Joint Intervenor Multi-Attribute Model
Defining and Evaluating the "Test-Drive"

Charles D. Feinstein, PhD
Jonathan A. Lesser, PhD

The August 16, 2016 CPUC Order requires a "test-drive" of the Joint Intervenor's multi-attribute modeling approach for risk management.[1] The CPUC Order specifies that the modeling approach will be tested "using a small set of detailed test problems (at least five) which are common across more than one utility."[2] This paper provides a suggested roadmap for readers of how the test-drive could be performed, including: (a) the steps in the analytical process, (b) what will be required from the utilities to enable a successful test-drive, (c) how the results of the test-drive can be evaluated and compared with the structure and outputs of the current utility methodology, and (d) selecting the test problems to be evaluated.[3]

I. THE TEST-DRIVE ANALYSIS

We believe the test-drive results and analysis should be considered *illustrative* only. That is, we do not believe it appropriate to require the utilities to adopt the specific recommendations of the test-drive as utilities prepare future filings before the CPUC, even if the test-drive is considered to be successful by the CPUC. (Of course, the utilities may choose to rely on the test-drive analysis results, or an update of those results.)

The test-drive will require completion of five distinct steps. These are as follows:

Step 1: Identify the value attributes, attribute scales, and attribute weights for the multi-attribute value function that will enable us to measure risk and risk reduction associated with specific mitigation actions. This exercise need only be done once and will apply to all of the test problems.

Step 2: Develop the condition-dependent hazard rates for each asset type being evaluated.

---

[1] CPUC Order, p. 2. The model is also called the "EPRI Model" in the CPUC Order.

[2] CPUC Order, Ordering Par. 1(b).

[3] For additional detail, see the Joint Intervenor Whitepaper of January 28, 2016 and the accompanying Technical Appendix.

Step 3: Develop probability distributions for the consequences of failure (CoF) for each asset type, which reflect the changes in the levels of the identified attributes (from Step 1) that result from asset failure.

Step 4: Identify the specific risk mitigation actions to be modeled for each asset type. For each mitigation action, specify the resulting condition-dependent hazard rates (LoF) and the probability distributions for the changes in attribute levels that are the consequences of failure (CoF) for each asset type. (Note that a mitigation action may change either LoF or CoF or both.)

Step 5: Based on the alternative mitigation actions identified, assume an illustrative budget constraint and (i) rank the mitigation actions *individually* in terms of risk-reduction per dollar spent and (ii) select a portfolio of risk mitigation actions given a budget constraint.

Below, we discuss each of these steps in more detail.

A. Step 1: Attribute Identification and Weighting

Because the test problems are designed to apply to more than one utility and to reduce the test-drive workload, there should be a single set of value attributes, attribute scales, and attribute weights that measure the consequences of failure (CoF) for all problems in the test-drive analysis. (However, it will be possible to change the attribute weights as part of the test-drive analysis.)

We recommend that SED staff work with the utilities to select a group of utility SMEs who can identify the value attributes, specify the attribute scales, and determine a consistent set of attribute weights. This group could meet in a working group that we will lead. (We have done exactly this process many times in our practice.) We suggest that SED staff participate in the attribute identification/scaling/weighting working group. In addition, the CPUC may wish to invite other intervenors to participate.

The attribute selection process begins with high-level attributes (e.g. “safety”, “reliability”, etc.) and then identifies a complete set of value-independent, *measurable* attributes. For example, reliability may be measured by hours of lost load or the dollar impacts of that lost load. Safety may be measured in terms of potential deaths/injuries of employees and the general public. (We hasten to say that such measurements involve specification of uncertainty.)

Once the final set of measurable attributes is determined, each attribute will be scaled. The scaling function for each attribute specifies the relative value of a change in attribute level. The scaling function ranges from 0 to 100 over the possible attribute levels (from best to worst). The scaling functions will be determined in consultation with the SMEs at the working group session.

Scaling is important because it permits us to specify a *consistent* set of attribute weights, as we explained in our Whitepaper.[4] (Alternatively, because the CPUC says safety must be the most heavily weighted attribute, we could agree for the test-drive to set the safety weight in advance and calculate the remaining attribute weights in a consistent manner.)

The specific tasks to be completed in the working group session are as follows:

1-1. Define the attributes (e.g., safety, reliability, environmental quality, etc.).

