Opinion on
Energy Imbalance Market (EIM) Resource Sufficiency Evaluation Enhancements

James Bushnell, Member
Scott M. Harvey, Member
Benjamin F. Hobbs, Chair

Members of the Market Surveillance Committee of the California ISO

February 2, 2022

1. Introduction and Summary

The Market Surveillance Committee has been asked to comment on the proposed enhancements to the EIM resource sufficiency evaluation (RSE) process.[1] That initiative is in response to the events of August 2020,[2] and is a continuation of the ISO's evaluation of resource sufficiency evaluation in the EIM that began with the summer 2021 readiness initiative.[3]

As explained in the draft final proposal,[4] a purpose of the resource sufficiency evaluation is to evaluate whether each balancing area authority (BAA) is capable of meeting its own net load with its own available resources before allowing it to exchange power through the EIM with other BAAs. The capacity and flexible ramping tests are designed to minimize the likelihood of a BAA having to "lean" on other EIM entities, and are the focus of this opinion. In addition, there are balancing and feasibility tests that ensure, respectively, that an entity's base schedule of supply and demand are balanced and that the base schedule can feasibly flow over the network. If an entity fails either the capacity or ramping test for a particular market interval, the consequence is that EIM transfers into (or out of) the BAA are capped over the period in which the sufficiency evaluation was failed. The sufficiency evaluation was implemented because the BAAs who established the EIM viewed its purpose as the promotion of economic power exchanges in real-time, and not the enhancement of west-wide reliability by allowing balancing areas (BAs) to make up for short-run capacity deficits by importing power from BAs who are long. In short, the goal was to prevent behavior that could be interpreted as "leaning."

---

[1] D. Johnson and B. Dean, *EIM Resource Sufficiency Evaluation Enhancements, Phase 1: Revised Draft Proposal*, CAISO, Dec. 16, 2021. www.caiso.com/InitiativeDocuments/RevisedDraftFinalProposal-EIMResourceSufficiencyEvaluationEnhancements.pdf

[2] See California Independent System Operator, California Public Utilities Commission, and California Energy Commission, *Root Cause Analysis, Mid-August 2020 Extreme Heat Wave, Final Report*, January 13, 2021 www.caiso.com/Documents/Final-Root-Cause-Analysis-Mid-August-2020-Extreme-Heat-Wave.pdf

[3] *Market Enhancements for 2021 Summer Readiness: Final Proposal.* California Independent System Operator. March 19, 2021, www.caiso.com/InitiativeDocuments/FinalProposal-MarketEnhancements-Summer2021Readiness.pdf

[4] See Johnson and Dean, op. cit., for details.

This opinion is structured as follows. In Section 2, we summarize the objectives of the RSE. We then turn to discussions of several issues with the present implementation, starting with two statistical issues in Section 3, which also includes a review of the CAISO's proposed quantile regression approach. We emphasize the well-known issues associated with estimating the probabilities of extreme events using small samples.

Then in Section 4, we address expressed concerns that that the level of CAISO EIM transfer imports reflects undue leaning on the pool, and that this leaning is facilitated by the use of load conformance adjustments by CAISO operators. In that section, we discuss the EIM transfer data, and then discuss implications of consideration of load conformance adjustments in the resource sufficiency evaluation. We conclude that there no leaning on the rest of the system as a result of using load conformance adjustments in HASP and RTPD. We then examine the concern with the potential for EIM transfers to enable EIM entities to pass the flexibility test of the RSE.

Section 5 addresses issues concerning the interaction of the Hour-Ahead Scheduling Process (HASP) and the two real-time markets (real-time pre-dispatch (RTPD) and real-time dispatch (RTD)). We first discuss issues with the structure of HASP in relation to the capacity test of the RSE, focusing on how the structure of the CAISO HASP creates challenges for CAISO operators in passing that test. We then turn to a discussion of issues with the structure of RTD, and to a lesser extent FMM and the hourly short-term unit commitment (STUC), in relation to the flexibility test of the RSE.

Section 6 concerns whether and how off-line capacity that could be committed in STUC should be counted in the RSE's capacity test. We point out two key considerations that have to be balanced. One is that when there is capacity that cannot come on-line in time to provide the needed capacity in the period being tested, that capacity is unable to meet capacity needs in that period. But, on the other hand, it would be highly inefficient to commit thermal generation that is unneeded when intermittent resource output is high, and its commitment would either require curtailing zero emission generation or decommitting other resources.

