 for all ACRs except ACR4, for which the modeled geographical distribution is actually also
267 quite realistic. The model's convection seems to be expressed mostly via a generous production
268 of $CVS = H$ clouds which are featured in ACR2, to the detriment of $CVS = HML$ clouds featured
269 in ACR1, making ACR1 quite rare in GEOS. ACR3 is quite reasonably simulated in the model;
270 but this not the case for ACR6, with apparently many of the model's marine stratocumulus (CVS
271 $= L$ clouds) going to to ACR5. But ACR5 on the other hand suffers from little southern ocean
272 presence, as it is replaced in that region by ACR8. While ACR9 is mostly continental in
273 observations, it has nearly as much oceanic presence in GEOS and emerges as the most populous
274 ACR. Finally, the model seems to have a cloudier Antarctic continent, with ACR4, ACR5,
275 ACR8, and ACR9 presence, whilst in the observations the least cloudy ACR10 dominates.

276 In the following section we examine how these differences translate to biases in the breakdown
277 by ACR of radiative flux and diabatic heating. The analysis reveals that even if overall the
278 model's global cloudiness values are (unsurprisingly) close to observations, their distribution
279 among the world's major cloud systems, expressed via ACRs, remains deficient.

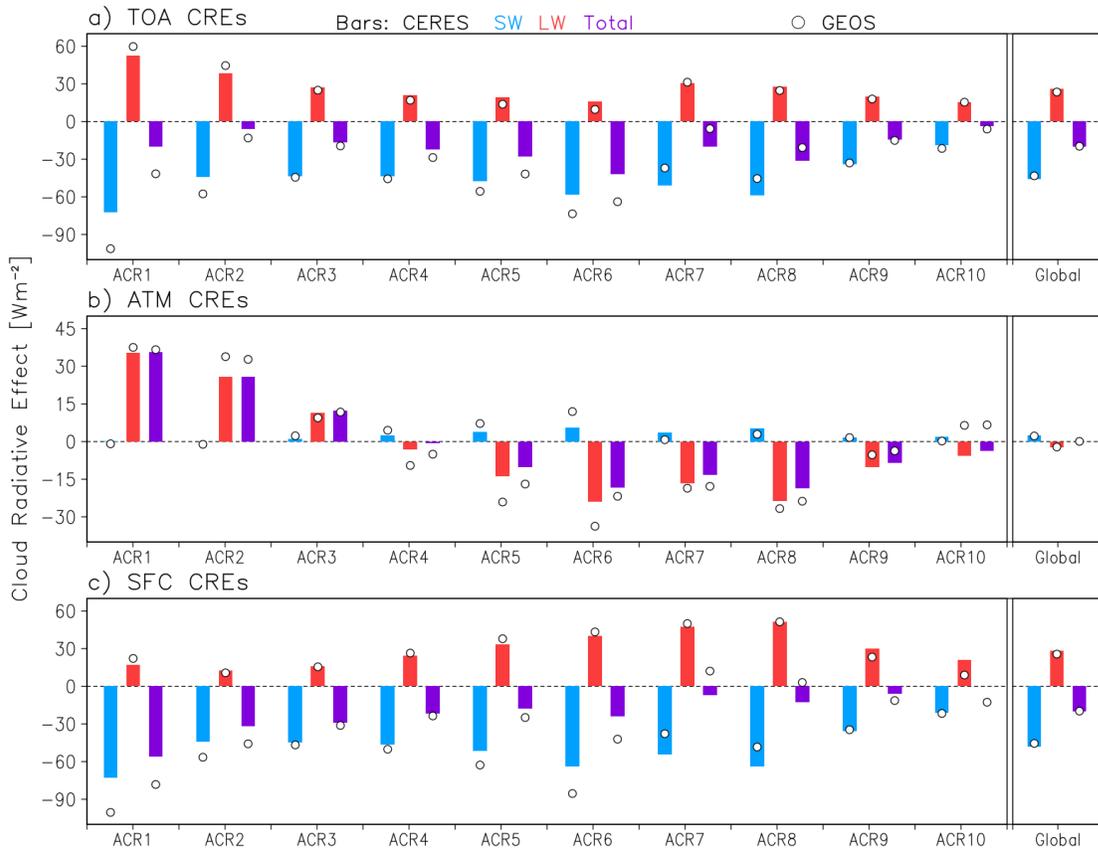
280 **5. Energetics of observed vs modeled ACRs**

281 This section examines how the world's cloud systems, when classified via ACRs, regulate major
282 energetic components of the atmosphere and the planet as a whole. First, we look at how ACRs
283 change radiative fluxes and heating rates with respect to cloudless skies in both the shortwave
284 (SW) and longwave (LW) part of the spectrum, and then how they affect the atmospheric
285 diabatic heating, especially the dominant radiative and latent heat components.

286 *5.1 Integrated cloud radiative effects in the atmosphere-surface column*

287 The Cloud Radiative Effect CRE is the difference between net (down minus up) all-sky and
288 clear-sky fluxes and serves as a measure of the impact of clouds on atmospheric radiative fluxes.
289 The observed mean CRE of each ACR for SW (blue bars), LW (red bars), and total = SW+LW
290 (purple bars) at the boundaries of atmospheric column, namely the top-of-the-atmosphere (TOA)
291 and the surface (SFC) is shown in Fig. 5's top and bottom panels, respectively. The difference
292 between TOA and SFC CREs is the atmospheric CRE which is shown in the middle panel. CRE
293 comes from CERES-EBAF (Loeb et al. 2018; Kato et al. 2018), specifically the latest version
294 Ed4.2 which provides monthly values on a 1° grid, and has thus been coarsened to 4° for our
295 purposes to match the ACR spatial resolution. CRE is then averaged separately for each ACR.
296 The ability to use CERES-EBAF in conjunction with CALIPSO and CloudSat cloud retrievals is
297 a testament to the utility of the ACR concept. Vertical cloud masks from active observations
298 originally resolved in the 2BCL dataset at ~2 km have been transformed into a gridded dataset
299 via ACRs and can now be combined with another gridded dataset. The only previous choice for

300 combining radiative fluxes with CloudSat/CALIPSO 2BCL clouds was the 2BFL product (Matus
 301 et al. 2017; Oreopoulos et al 2017; L'Ecuyer et al. 2019; Lee et al. 2020).