1-2. Create an attribute structure such that measurable attributes are at the bottom. In other words, we typically start with the high-level attributes (e.g., safety, reliability, etc.) and provide additional details about what comprises the high-level attribute and how the high-level attribute is actually observed and measured. The specification of each attribute ends when we reach a mutually agreed set of *measurable* attributes that completely describe each high-level attribute. (For example, reliability may be measured in terms of commercial customer reliability and non-commercial customer reliability. Commercial customer reliability may be measured by the annual rate of service interruptions and the duration of each interruption. Non-commercial customer reliability may be measured by the number of customer-minutes lost per year. These are the *natural units* of the attributes; see 1-3, below.)

1-3. Specify the natural units for each attribute and the ranges of those natural units. Natural units are those that one would commonly use to express an observed value for a given attribute. For example, the natural unit for expressing safety, as discussed above, may be the number of deaths and injuries to employees and the general public. An environmental quality attribute might include acres of land burned, and so forth. The range of natural units for each attribute identifies the best and worst observable levels in natural units for each attribute. The best level for the natural unit of an attribute is typically, but not always, the observed attribute level when no failure event has occurred. The worst level for the natural unit of an attribute is typically, but not always, the observed attribute level when the most consequential failure event has occurred.

1-4. Specify scaled units for each attribute (these are the attribute scaling functions, internal to each attribute, with each scale range from 0 to 100). The scaled units determine the value of changing attribute levels for a given attribute. In some cases, such as attributes measured in dollars, the scales are linear. In other cases, the scales will not be linear. (For example, reliability may be viewed as having consequences that increase non-linearly. Thus, a 10-hour outage might have consequences that are worse than twice the value of a 5-hour outage.) The scaling function associates each level of an attribute in natural units with a scaled value from 0 to 100. The scaled value of a change in attribute

---

[4] Joint Intervenor Whitepaper, pp. 19 – 22.

level between two levels of natural units is the difference between the scaled values. Thus, changing from worst level to best level has a scaled value of 100 because 100 – 0 = 100. The way the scaled value is interpreted is that all changes in attribute levels that have the same difference in scaled units are equally valuable for a given attribute. (That does not apply when comparing the value of scaled changes for different attributes unless the attribute weights are equal.) Also, cardinal comparisons apply: a change in attribute levels that has a scaled value of 5 has one-twentieth the value of changing from the worst level to the best level (because 5/100 = 1/20).

1-5. Derive the attribute weights. The attribute weights are determined by a collection of tradeoffs made using the attribute structure. The tradeoffs can be pairwise comparisons or a direct assignment of relative values or some combination. Although the calculation of the attribute weights is a simple algebraic exercise, the collection of tradeoffs will be made in the working group session by the SMEs.

As should be clear, these first five tasks *require* utility participation (as well as SED staff and possibly other intervenors, if the CPUC wishes.) Our role solely will be to *facilitate* the discussions among the group and enforce the logical constraints imposed by the attribute structure. We do not, and cannot, determine the attributes, the attribute scales, and the attribute weights; that must be the role of utility SMEs, SED staff, and other intervenors.

1-6. Evaluate and revise, if necessary, the attribute definitions, structure, ranges, scales, or weights. This final step is required to ensure that the working group session participants are all satisfied with the results. For example, we may calculate an attribute weight for reliability that participants believe to be too low (or too high). They might decide that the worst case outcome for reliability was overly pessimistic, or not pessimistic enough. This last step is designed as a check to ensure all participants agree with the attribute structure and the multi-attribute value function before proceeding to the rest of the analysis.

Steps 1-1 - 1-6 provide a complete specification of the multi-attribute value (utility) function.

B. Step 2: Develop Condition-Dependent Hazard Rates

Condition-dependent hazard rates are what we use to measure LoF. These hazard rates reflect the fact that the likelihood an asset will fail (e.g., a segment of vintage pipe, a wood utility pole, etc.) in a given time frame typically depends on the asset's condition. For example, pipe that is corroded is more likely to fail than pipe that has no corrosion. (In some cases, LoF does not depend on asset condition. This is just a special, simpler case, in which the condition-dependent hazard rates are the same for all possible asset conditions.)

As with Step 1, developing condition-dependent hazard rates will also require our working with the utility SMEs and evaluating available data. This step will also consider interactive threats, to the extent they are identified by SMEs, as well as cascading (dependent) failures (i.e., failure of asset A causes failure of asset B, which causes failure of asset C, etc.), such as multiple transformer failures. For example, if vintage pipe is selected as one of the test-drive analyses, the interactive threats to such pipe include earthquakes, corrosion, and manufacturing defects (and any other factors the SMEs identify). The conditional hazard rates vary depending on the arrival of each of the interactive threats.

