Large Ensembles Workshop Breakout Questions

Day 1 (Value of LEs for attribution, model evaluation, and unraveling projection uncertainty)

(1) In your subfield/specialty, what approaches are used to identify anthropogenic influences in the observational record? What are the main hurdles to doing this and how are they best overcome? In what aspects of the observed record is there confidence in an anthropogenic influence and its magnitude and in what aspects are there not?

(2) How are model mean state, trends, and /or variability typically validated in your particular subfield/specialty? Can this be improved upon with large ensembles and statistical methods applied to observations (i.e., an "observational large ensemble"), and if so, how?

(3) Within your subfield, what are the largest contributors to projection uncertainty (model structure, internal variability, forcing scenario) and what is the potential for narrowing them? How important are Large Ensembles for doing this?

Day 2 (LE design considerations, new research directions)

(1) There are multiple types of ensembles that can be run using climate models. These include:
a. *Initial condition* ensembles, in which the ensemble members differ only by their initial conditions;
b. *Physics ensembles*, in which ensemble members use the same basic model but alter physics parameters (this would include stochastic physics ensembles);
c. *Multi-model ensembles*, in which different models are run once with similar forcings and are combined to form an ensemble.
d. *Ensembles with single forcings*.
With finite computer resources there are always tradeoffs. If there were an internationally coordinated 'Large Ensemble' program, what type of ensembles should be the focus? What topics/questions would be key to address? For what types of questions are large ensembles necessary, and for what questions are other approaches potentially more desirable, such as analyses derived from long control simulations?

(2) There are many attributes of the models used in an ensemble, including resolution, complexity (types of processes included, such as biogeochemical processes), number of ensemble members, and duration of simulations. If you had 100 units of computing to spend to address your science question(s), how would you allocate those units in terms of these attributes, and why?

(3) For an initial condition ensemble, there are multiple ways to initialize individual members. One is to have members differ by bitwise perturbations to the initial conditions, with the differences growing quickly. Another is by using initial conditions taken from different points in a long control integration. Given these choices and their implications, do you have a preferred method for creating initial conditions for ensemble members?

(4) How do we decide how many ensemble members are needed? In this context, what defines "large"?

(5) What are 3 new research directions (within your subfield) that could benefit from initial-condition (or other types of) Large Ensembles? Do these require new models and/or additional model output?