302
 303 **Figure 5.** SW, LW, and total=SW+LW CRE from CERES-EBAF for (a) TOA , (b) ATM, (c) SFC
 304 averaged across ACR occurrences (i.e, not RFO-weighted which are shown in the Supplemental
 305 Material Fig. S1). Observed values are shown as bars while modeled values as open circles. TOA
 306 and SFC CRE is defined as the double difference between downward and upward all-sky minus
 307 clear sky fluxes. ATM CRE is the difference between TOA and SFC CRE.

308 Open circle symbols in Fig. 5 correspond to the counterpart CRE breakdown by ACR from
 309 GEOS. It is important to underscore that the CRE means are averages across ACR occurrences
 310 only, i.e., showcase the average strength of CRE for each ACR *when it occurs*. CREs weighted

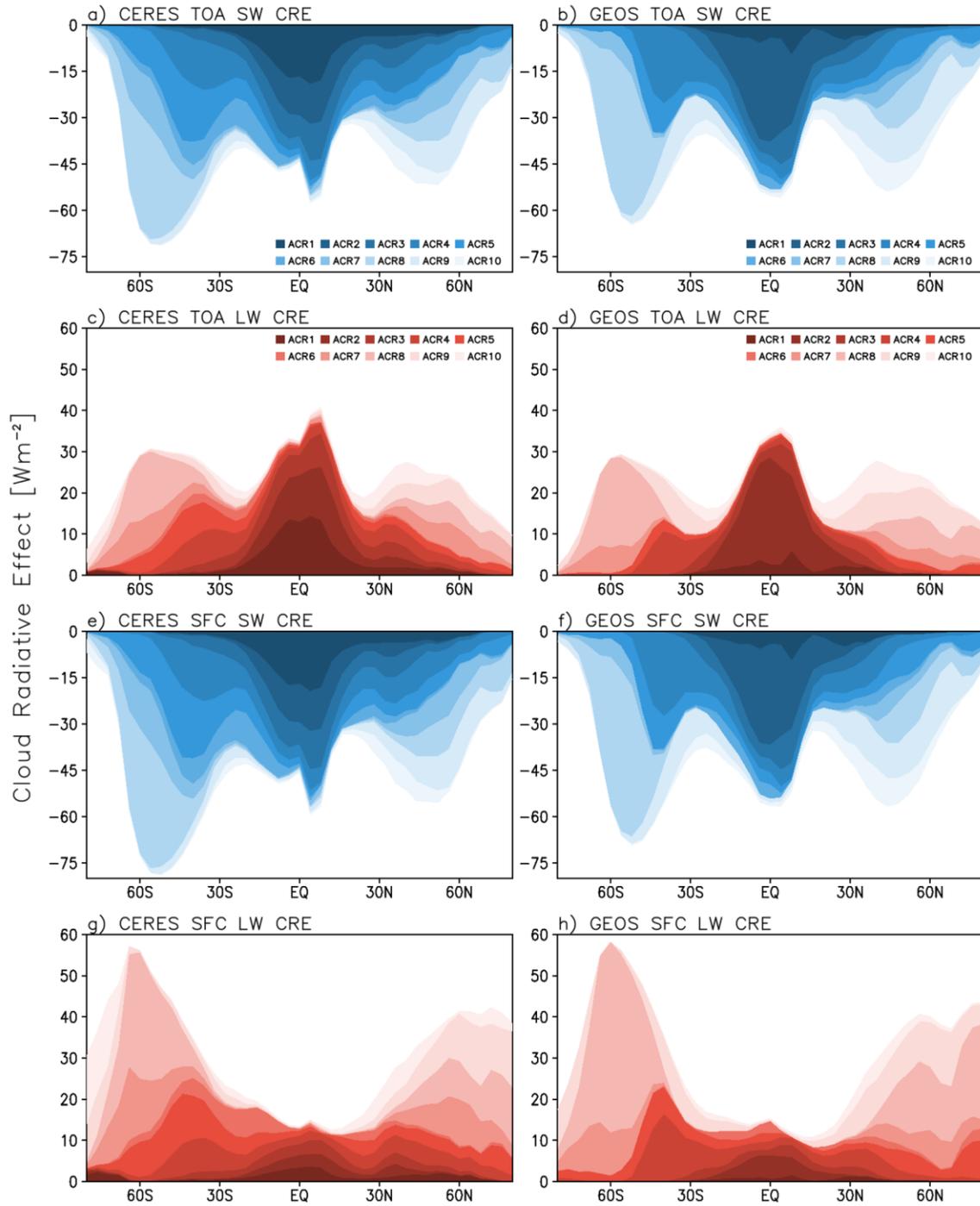
311 by ACR RFO to indicate the contribution of each ACR to the global CRE are shown in the
312 Supplementary Information (Fig. S1). The TOA and SFC CREs indicate whether ACRs cool
313 (negative values) or warm (positive values) the planet as a whole (TOA) or its surface (SFC).
314 Their difference (ATM CRE) indicates how much ACRs reduce (positive values) or enhance
315 (negative values) the atmospheric clear-sky radiative cooling. The SW ATM CRE is typically a
316 small positive value or near zero (due to the competition between added cloud absorption and
317 reduced water vapor absorption below clouds), while the LW ATM CRE can be either positive
318 or negative.

319 Figure 5 illuminates various interesting aspects of the global radiative role of ACRs. All ACRs
320 have a planetary cooling effect, as seen by the negative purple bars at the TOA, stemming from
321 the SW CRE being larger in absolute value than the LW CRE. They also have a cooling effect at
322 the surface across all ACRs, but the magnitudes of total CRE differ at the two boundaries. The
323 strongest TOA cooling corresponds to ACR6 dominated by $CVS = L$ which has a much weaker
324 TOA LW CRE, because low clouds do not change much the height of atmosphere's effective
325 emission to space, than TOA SW CRE. But at the SFC where the SW CRE remains strong, the
326 LW CRE is also strong because the increased atmospheric opacity of low clouds at warm
327 tropospheric temperatures results in an enhancement in the downward LW flux. The weak
328 cooling at the SFC compared to the strong cooling at the TOA means that substantial cooling
329 occurs within the atmosphere. This is indeed shown by ACR6's purple bar in the middle panel of
330 Fig. 5. A similar behavior is also observed for ACR8 which has proportionally less $CVS = L$, but
331 apparently enough to behave similarly. Contrast this with ACR2 which has many $CVS = H$
332 clouds: SW and LW CRE at TOA almost cancel (near zero total CRE), but the SW CRE
333 dominates at the SFC (approximately the same value as in TOA, similarly to all other ACRs)
334 because the LW CRE is small: despite the added opacity, clouds are too high to have much of an

335 impact to the downwelling LW at the SFC which originates mostly from lower atmospheric
336 levels (in other words, ACR2 does not reduce much the effective height of downward emission).
337 This makes ACR2's surface cooling much stronger than its TOA cooling which in turns means
338 that large atmospheric warming takes place, something confirmed by the middle panel.