We recommend that SED staff work with the utilities to select a group of utility SMEs who can identify the asset condition, the condition-dependent hazard rates, and the scope and accuracy of asset condition testing. This group will meet in a working group session that we will lead. The working group may require more than one session. (We have done exactly this many times in our practice.) We suggest that SED staff participate in the working group session. In addition, the CPUC may wish other intervenors to participate.

The specific tasks involved are as follows:

2-1. Define asset conditions. Asset condition is typically a set of discrete states in which an asset can be characterized as being in. We often begin with three mutually exclusive, collectively exhaustive states, such as “good,” “fair,” and “poor.” (”Mutually exclusive and collectively exhaustive” conditions are defined such that an asset will be in one of the conditions at any time and that there are no other possible conditions. In other words, if we define three conditions (e.g., “good,” “fair,” and “poor”), then an asset must be in exactly one of those three conditions at any given time. Thus, for example, an asset cannot both be in “good” and “fair” condition, nor can the asset be in any condition other than “good,” “fair,” or “poor”.) The hazard rate must be specified for each condition. In order to define such condition-dependent hazard rates for the test-drive, we will ask the appropriate SMEs to identify the asset conditions. The essential ideas in this identification are that different asset conditions: (a) imply different hazard rates, (b) vary over time (see B2, below), and (c) are revealed by testing the asset (see B4 below).

2-2. Specify the dynamic behavior of asset condition. In the future, at any point in time, the asset condition is uncertain. Therefore, at each point in time, we find the probability that the asset is in a given condition.

Because the asset condition varies over time, the probability that the asset is in any one of the possible conditions also changes over time. Thus, at any future time, we find the probability that the asset is in any one of the possible conditions. We then forecast how asset condition will change over time by specifying how condition changes periodically, say annually. In other words, at the working group session, working with the SMEs, in this step we determine the likelihood that an asset's condition changes from, say, "good" to "fair" or "poor" over a single period. Absent replacement, an asset's condition is typically assumed not to improve over time. For example, if a pipe segment is corroded today and judged to be in "poor" condition, the corrosion will not suddenly disappear and restore the pipe segment to "good" condition. We build a mathematical model that computes the probability distribution on asset condition at any time in the future. A Markov model is a popular method for describing the dynamic probabilistic behavior of asset condition. At the working group session, we will work with the SMEs to specify the parameters of the model.

The parameters of a Markov model can be presented in a matrix, as the illustrative example below shows. In this example, the probability that an asset in good condition at the start of the period (say, one year), still will be in good condition at the end of the year is 0.60. The probability it will be in fair condition is 0.30, and the probability it will be in poor condition is 0.10. These are known as *transition* probabilities. (Note that the sum of the transition probabilities across any row of the matrix must be 1.00.)

		Probability Asset will be in Stated Condition at End of Period		
		Good	Fair	Poor
Condition of Asset at Start of Period	Good	0.60	0.30	0.10
	Fair	0.00	0.50	0.50
	Poor	0.00	0.00	1.00

2-3. Test the Markov model and revise parameter estimates as needed. In this step, we will work with the SMEs to determine whether their initial estimates of how asset condition changes over time are reasonable or require further revision. This can be done between working group sessions.

2-4. Identify and describe the tests available to observe the condition of an asset. Asset condition is uncertain, but tests exist that can resolve the uncertainty somewhat, if not completely. For each asset, the applicable tests will be identified, their costs will be determined, and their accuracy will be described. Test accuracy is usually described in terms of the likelihoods that a test reports either accurately (a test says that the condition is "fair" when the true condition is "fair") and inaccurately (a test says that the condition is "good" when the true condition is "fair"). The results of a test applied to an asset at any time modify the probability distribution on asset condition. Therefore, a test may have value because the test results can motivate useful changes in risk mitigation strategy as an asset becomes more or less risky to operate. However, the value of any information provided by the test must be sufficient to justify the cost of the test.

