## Defining a common reliability standard

1. In order for the CPUC and parties to compare different models and their results, we would like to develop some standardization of outputs between the models. What should the first step be?

2. What reliability standard should each model try to achieve?  Is the reliability level an input or an output to the model?

   a. A common standard is referred to as "1 day in 10 years" or "1 event in 10 years".  How do we define precisely what these mean (e.g. in terms of duration)?

   b. What are the different types of events that need to be measured?  Is an event the same as an outage or is an outage just one type of event?  What are event/outage types? (e.g. capacity, upward ramp, overgeneration, etc.)

   c. How should we distinguish and report intra-hour flexibility shortfall, as opposed to hourly or multi-hour ramping (up or down) shortfall?  How does unit commitment logic and the assumed fleet affect this?

3. What are the thresholds to trigger an event (e.g. supply is only 3% above demand)?  How do such thresholds relate to existing reliability standards?

## Standardizing calculation of reliability need

4. How should models calculate and characterize system shortages or deficiencies relative to a reliability requirement?

   a. What are the parameters necessary to adequately describe and quantify the reliability need (is it just MW, MW with what characteristics, MW of a standard proxy resource, and what is the definition of the proxy resource)?

5. How can we standardize the handling or curtailment/overgeneration across different models? Is it possible to have common assumptions about how much/how long curtailment is allowed? How should models quantify downward flexibility need?

6. How should models address imports as a resource to support load or exports as a sink for "over-generation"?

## Methodological differences between models that impact comparability

7. Some models draw from a single day (as opposed to whole year simulations) to reflect a month or a season (what month ranges are seasons?). What are the additional assumptions necessary to interpret results? For example, how do we translate from events per spring day to something like hours per year?

8. How does model X calculate capacity needed for flexibility? For example, does it assure that traditional measures of reliability are satisfied, and then add resources for flexibility, or does it satisfy both simultaneously? Are there methodological differences between how different models calculate shortfalls relative to a reliability standard that impacts the comparison between different models?

9. Production simulation models generally operate as though all resources are available for perfect optimization. Various real world constraints (e.g. committing capacity to satisfy local requirements, self-scheduling, etc.) may not be reflected in their projections. How important is it to consider that some models assume perfect foresight and optimized unit commitment and dispatch, while some models may account for real world constraints or uncertainty?