339 Figure 6 shows the annual zonal distribution of the contribution of each ACR to the TOA and
340 SFC SW and LW CRE. These plots are based on RFO-weighted CREs (the ACR zonal RFOs
341 themselves are shown in Fig. 9), and are similar to those in L'Ecuyer et al. (2019) who showed
342 zonal contributions to 2BFL CRE by 2BCL cloud types. In this type of plot the zonal CRE of
343 each ACR is added to the cumulative CRE of all ACRs that preceded it in our (somewhat
344 arbitrary) ordering of ACRs. So, for example, the zonal CRE values shown for ACR6 are
345 actually the cumulative CRE values from ACR1 to ACR6, and the ACR10 zonal values are thus
346 the cumulative zonal CRE contributions of all ACRs. For illustration purposes the area under the
347 curves representing the successive cumulative values is color-filled so that the incremental
348 contribution of each ACR to the zonal CRE values can be visualized by the area each ACR adds
349 to the cumulative zonal curve of all previous ACRs. The left column of panels shows observed
350 values from CERES EBAF, while the right column shows counterpart GEOS values. These
351 zonal plots make clear the tropical dominance in observations of ACR1 and ACR2 for SW and
352 LW TOA CRE, and SW SFC CRE; their LW SFC CRE contribution is more muted which is
353 consistent with the weak values of this CRE in the tropics where the high humidity of the lower
354 troposphere suppresses the downward longwave cloud emission reaching the surface. Fig. 6 also
355 reveals a strong hemispheric asymmetry at midlatitudes with ACR8 making much larger

356 contributions to all four CRE components in the Southern Hemisphere because of its large
 357 presence in that part of the world (see ACR RFO panel in Fig. 9).



359 **Figure 6.** Contributions of each ACR to the annual zonal TOA and SFC SW and LW CRE from
360 CERES EBAF RFO-weighted CRE values (the zonal values of the RFOs themselves are shown
361 in Fig. 9). The values corresponding to each ACR are cumulative and include the contribution of
362 that ACR along with that of all those that precede it in our ACR ordering. See text for more
363 details. The left column of panels corresponds to observations while the right column to GEOS.

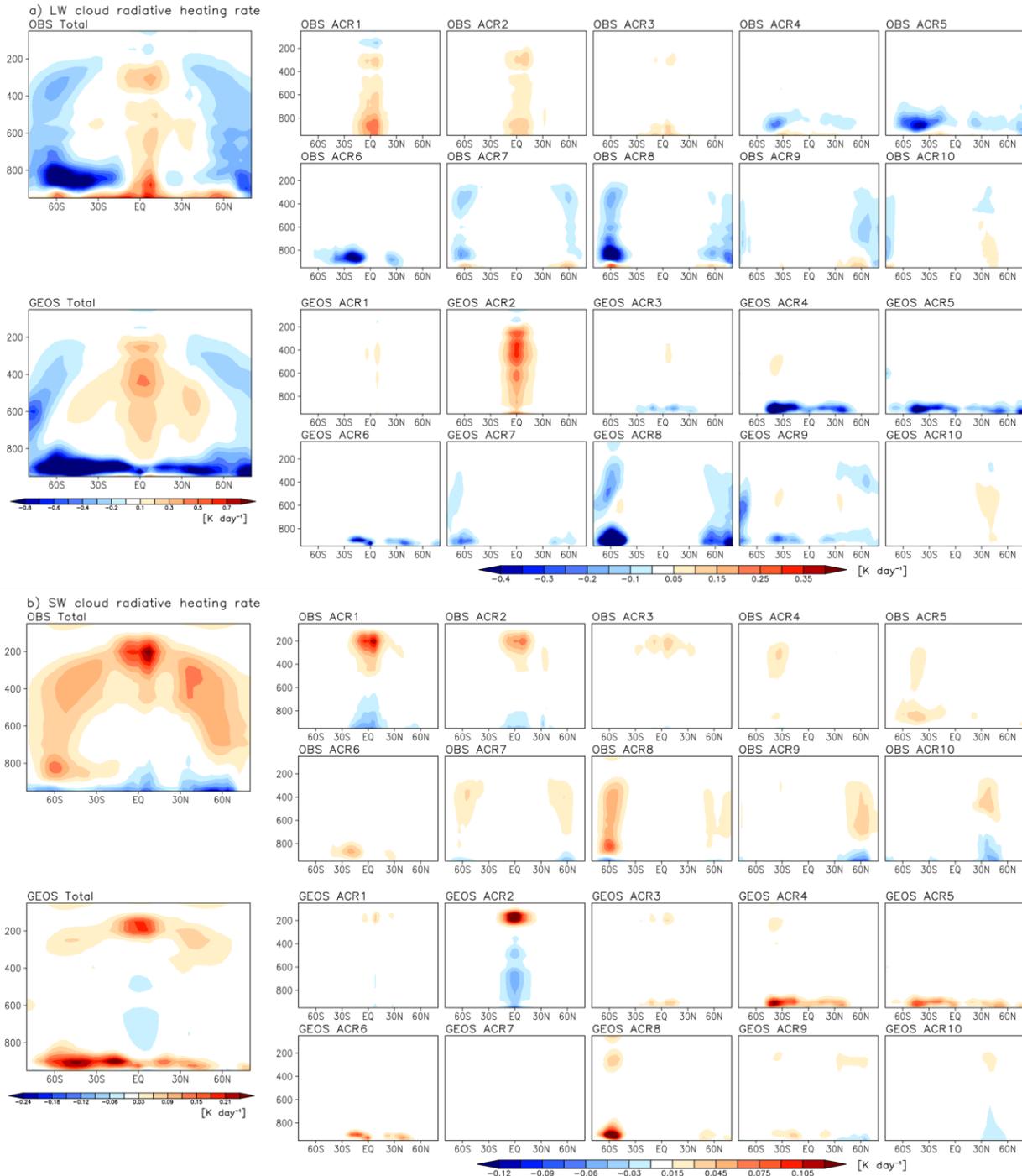
364 Figures 5 and 6 make apparent the great diversity of CRE stemming from the planet's various
365 cloud systems. Given strong seasonal variations of solar illumination outside the tropics, even
366 more SW CRE variability throughout the year would have been seen for ACRs prevailing at
367 those latitudes, if seasonal CREs were to be examined. A lot of CRE diversity is also seen in
368 GEOS (open circles in Fig. 5), but the breakdown of CRE by ACR generally deviates from
369 observations with a few exceptions, such as ACR3 and ACR4. Also, when different CRE
370 components are combined, e.g., adding SW and LW CRE or subtracting SFC from TOA CRE,
371 the model can fortuitously approach the observed CRE. See for example how well ACR1's total
372 ATM CRE is simulated despite large TOA and SFC discrepancies. Moreover, also note how well
373 the model simulates the global CRE (rightmost part of the Fig. 5 panels): despite errors in the
374 global CRE breakdown by ACR, when everything is added together the biases cancel out almost
375 perfectly. The TOA and SFC CRE observation-model discrepancies at the ACR level can be
376 further scrutinized with the Fig. 6 zonal plots. When comparing the panels of the left and right
377 columns of Fig. 6 one can see again the contrast in ACR1 and ACR2 CRE contributions; the
378 model's weaker SW CRE and weaker LW CRE in the in the Southern Hemisphere midlatitudes