2-5. Specify the condition-dependent hazard rates. The hazard rate is the probability that an asset will fail before the end of a time interval, given that the asset had not failed at the beginning of the time interval.[5] The hazard rate is related to the survivor rate of an asset. The hazard rate depends on the condition of the asset. Therefore, as the condition of the asset changes over time (described in steps 2-1 - 2-3, above), the hazard rate of the asset changes over time. Notice that the condition of an asset can change without the occurrence of asset failure. What happens instead is that, as the asset's condition changes, the likelihood of failure changes. Indeed, it is possible that a pipe segment could be in good condition today and still fail tomorrow, while another segment could be in poor condition and not have failed a year from now.

The condition dependent hazard rates will be specified by the SMEs at a working group session. In most cases, the condition dependent hazard rates are found by a combination of data analysis and expert judgment.

2-6. Test unconditional hazard rates and revise the parameter estimates as needed. The forecast of the change in asset condition over time (found in steps 2-1 - 2-3) can be combined with the condition-dependent hazard rates (found in step 2-5) to estimate the probability that a randomly selected asset will fail at any point in time. (In other words, by combining the probability that a randomly chosen asset is in a given condition and the probability that an asset in a given condition will fail we derive the unconditional probability that this randomly selected asset will fail.) This estimate can be applied to the asset inventory to forecast the expected number of asset failures in the future. That forecast will be reviewed by the SMEs and the estimates of the parameters of the asset condition model and of the condition-dependent hazard rates can be modified, as needed.

2-7. Define threats and specify the effects of threats on condition-dependent hazard rates. At the next working group session, we will work with the SMEs and identify the threats (external events that are different than asset condition) that can affect the probability that a given asset will fail. For example, pipe in poor condition, perhaps because of corrosion, might have a failure probability of 20% in the next year. However, if an earthquake occurs, the condition-dependent hazard rate might increase to 90% for pipe in poor condition. The purpose of this step is to identify all of the specific threats that will affect the probability of asset failure and measure the effect of the threat on the condition-dependent hazard rate. One way to measure the effects is to specify a set of hazard rate multipliers.

---

[5] See Joint Intervenor White Paper, p.8. n.10, ff.

2-8. Specify interactive threats and effect of interactions on condition-dependent hazard rates. Threats can also interact, which means they may arrive at the same time and jointly affect the condition-dependent hazard rate. For example, wooden utility poles that have suffered a third-party contact (perhaps a car crash) may be more susceptible to failure because of fire than poles with no such contact. Third-party contact and fire are two threats that interact and, in doing so, affect the condition-dependent hazard rate. Identifying important interactions will be addressed at a working group session.

2-9. Specify arrival rates of threats and specify conditional dependence among non-independent threats. The arrival rate of a threat is a parameter that can be used to find the probability that the threat occurs in a given period, typically one year. As discussed in the Whitepaper, we convert arrival rates into probabilities using the Poisson distribution.[6] The SMEs will specify the arrival rates at a working group session. Further, the arrival rate of a threat may depend on the occurrence of another threat. The arrival rates will be addressed at a working group session.

2-10. Test unconditional hazard rates and revise parameter estimates as needed. Similarly to step 2-6, the forecast of the probability that a randomly selected asset will fail at any future time will, after the results of steps 2-7 – 2-9, include the effects of threats on hazard. This estimate can be applied to the asset inventory to forecast the expected number of asset failures in the future. That forecast will be reviewed by the SMEs and the estimates of the parameters describing the arrival and interactions of the threats can be modified as needed.

2-11. Specify failure dependencies. In some cases, failure of any asset changes the probability that another asset in the asset inventory will fail. Such *dependent* failures may be an important aspect of risk mitigation strategy. At another working group session, the important dependencies will be defined and the dependencies will be measured. One approach to specifying dependencies is to set dependent hazard rates.

2-12. Identify consequential non-asset related events. There may be events that have consequences that can be mitigated without considering the behavior of assets. For example, some aspects of cybersecurity have adverse effects that do not entail asset failure. All such events will be identified at a working group session. The arrival rates for such events will be specified by SMEs.

---

[6] Joint Intervenor Whitepaper, p.14, fn.13.

Steps 2-1 - 2-11 provide a complete description of the LoF for any asset. Step 2-12 provides the LoF (which we might more accurately call "LoO," the likelihood of occurrence) for risks that are not asset-related.

