Response to Reviewer 1

This manuscript delivers an elegant and much-needed synthesis of how and when snapshot observations of clouds can justifiably be interpreted as proxies for time-resolved processes. Its intellectual clarity and breadth of examples promise to influence both observationalists and modelers.

Inferring process from snapshots of cloud systems is a thought-provoking synthesis: it distils scattered intuitions about when spatial statistics can stand in for temporal evolution, and it offers a clear vocabulary (ergodicity, D-number, Type 1 vs. Type 2) that would help future cloud research.

In my opinion, apart from a few minor modifications, the paper should be published. It is an "opinion" paper, and as such, it highlights an important question that is highly relevant to cloud/rain/aerosol/climate research in a rather qualitative manner.

We thank the reviewer for their positive comments!

Minor comments:

1) In general, I miss a consideration of the variance of the explored processes. In all of the examples, the mapping is not one-to-one. The relevant variables are represented by a distribution (r_e or LWP). When sampling the state space, one must be sure that the variability is covered.

The variance is not necessarily a reflection of the many states. It could reflect variations around a given state. The text in Line 141, for example. The step of translating the satellite snapshot into a few hours of observation is critical. It distills the essence of the paper and should be better explained. We know that the r_e slices (per T, Z, or P) can be highly variable. I miss a discussion of the need for fully covering the statistical variance.

The question is a good one. We will distinguish between three different sources of variance: i) variance caused by varying meteorological conditions ("varying rules of the game", Type 2; ii) variance in cloud top $r_e(z)$ across the domain, which is required for the cloud-scale profiling to be effective; and (iii) variance rooted in cloud-scale fluctuations, especially from turbulent mixing, which is not the scale of interest in the problem posed here. For our targeted cloud-scale process timescale, in order to capture this variance and obtain sufficiently precise mean values, one has to sample a sufficient number of clouds over the range of z of interest. The sample size will vary with altitude, there being far more small/shallow clouds than large/deep clouds. To improve statistics, one would like to have a large enough domain with the proviso that there is no significant change in the meteorological state across the domain.

The LES results represent a very large statistical sample of clouds and remarkably show a tight $r_e(z)$.

In the interests of keeping the manuscript conceptual, we have not attempted to quantify variance that might derive from process changes (e.g. fluctuations in entrainment), meteorological changes (large scale, or perhaps associated with clouds themselves affecting their local environment), or from sampling statistics. We have, however, added clarifying comments to the text, particularly regarding the scale separation inherent to the different examples.

The text has been modified on lines 152-157:

When applying this method to satellite-based observations one has to sample a sufficient number of clouds over the range of heights of interest. The sample size will vary with altitude, there being far more small/shallow clouds than large/deep clouds. To improve statistics, one has to balance the desire for a large domain with the risk of gradients in the meteorological state across the domain. We have not attempted to quantify the variance associated with small-scale process changes (e.g. fluctuations in entrainment-mixing), meteorological changes (large scale, or perhaps associated with clouds themselves affecting their local environment), or sampling statistics.

2) On the same note, what is a sufficiently large snapshot? How to scale the spatial length of it to the time it covers? What is the right mapping constant? Is it advection?

As noted above, a sufficiently large snapshot would be one in which one has a sufficient number of clouds over the range of z of interest the domain is large enough that there is no significant change in the meteorological state across the domain. The overall time duration for this case is connected to the range of cloud heights that the cloud system spans rather than an advective timescale.

See text on lines 161-163: "*The relevant process timescale τ_{proc} is associated with the evolution of an individual cloud or, more specifically, the height increment between the individual clouds, ordered by height (say 100 m). For typical updrafts on the order of meters per second τ_{proc} would be on the order of minutes.*"

3) Line 237: "Because the data derive from many different conditions, the observation timescale t_obs is on the order of many days …" Please explain why, when mixing many observations of different thermodynamical states, we can scale the observation time to days? I guess that by doing so, we average over many thermodynamic scenarios? Again, in this case, the variance of the timescale is important.

First, increasing the number of days included in the analysis naturally corresponds to a large observation timescale because many different cloud and environmental states are being observed. However, the crux here is that when aggregating data from many days one encompasses many cloud and environmental conditions. Individual systems therefore experience different environments and therefore respond differently, obfuscating the underlying processes of interest, leading to a Type-2 designation.

See text on lines 253-257:
The large number of days included in the analysis corresponds to a large observation timescale because many different environmental states--which translate to different cloud dynamical states--are being observed. The changing dynamical state is important because it brings in processes such as droplet nucleation, in-cloud residence times, and fallout, which could obscure the collision-coalescence process.

4) What is the meaning of ergodicity in the case of averaging many states? I think that discussing it in the introduction would be beneficial to the general message of the paper. Can any system be averaged such that taking enough samples to cover the state distribution will yield an ergodic system?

Averaging over many systems that evolve under different states can be considered equivalent to a very long timeseries that explores all these states over time. Averaging over many states thus increases the observational timescale and accordingly means that slower processes are sampled. As one does so, one increases the range of spatial/temporal scales that are included in the analysis, leading to less and less useful results. For example, if one were to average $r_e(z)$ over a very large area—say global, just to make a point—one would obtain an $r_e(z)$ profile that is representative of global conditions but has very little relevance to microphysics. In other words, with progressive averaging one loses the timescale of interest until the result isn't very useful because it encompasses a range of scales large enough that one cannot learn anything specific about the target (small-scale) process.

Regarding the latter part of this question, the answer is 'no' because the rules keep changing and the range of processes keeps increasing.