379 and subtropics, respectively; and the bigger importance of its ACR8-ACR10 in the Northern
380 Hemisphere midlatitudes.

381 *5.2 Vertically resolved atmospheric cloud radiative effects*

382 The ATM CRE can be further resolved vertically using 2BFL flux profiles which we composite
383 by ACR for the 2007-2010 period. Since the 2BFL dataset matches spatiotemporally 2BCL, a
384 breakdown of vertically resolved ATM CRE over the same period could have also been
385 performed by CVS, and would have also been insightful, but would not reflect the reality of CVS
386 mixing regionally. Annual zonal averages (resolved in 4° latitude) of the vertically resolved
387 cloud radiative heating rate CRHR for each ACR for both observations and GEOS are shown in
388 Fig. 7a for LW, and Fig. 7b for SW. The CRHR is basically the difference of the net (down-up)
389 flux profiles between cloudy skies and clear skies expressed in units of K/day, and represents
390 physically the cloud-induced change of the radiative heating/cooling rate of clear skies. The
391 CRHRs of the ACRs add to the total CRHR shown in the leftmost panels of Fig. 7. The SI shows
392 counterpart seasonal (winter and summer) plots (Fig. S2) which allow us to glean radiative

393 effects on the atmosphere of seasonal movements of cloud systems in the two hemispheres and
 394 to compare with similar plots by Haynes et al (2013).



396 **Figure 7.** (a) Zonal-height plots of the annual LW CRHR corresponding to the 10 ACRs; (b)
397 Same as (a), but for SW CRHR. For both the LW and SW CRHR, the top two rows of panels
398 correspond to observations while the bottom two rows correspond to GEOS. In both figures the
399 LW and SW zonal-height CRHR distribution regardless of ACR is also shown (the bigger panels
400 on the left).

401 The top two rows of Fig. 7a reveal that ACR1-ACR3 have a radiative warming effect, i.e.,
402 induce a reduction in the rate of clear-sky radiative cooling, on a range of tropospheric altitudes.
403 Other than the highest tropospheric levels where there is enhanced (relative to cloudy skies)
404 cloud emission to space and thus enhanced cooling, ACR1 suppresses the atmospheric cooling in
405 the rest of the troposphere by enhanced downward emission counteracting upward emission, i.e.,
406 reducing the efficiency of radiative cooling. ACR2 exhibits weaker cooling reduction and no
407 measurable upper tropospheric cooling since its highest clouds are not as optically thick as those
408 of ACR1. ACR3's clouds are not as vertically developed, hence their trapping and downward re-
409 emission of LW radiation is much weaker. The other ACRs with substantial cloud amounts
410 (ACR4-ACR9) enhance atmospheric LW radiative cooling (Fig. 5) to various degrees, with the
411 exact profiles of radiative cooling enhancement depending on the exact CVSs that are prevalent
412 within the ACR. For example, ACR6's cooling is much more concentrated in the lower
413 troposphere because the dominance of $CVS = L$ is stronger than for ACR8 which has also a
414 considerable proportion of middle and low clouds (consolidated $CVS = ML$ and $CVS = HML$)

415 Fig. 7b, the counterpart of Fig. 7a for SW, illustrates in the top two rows the observed
416 enhancement or reduction at various tropospheric levels of clear-sky SW radiative heating. In
417 general, ACRs enhance SW radiative heating because of the additional solar absorption provided
418 by cloud particles, except at low altitudes where reduction in heating can occur as overlying

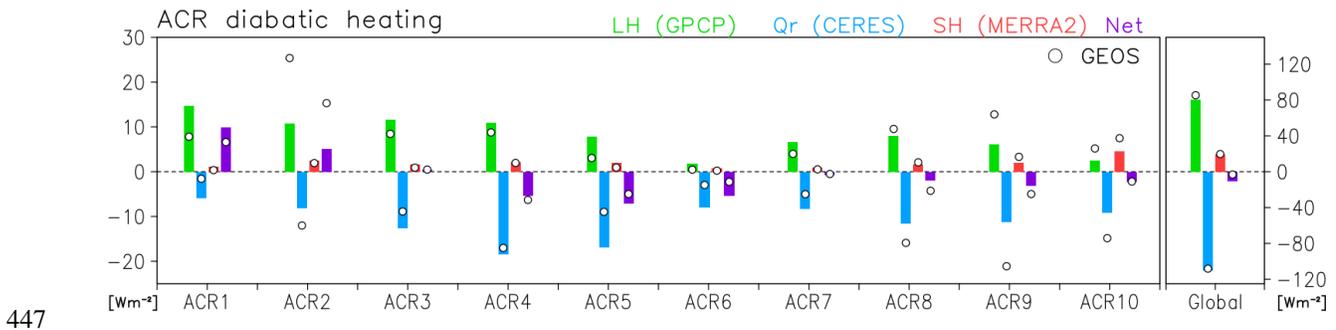
419 clouds block radiation that would have otherwise propagated to the lower tropospheric levels
420 where boundary layer clouds and low-level water vapor await to induce absorption. When there
421 are no substantial amounts of overlying high clouds such in ACR6, the solar absorption added by
422 clouds manifests only in the lower troposphere where most clouds reside.