C. Step 3: Develop Probability Distributions for Consequences of Asset Failure

Once the multi-attribute value function (attributes, attribute scales, and attribute weights) and the condition-dependent hazard rates are specified, the next step is to work with the various SMEs to determine the consequences of failure (CoF). The consequences of asset failure (as well as those of non-asset-related events) are measured by the *changes* in the attribute levels that are caused by the failure or other event. (If an asset fails, but none of the attribute levels changes, then there are no consequences of failure.) Our preferred approach is to recognize that the CoF are uncertain and therefore we measure CoF by the probability distributions of attribute levels caused by the failure or other event.

We recommend that SED staff work with the utilities to select a group of utility SME's who can identify the consequences of the occurrence of asset failure or the arrival of non-asset-related events in terms of the attribute levels in the value model found in Step 1. This group will meet in a working group session that we will lead. The working group may require more than one session. (We have done exactly this many times in our practice.) We suggest that SED staff participate in the working group. In addition, the CPUC may wish other intervenors to participate.

The specific tasks involved are as follows:

3-1. Specify the consequences of failure in terms of the changes in attribute levels. SMEs will provide either point estimates or probability distributions. One way to specify a probability distribution is to provide the 10-50-90 ranges that measure the tenth, fiftieth, and ninetieth percentiles of the distribution. In general, the consequences of an asset failure won't be known precisely. For example, if a transformer catches fire or explodes, the resulting loss of electric service likely will depend on the location of the transformer and how quickly service can be restored. That is why probability distributions on CoF are both necessary and natural.

3-2. Specify the effect of multiple failures on the changes in attribute levels associate with the occurrence of the failure event. For example, if transformer A fails, it may cause transformer B to fail. The LoF for such dependent failures was addressed in task B10, above. The CoF for dependent failures will be addressed by SMEs in this task. The simplest approach is to assume that the CoF is additive for multiple failures. However, that need not be true. For example, it is possible that the consequences of the joint failure of two transformers could be less than the sum of the individual failure consequences;

i.e., CoF(A + B) < CoF(A) + CoF(B). Alternatively, the consequences of failure of both could be greater than the sum of the individual failure consequences; i.e., CoF(A + B) > CoF(A) + CoF(B). SMEs will specify the interactive effects.

Steps 3-1 - 3-2 provide a complete description of the consequences of failure (CoF).

D. Step 4: Identify Alternative Mitigation Measures

The next step in the test-drive process will be for SME's to identify the different mitigation measures available for each asset class. For each mitigation alternative, SMEs will be asked: (i) how the mitigation affects the LoF and (ii) how it affects CoF. Again, these SME estimates can be point estimates or 10-50-90 ranges.

We will also need to know if mitigation measures are mutually exclusive (i.e., if mitigation measure A is applied, then mitigation measure B cannot be applied, etc.) or whether certain mitigation measures can be combined and, if so, how LoF and CoF are affected by the combined measure. In most cases, the effect on CoF of a combined measure is not the sum of the effects of each measure, while the effects on LoF will depend on the nature of the threats that influence failure.

It is also possible that a mitigation measure focused on one type of equipment may affect the failure rates for other equipment. For example, transformers attached to wooden utility poles are more likely to fail if the wooden pole catches fire or falls down in high winds. Replacing a wooden utility pole with a metal one not only mitigates the failure risk for the pole, but also reduces the failure risk for the transformer. Depending on how comprehensive the test-drive is required to be, these sorts of joint effects can be evaluated.

We recommend that SED staff work with the utilities to select a group of utility SME's who can identify the effects of the mitigation measures on the likelihood of failure (LoF) and the consequences of failure (CoF). This group will meet in a working group session that we will lead. The working group may require more than one session. (We have done exactly this many times in our practice.) We suggest that SED staff participate in the working group. In addition, the CPUC may wish other intervenors to participate.

The specific tasks involved are as follows:

4-1. Identify the alternative risk mitigation measures. Specify the cost of undertaking each risk mitigation measure. Assuming we are not performing a dynamic analysis in the test-drive, this cost should be the present value of all cash flows associated with implementing the risk mitigation measure.

4-2. Express the consequences of applying the risk mitigation measures both before and after a failure, including the effect of the risk mitigation measures on multiple failures. The consequences of applying the risk mitigation measure should be expressed in terms of changes to LoF or CoF or both. That is, a risk mitigation measure may result in failure being less likely or the consequences of a failure less costly or both.

Steps 4-1 - 4- 2 provide a complete description of the effects of any risk mitigation measure on both likelihood of failure (LoF) and consequences of failure (CoF).