423 The bottom two rows of panels of the two parts of Fig. 7 show the decomposition of CRHR by
424 ACR in GEOS. We first note that the overall CRHR height-latitude distribution in GEOS differs
425 substantially from observations. GEOS's LW CRHR plots reveal a pervasive absence of middle
426 and high clouds to prevent or temper the LW cooling of the lowermost clouds. As a result, low
427 cloud cooling is present at all latitudes and confined to the lowest 200 hPa of the troposphere. On
428 the SW side, the absence of middle-level clouds and the overall smaller amounts of low clouds
429 close to the top of the boundary layer or just above it, suppresses the cooling that would have
430 occurred due to reduced water vapor absorption in the lowermost troposphere. Instead, SW
431 absorption takes place by the boundary layer clouds themselves either when there is not much
432 overlying cloudiness to dampen the downward propagation of solar flux (ACR4, ACR5, and
433 ACR6), or even when there is (ACR8).

434 *5.3 Atmospheric Diabatic Heating*

435 The breakdown by ACR of the atmospheric diabatic heating and its three components (latent
436 heating LH, atmospheric radiative cooling Q_r , and sensible heating SH) according to
437 observations/reanalysis and GEOS are shown in Fig. 8. Observed LH (green bars) comes from
438 GPCP monthly precipitation (Version 3.2, Huffman et al., 2023) values converted to Wm^{-2} by
439 multiplying the rain rates with the latent heat of vaporization using an approximate relationship,
440 instead of surface evaporative fluxes. The observed Q_r (blue bars) is the atmospheric radiative
441 flux divergence and is calculated from CERES EBAF Ed4.2 as the difference between TOA and

442 surface combined SW and LW all-sky net (=down-up) fluxes. SH (green bars) comes from
 443 MERRA-2 reanalysis (GMAO 2015, Gelaro et al., 2017). The values shown have been weighted
 444 by the annual global RFO of each ACR, so represent ACR contributions to global values of
 445 diabatic heating and its components (i.e., they yield the proper global value when added). The
 446 corresponding GEOS values are shown with open circle symbols.



447
 448 **Figure 8.** The observed total atmospheric diabatic heating (purple bars) and its components
 449 latent heat LH, radiative flux divergence Q_r , and sensible heat SH (green, blue, red bars,
 450 respectively). Counterpart values for the atmospheric GEOS model are shown as open circles.
 451 Global values irregardless of ACR are shown in the rightmost part of the plot.

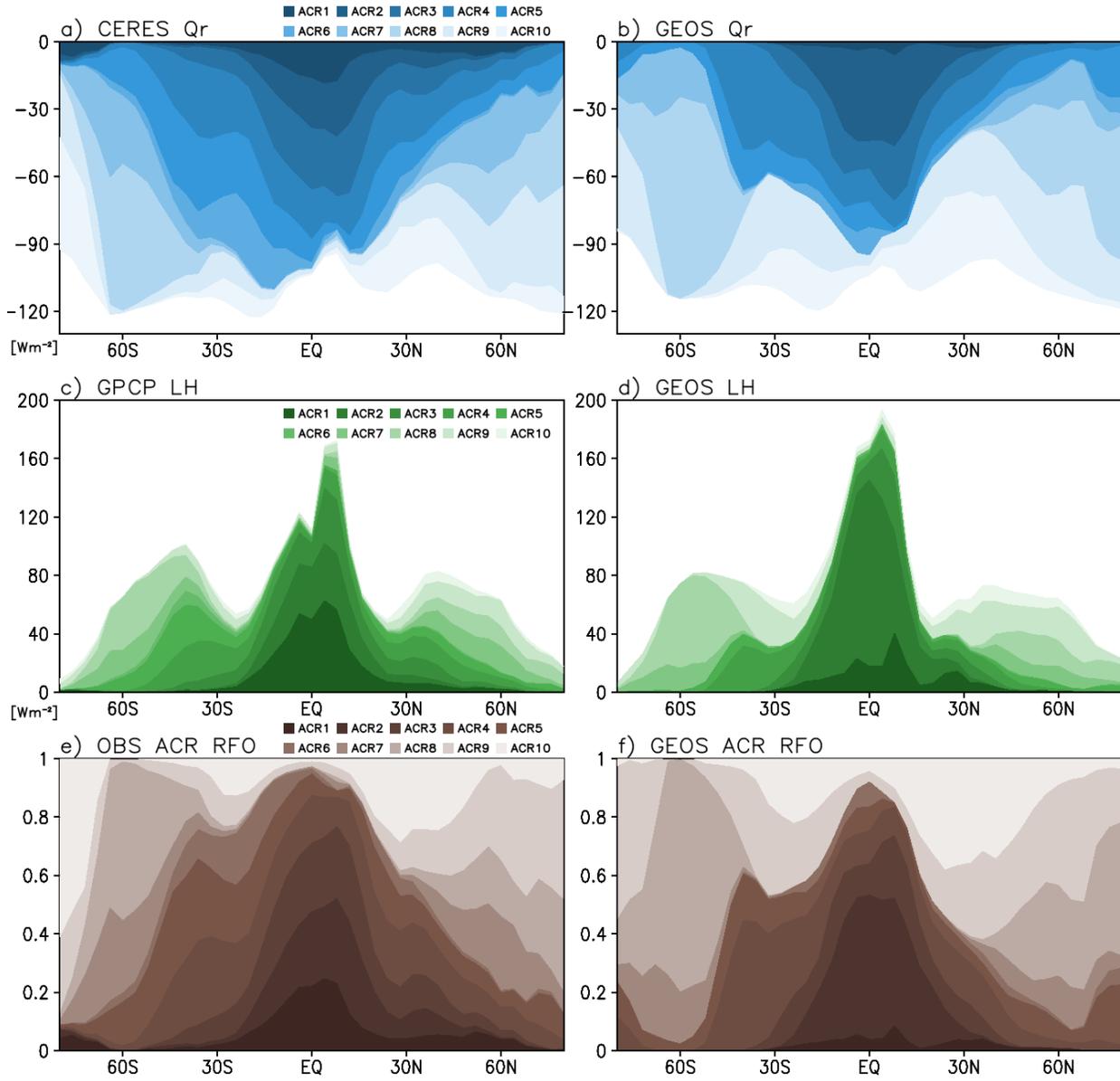
452 According to Fig. 8, only the two mostly tropical ACRs, ACR1 and ACR2 have positive diabatic
 453 heating mainly because of the small Q_r in the presence of plentiful high clouds that have a
 454 positive ATM CRE. ACR1's convection not only reduces greatly (in absolute terms) Q_r , but also
 455 produces large amounts of precipitation, and thus yields the largest diabatic heating. The largest
 456 (in absolute terms) contribution to the (negative) global Q_r does not coincide with the ACR that
 457 has the largest proportion of CVS = L clouds (ACR6), but rather with ACR4 which is much
 458 more prevalent. ACR6 is though associated with the lowest precipitation, making its diabatic
 459 heating substantially negative, as there is not much counteraction for the negative Q_r . Despite

460 higher precipitation than ACR6, ACR5's negative total diabatic heating (the largest of all ACRs)
461 is larger because the radiative cooling added by its clouds is so much larger.

462 The global diabatic heating shown in the rightmost part of the plot is near-zero (within the
463 uncertainties of the datasets), as expected, in accordance with the planet being in radiative-
464 convective equilibrium, where the atmospheric radiative cooling is compensated by the surface
465 turbulent fluxes (or their atmospheric surrogates). The GEOS global values agree well with
466 observations despite large disagreements in the breakdown by ACR. Particularly striking is the
467 completely different energy distribution of GEOS for ACR1 and ACR2. The model's version of
468 ACR1 produces much less precipitation for ACR1, but compensates with excessive precipitation
469 for ACR2 despite the fact that its ratio of $CVS = HML$ to $CVS = H$ clouds is much lower than
470 observations, which in turns implies that modeled $CVS = H$ clouds are strong precipitation
471 producers. Combined with a larger reduction of Q_r in GEOS, the total diabatic heating of ACR2
472 in GEOS is more than twice as large as in observations and actually overcompensates for the
473 underestimate of diabatic heating by ACR1.

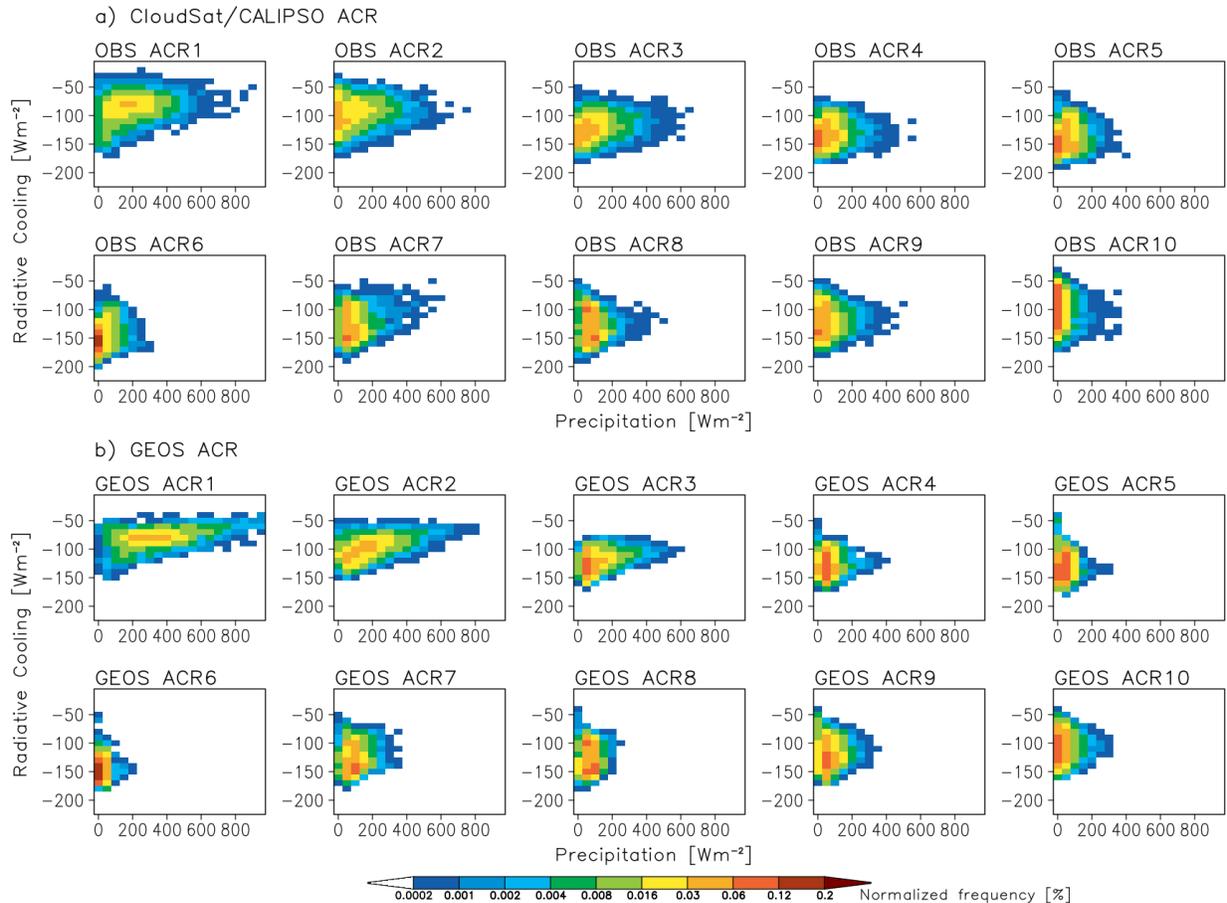
474 Figure 9 shows the contribution of each ACR to the annual zonal variation of two main adiabatic
475 heating components (Q_r and LH) in the same manner as Fig. 6. The left column of panels
476 corresponds to observations while the right column to GEOS; the bottom row of panels shows
477 the zonal variation of ACR RFO. While the zonal distributions of the cumulative Q_r and LH (i.e.,
478 regardless of ACR) are quite similar between observations and model, ACR contributions are
479 quite dissimilar. The diminished role in GEOS of ACR1 relative to that of ACR2 is again clearly
480 seen in the tropics for both Q_r and LH. ACR8 and ACR9 are bigger contributors in the model's

481 Northern Hemisphere than in observations because of their larger RFOs; ACR5 is a much
 482 smaller contributor everywhere compared to observations because of its global scarcity.



483

484 **Figure 9.** Contributions by ACR to the zonal variation of the two main adiabatic heating
 485 components (Q_r and LH) shown in the same manner as Fig. 6. The left column of panels
 486 corresponds to observations while the right column to GEOS; the bottom row of panels shows
 487 the zonal variation of ACR RFO.