Steps 1 - 4 complete the work of the SMEs and other utility experts. The result of these steps is a complete set of inputs to the Joint Intervenors risk management methodology. The next step converts the inputs into the methodology outputs, the mitigation measure rankings and selections based on risk reduction.

E. Step 5: Analysis and Ranking of Risk Mitigation Alternatives

The final step in the test-drive process will be ranking the alternative mitigation measures and costs, in terms of risk reduction per dollar spent. Depending on the scope of the test-drive, we may also provide risk mitigation measure portfolios subject to applicable constraints. (We would have to specify the amount of constrained resources consumed by each risk mitigation measure.)

The risk-reduction-per-dollar rankings do not change in the presence of budget constraints. Because of that, a common approach to selecting ranked alternatives in the presence of a single budget constraint is to select the alternatives according to the ranking and stop at the point the budget is exceeded. That selection process is called a heuristic and is not generally optimal.

In fact, the optimal constrained risk mitigation measure portfolios need not be comprised of the risk mitigation measures that follow the ranking even if there is only a single budget constraint. In other words, if risk measures are ranked A – B – C – D – E – F, in order of descending risk reduction per dollar, the optimal choice of measures constrained by budget may exclude, say, measure C, and include, say, measure F, even though the latter has a lower risk reduction per dollar value. In our methodology, the optimal set of risk mitigation measures can be determined using integer programming methods (including ones available in Microsoft Excel).

Finally, we recommend that, in addition to developing separate rankings for each of the test-drive problems, we rank the mitigation measures applied to each problem jointly (and create portfolios of risk mitigation measures depending on the applicable budget constraints). In this way, we can demonstrate how rankings can change when various risky assets are considered jointly.

The specific tasks involved are as follows:

5-1. Compute the risk-reduction for each mitigation measure. The risk reduction is simply $(\text{LoF } x \text{ CoF})_{\text{Before}} - (\text{LoF } x \text{ CoF})_{\text{After}}$.

5-2. Rank the alternatives and report the results.

5-3. Impose constraints as applicable and find the portfolios of risk mitigation measures that maximize the risk reduction achieved subject to the applicable constraints.

5-4. Sensitivity studies. Depending on the scope of the test-drive, we can examine how the rankings and portfolios change when the underlying assumptions change. For example, suppose the SMEs' LoF values for vintage pipe are increased? Does that change the selection of mitigation measures? Do we end up prioritizing vintage pipe replacement over (say) wooden pole replacement? Sensitivity studies can examine the impacts of changing assumptions and, by doing so, can identify areas where additional data collection is valuable.

Steps 5-1 - 5-4 provide a complete evaluation of the risk mitigation alternatives based on risk reduction. The complete ranking is found and optimal portfolios are determined that maximize risk reduction subject to applicable constraints.

II. EVALUATING THE SUCCESS OF THE TEST-DRIVE

After the test drive is completed, two key questions will likely arise. First, how do we evaluate whether the test-drive is successful? Second, how do we compare the test-drive results to the results arising from utility models?

A. Evaluating Test-drive Success

We propose five criteria to measure the success of the test-drive.

1. Quality of the outputs: Does the Joint Intervenor methodology provide measures of risk reduction and risk-spend efficiency? Does the methodology select portfolios of risk-reduction measures that provide the greatest risk reduction, given any constraints we incorporate?

2. Transparency: Are the methodology's inputs, outputs, and transformation of inputs into outputs clear and easily followed? Are the computations that determine the portfolios of measures from the inputs to the methodology clear?

3. Logical and Explainable: Is it logical and explainable why the methodology operates the way it does, including the operations to determine the portfolios of risk-reduction measures,

calculate the risk-reduction provided by those portfolios, and calculate the risk-spend efficiency?

4. Sensitivity analysis: Does the methodology allow us to easily evaluate how the answers change when inputs change? And, consequently, does the methodology allow us to identify areas where collecting additional data is most valuable?

5. Ease of Considering Alternatives: Does the methodology provide a straightforward means for the Commission and parties to examine alternative impacts on cost and risk reduction under alternative portfolios of mitigations?

B. Comparing the Test-drive Results with Utility Analyses

We have previously claimed that, even if the inputs were the same, the utilities' methods will provide different answers compared with the Joint Intervenors methodology. A fundamental question is how best to compare the results our methodology with the results of the utilities' own models. We believe that such a comparison should address the following questions:

(a) How do the rankings of the risk mitigation measures differ? Why?

(b) Given the same set of constraints, how do the portfolios of risk mitigation measures differ? Why?

(c) How does the total quantity of risk mitigation, based on the mitigation measures identified by the utility models, compare with the risk mitigation amount identified by the multi-attribute model?

(d) What is the total risk addressed by the risk mitigation measures in the different portfolios? Does that matter? Should risk addressed be used to determine which risk mitigation measures to adopt?

(e) Do both methodologies report consistent levels of risk before and after application of the risk mitigation measures?

One difficulty confronting such an evaluation is that several of the existing utility models do not measure risk mitigation. Instead, the utilities select a group of mitigation projects based on methodologies that are not always clear, although they have claimed that the basis of the selection is the level of risk prior to implementation of a risk mitigation measure. That is why both questions (c) and (d) above should be answered. (Of course, if the utilities have now implemented revised models that do measure risk mitigation, then we can compare the results of the test drive with the results produced by these revised utility models.)

III. SELECTING THE TEST PROBLEMS TO BE EVALUATED

The foregoing has shown that to reasonably test the operation of the Joint Intervenor approach with respect to any aspect of utility operations requires a dedication of resources by the utilities, Joint Intervenors, Commission staff, and other interested parties. For each aspect of utility operations (e.g., type of utility asset) that is considered, Steps 2-4 described in Section I above would need to be independently performed. (Note that Step 1 would be common to all aspects of utility operations that are assessed under the Joint Intervenor approach).

Accordingly, we believe that a reasonable way to approach the test drive process is to plan on studying two different aspects of utility operations – one electric and one gas – and to assess whether further areas of operations need to be studied after completing that work.

Although the final selection of test-drive problems will require discussion with all parties involved, we suggest the following as candidates for study under the Joint Intervenor approach.

1. Wooden utility poles (electric):

   Wooden utility poles are common to all of the utilities. We believe this is an appropriate test case for the methodology because of the Commission's and other intervenors' concerns about fire-related risks that can include failure of wooden poles. Moreover, utilities typically have test data regarding wooden pole condition available, which will be useful in developing the condition-dependent hazard rate functions needed for the analysis.

2. Vintage pipe (natural gas):

   The analysis of vintage pipe parallels that of underground electric cable. Because we have performed studies of the latter, we believe analyzing vintage pipe is an appropriate exercise for a test-drive.

3. Transformers (electric):

   Transmission transformers are another asset common to integrated electric utilities and, of course, transmission companies. We have applied the methodology to inventories of transmission transformers (notably for PJM Interconnection, LLC). One interesting aspect of the problem of risk management for an inventory of transmission transformers is the important role of dependent failures.

4. Breakers (electric):

The inventory of distribution system breakers is an asset common to all electric utilities. We have previously applied aspects of the methodology to an inventory of distribution system breakers for electric utilities.

5. <u>Compressors (natural gas)</u>:

   Compressors are a key component of reliable operation of natural gas transmission and distribution systems. It may be useful to address compressors as a means to assess the applicability of the methodology to gas system assets.

6. <u>Valves (natural gas)</u>:

   Valves are a key component of reliable operation of natural gas transmission and distribution systems. Moreover, the ability to close valves (whether manually or automatically) can affect the possible consequences of a pipeline rupture. Therefore, the asset behavior can influence the consequences of failure of another asset. That may provide an interesting application.

7. <u>Cybersecurity and Terrorism</u>:

   Cybersecurity and terrorism risk are not asset-specific, but the consequences of a successful breach of cybersecurity or a successful terrorist act may involve the failure or other malfunction of utility assets. If data are available (we do not know, because of issues involving national security), or if SMEs can be consulted, we could evaluate different types of risks affecting assets. We note that cybersecurity risks that do not affect utility assets (e.g., a hacker steals customers' information) can be evaluated by less complex methods. First, we need to specify the arrival rate of a successful cybersecurity attack. Then, if a specific risk doesn't affect any utility assets then the consequences of the occurrence of a successful attack can be expressed directly using the multi-attribute value function. What can be done, of course, is to evaluate the tradeoffs among devoting resources to risk mitigation actions that address such non-asset risks and devoting resources to addressing other risks that are directly related to assets.

Perhaps more interesting for our purposes is a case in which a hacker or terrorist causes specific components of the electric or natural gas system to malfunction. Here, the issue is again asset failure. Therefore, it fits into the test-drive's analytical structure and we can address such a risk by our methodology.