488

489 **Figure 10.** Density plots of the two major diabatic heating components, namely atmospheric flux
 490 divergence (radiative cooling) vs latent heat (i.e., precipitation in Wm⁻²) for the 10 ACRs. The
 491 top two rows of panels (a) correspond to observations, while the bottom two rows (b) to GEOS.
 492 Joint histograms (density plots) of Q_r (from CERES-EBAF Ed4.2, as above) and precipitation
 493 (LH) from GPCP are shown in Fig. 10. This type of plot was introduced by Jakob et al. (2019) to
 494 elucidate how the competition between radiative cooling and latent heating due to precipitation
 495 results in a radiative-convective quasi-equilibrium state in the tropics. At first glance, it appears
 496 that the model's density plots exhibit substantial resemblance to observations, but more careful
 497 comparison reveals several differences. It is interesting that despite the overall weaker
 498 precipitation of its ACR1, the model produces occasionally extreme high values not seen in the

499 observations. One should however not place unwarranted attention to these extremes as they are
500 quite rare (note the colorbar's logarithmic scale). The same also happens to some extent for
501 ACR2. For the remaining ACRs the opposite is true, with the real world having a wider
502 precipitation range. A wider range of values in observations is also seen for atmospheric cooling,
503 although once again, this impression is exaggerated by the non-linear color scale. The model's
504 overestimation of precipitation for ACR2 and ACR9 is seen here as a rightward shift of peak
505 populations. On the flipside, these ACRs have smaller populations in the weak radiative cooling
506 part of the phase space, reflecting the overestimation of radiative cooling needed to compensate
507 for the exaggerated precipitation.

508 **6. Discussion and conclusions**

509 We propose a new way to characterize sub-seasonal (monthly) cloudiness over large ~400 km
510 regions that essentially classifies recurrent patterns of cloud vertical morphology. The coarse
511 spatial resolution is a necessary compromise that previous investigators (Haynes et al 2013;
512 L'Ecuyer et al 2019; Hang et al 2019) had to also accept when presenting climatological cloud
513 characteristics from spatially sparse CloudSat and CALIPSO observations. We show that 10
514 classes of so-called active cloud regimes (ACRs) describe effectively and meaningfully where
515 and how frequently different types of vertical cloud configuration coexist in approximately
516 recurrent proportions. We demonstrate that the ACRs are particularly suitable for understanding
517 the radiative impacts of different cloud vertical structure mixtures, namely their relative
518 contributions to planetary and surface cooling and to enhancements or reductions of the full-
519 column and vertically resolved radiative cooling of the atmosphere. Furthermore, we used ACRs

520 as a foundational dataset to understand different states of the competition between atmospheric
521 radiative cooling and latent heating.

522 Simulating realistic ACRs in Global Climate Model is a stringent requirement that cannot even
523 be assessed unless the model can produce CVS, which in turn requires a cloud subcolumn
524 generator. NASA's GEOS implements such a generator, but when it comes to producing credible
525 CVS the model has mixed success (Lee et al. 2020). This is not only because of the generator's
526 inadequacies, but also (primarily) of the flawed cloud fraction profiles supplied to the generator.
527 When CVS are unrealistic, so will also inexorably be ACRs. Therefore, when assessed in terms
528 of ACRs, GEOS does not seem to have enough vertically developed clouds, has too many
529 isolated high clouds which produce unphysical amounts of precipitation, lacks mid-level clouds,
530 and forms low clouds too close to the surface. Its global energetics are good, but the breakdown
531 of energy fluxes by ACR deviates from the real world.

532 We view ACRs as complementary to cloud classifications based on passive observations such as
533 cloud regimes (CRs; Oreopoulos et al. 2014; Oreopoulos et al. 2016; Cho et al. 2021) and
534 regimes of regimes (RORs; Cho et al 2024, accepted manuscript). While ACRs contain
535 information about the vertical structure of clouds, they do not explicitly contain information on
536 cloud opacity which primarily controls propagation of solar radiation. While cloud vertical
537 extent correlates with opacity, there is no strict one-to-one correspondence since hydrometeor
538 concentrations can vary widely. Passively-derived CRs are similarly deficient in providing
539 unambiguous descriptions of how clouds interact with thermal infrared radiation, especially
540 when it comes to radiation emitted towards the surface. Thus both active and passive
541 classifications have to be used synergistically for a more complete picture of cloud climatologies,
542 and to understand better how broad cloud classes regulate the planet's energy budget. We plan to

543 conduct such work in the future, conceivably under a framework of hybrid passive-active cloud
544 regimes.

545 **Acknowledgments**

546 Resources supporting this work were provided by the NASA High-End Computing (HEC)
547 Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight
548 Center. Support by NASA's CloudSat-CALIPSO Science Team program is gratefully
549 acknowledged.

550

551 **Open Research**

552 Observed CVS, and observed/modeled ACRs and the data used to produce the manuscript's
553 figures are available at <https://doi.org/10.5281/zenodo.13891853>. The 2B-CLDCLASS-LIDAR
554 and 2B_FLXHR-LIDAR datasets are available at <http://www.cloudsat.cira.colostate.edu/> and
555 CERES EBAF Ed4.2 dataset (NASA/LARC/SD/ASDC., 2023) is available at
556 <https://ceres.larc.nasa.gov/data/>. GPCP v3.2 dataset (Huffman et al., 2022) can be downloaded
557 from <https://measures.gesdisc.eosdis.nasa.gov/data/GPCP/GPCPMON.3.2/>. MERRA-2
558 tavgM_2d_flx_Nx dataset (GMAO, 2015) is available at
559 https://goldsmr4.gesdisc.eosdis.nasa.gov/data/MERRA2_MONTHLY/M2TMNXFLX.5.12.4/.
560 The baseline GEOS model source code can be accessed at [https://github.com/GEOS-](https://github.com/GEOS-ESM/GEOSgcm/releases/tag/v10.21.0)
561 [ESM/GEOSgcm/releases/tag/v10.21.0](https://github.com/GEOS-ESM/GEOSgcm/releases/tag/v10.21.0).

562

563 **References**

564 Cho, N., Tan, J., & Oreopoulos, L. (2021). Classifying planetary cloudiness with an updated set
565 of MODIS cloud regimes. *Journal of Applied Meteorology and Climatology*, 60(7), 981-997.

566 Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A.,
567 Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C.,
568 Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G., Koster, R., Lucchesi, R.,
569 Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.
570 D., Sienkiewicz, M., & Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research
571 and Applications, Version 2 (MERRA-2). *Journal of Climate*, 30(14), 5419-5454.
572 <https://doi.org/10.1175/JCLI-D-16-0758.1>

573 Global Modeling and Assimilation Office (GMAO) (2015), MERRA-2 tavgM_2d_flux_Nx:
574 2d,Monthly mean,Time-Averaged,Single-Level,Assimilation,Surface Flux Diagnostics V5.12.4,
575 Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES
576 DISC), Accessed: [05/30/2024], 10.5067/0JRLVL8YV2Y4

577 Hang, Y., L'Ecuyer, T. S., Henderson, D. S., Matus, A. V., & Wang, Z. (2019). Reassessing the
578 effect of cloud type on Earth's energy balance in the age of active spaceborne observations. Part
579 II: Atmospheric heating. *Journal of Climate*, 32(19), 6219-6236.

580 Haynes, J. M., Vonder Haar, T. H., L'Ecuyer, T., & Henderson, D. (2013). Radiative heating
581 characteristics of Earth's cloudy atmosphere from vertically resolved active sensors. *Geophysical*
582 *Research Letters*, 40(3), 624-630.

583 Henderson, D. S., L'Ecuyer, T., Stephens, G., Partain, P., & Sekiguchi, M. (2013). A multisensor
584 perspective on the radiative impacts of clouds and aerosols. *Journal of Applied Meteorology and*
585 *Climatology*, 52(4), 853-871.

586 Huffman, G. J., Adler, R. F., Behrangi, A., Bolvin, D. T., Nelkin, E. J., Gu, G., & Ehsani, M. R.
587 (2023). The New Version 3.2 Global Precipitation Climatology Project (GPCP) Monthly and
588 Daily Precipitation Products. *Journal of Climate*, 36(21), 7635-7655.
589 <https://doi.org/10.1175/JCLI-D-23-0123.1>

590 Huffman, G.J., A. Behrangi, D.T. Bolvin, E.J. Nelkin (2022), GPCP Version 3.2 Satellite-Gauge
591 (SG) Combined Precipitation Data Set, Edited by Huffman, G.J., A. Behrangi, D.T. Bolvin, E.J.
592 Nelkin, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services
593 Center (GES DISC), Accessed: [05/30/2024], 10.5067/MEASURES/GPCP/DATA304

594 Jakob, C., Singh, M. S., & Jungandreas, L. (2019). Radiative convective equilibrium and
595 organized convection: An observational perspective. *Journal of Geophysical Research:*
596 *Atmospheres*, 124(10), 5418-5430.

597 Kato, S., F. G. Rose, D. A. Rutan, T. E. Thorsen, N. G. Loeb, D. R. Doelling, X. Huang, W. L.
598 Smith, W. Su, & Ham, S.-H. (2018) Surface irradiances of Edition 4.0 Clouds and the Earth's
599 Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data product. *Journal of*
600 *Climate*, 31, 4501–4527, doi:10.1175/JCLI-D-17-0523.1.

601 L'Ecuyer, T. S., Hang, Y., Matus, A. V., & Wang, Z. (2019). Reassessing the effect of cloud type
602 on Earth's energy balance in the age of active spaceborne observations. Part I: Top of atmosphere
603 and surface. *Journal of Climate*, 32(19), 6197-6217.

604 Lee, D., Oreopoulos, L., & Cho, N. (2020). An evaluation of clouds and radiation in a large-scale
605 atmospheric model using a cloud vertical structure classification. *Geoscientific Model*
606 *Development*, 13(2), 673-684.

607 Loeb, N. G., D. R. Doelling, H. Wang, W. Su, C. Nguyen, J. G. Corbett, L. Liang, C. Mitrescu,
608 F. G. Rose, & Kato, S. (2018). Clouds and the Earth's Radiant Energy System (CERES) Energy
609 Balanced and Filled (EBAF) top-of-atmosphere (TOA) edition-4.0 data product. *Journal of*
610 *Climate*, 31(2), 895–918.

611 Matus, A. V., & L'Ecuyer, T. S. (2017). The role of cloud phase in Earth's radiation budget.
612 *Journal of Geophysical Research: Atmospheres*, 122(5), 2559-2578.

613 NASA/LARC/SD/ASDC. (2023). CERES Energy Balanced and Filled (EBAF) TOA and
614 Surface Monthly means data in netCDF Edition 4.2 [Data set]. NASA Langley Atmospheric
615 Science Data Center DAAC. Retrieved from [https://doi.org/10.5067/TERRA-AQUA-](https://doi.org/10.5067/TERRA-AQUA-NOAA20/CERES/EBAF_L3B004.2)
616 NOAA20/CERES/EBAF_L3B004.2

617 Oreopoulos, L., Cho, N., Lee, D., Kato, S., & Huffman, G. J. (2014). An examination of the
618 nature of global MODIS cloud regimes. *Journal of Geophysical Research: Atmospheres*,
619 *119*(13), 8362-8383.

620 Oreopoulos, L., Cho, N., Lee, D., & Kato, S. (2016). Radiative effects of global MODIS cloud
621 regimes. *Journal of Geophysical Research: Atmospheres*, *121*(5), 2299-2317.

622 Oreopoulos, L., Cho, N., & Lee, D. (2017). New insights about cloud vertical structure from
623 CloudSat and CALIPSO observations. *Journal of Geophysical Research: Atmospheres*, *122*(17),
624 9280-9300.

625 Räisänen, P., Barker, H. W., Khairoutdinov, M. F., Li, J., & Randall, D. A. (2004). Stochastic
626 generation of subgrid-scale cloudy columns for large-scale models. *Quarterly Journal of the*
627 *Royal Meteorological Society*, *130*(601), 2047-2067.

628 Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., et al.
629 (2008). The GEOS-5 data assimilation system—Documentation of versions 5.0.1, 5.1.0, and
630 5.2.0. *NASA Technical Reports Series on Global Modeling and Data Assimilation*, NASA/TM-
631 2008-104606, Vol.27, 92 pp, *NASA Technical reports*.

632 Rossow, W. B., & Schiffer, R. A. (1999). Advances in understanding clouds from ISCCP.
633 *Bulletin of the American Meteorological Society*, *80*(11), 2261-2288.

634 Sassen, K., Wang, Z. & Liu, D.(2008). Global distribution of cirrus clouds from
635 CloudSat/Cloud-Aerosol lidar and infrared pathfinder satellite observations (CALIPSO)
636 measurements. *Journal of Geophysical Research: Atmospheres*, *113*(D8).

637 Stephens, G.L., Vane, D.G., Boain, R.J., Mace, G.G., Sassen, K., Wang, Z., Illingworth, A.J.,
638 O'connor, E.J., Rossow, W.B., Durden, S.L. & Miller, S.D. (2002). The CloudSat mission and
639 the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin*
640 *of the American Meteorological Society*, 83(12), pp.1771-1790.

641 Winker, D.M., Vaughan, M.A., Omar, A., Hu, Y., Powell, K.A., Liu, Z., Hunt, W.H. & Young,
642 S.A.(2009). Overview of the CALIPSO mission and CALIOP data processing
643 algorithms. *Journal of Atmospheric and Oceanic Technology*, 26(11), pp.2310-2323.