The proposed requirement for transmission tags at t-40 in order for inter-BA transfers to count towards the RSE capacity test is the subject of Section 7. This proposal addresses a concern that transactions not tagged at t-40 are less likely to flow in real-time compared to transactions that are tagged at t-40. We understand from the CAISO that very few transactions that are not tagged at t-40 are subsequently tagged and flow, and we support the change based on that understanding.

Section 8 considers failures of the RSE under emergency conditions and the possible consequences for system reliability. We conclude that section by stating that there would be negative efficiency and reliability consequences if the long run RSE design freezes EIM imports when there is adequate capacity across the west to meet load but individual balancing areas are short because of unexpected rapid changes in system conditions impacting their ability to meet load At the same time, the current design attempts, albeit imperfectly, to ensure that when load shedding is needed because there is not enough regional capacity, the cost of load shedding falls on the balancing areas that have inadequate resources to meet their load.

Our conclusions are comprehensively summarized in the closing section (Section 9).

2. Resource Sufficiency Evaluation Objectives

From the inception of the EIM, the resource sufficiency evaluation (RSE) was intended to "ensure each EIM entity is able to meet their demand with their own net supply prior to engaging in transfers … through the EIM in the real-time market."[5] In the EIM context, the participation of a BAA in the EIM without sufficient capacity and ramping flexibility to cover its expected needs is considered "leaning." This definition of leaning is consistent with earlier usage in in tight power pools, such as the original PJM pool, where it referred to participating in the pool without bringing enough resources to the pool to meet one's load and thereby creating reliability risks for other pool members. In the context of the EIM, if there were no sufficiency evaluation, a balancing area without enough resources could have its load met in the five-minute market (RTD) with the responsibility for load shedding borne by all.[6]

Therefore, the objective of the RSE is to incent participants to make advance arrangements to provide sufficient resource capability to meet their own needs, while accounting for the diversity benefit provided by the EIM. The EIM could provide opportunities to meet those needs at lower cost but would not be relied upon as a market from which critically needed resources could be procured in the absence of self-sufficiency. Even though the EIM could, in theory, provide a venue in which a resource deficient BA could procure its needed energy when its neighbors enjoy a surplus, such transactions were deemed undesirable, or at least a source of undesirable incentives, when the EIM was formed.

Is important to recognize the distinction here between long run capacity planning and a "real time" RSE. An entity would not necessarily have to have made adequate long-term resource plans in order to pass the RSE in a given interval. It simply needs to be able to procure those needed resources in some market prior to the RSE. In this sense the RSE works to push forward some market activity out of the EIM into daily and longer-term markets.

Of course an entity with inadequate resources under its control or long-term contract runs a risk that it would not be able to procure the capacity needed to meet its resource sufficiency evaluation requirement in a daily or intra-day market, so there is still a relationship. Conversely, long-

[5] California ISO, "EIM Resource Efficiency Evaluation Enhancements Issue Paper, May 28, 2021, p. 3.

[6]This definition of leaning is different than how it is sometimes used in non-ISO context, where leaning can also be taken to mean the forcing of unscheduled, out-of-market flows from a neighboring balancing authority through a failure to adequately balance local demand with procured supply. This type of "leaning" is a non-market practice with extremely serious potential reliability consequences. By contrast, leaning in the EIM context would involve a market transaction between an entity with available supply and an entity that failed to procure enough resources prior to the running of the EIM. Further, because the EIM is a voluntary market, non-CAISO BAAs could choose to withhold some generation from the EIM to protect against EIM induced exports, leading to shortfalls in their BAA. However, such selective participation in the EIM would degrade the efficiency and benefits of the EIM. Therefore the RSE should be implemented in such a way that enhances confidence and participation in the EIM, rather than discouraging participation and reducing its benefits.

term resource adequacy plans operate on expectations and are of themselves no guarantee of adequate resource availability when stressed or unexpected real time conditions arise. In addition, as discussed below, optimal day-of market decisions taken prior the timing of the RSE, for example in the Hour Ahead Scheduling Process (HASP), can move the procurement of some "available" resources, such as hourly imports, into the EIM. Therefore, even though the resources were available and could have been procured in HASP, the process of maximizing the efficient use of those resources via deployment in the EIM could trigger an RSE failure. This is one example of how a rigid application of the RSE can raise overall operating costs without necessarily providing much incentive for improved long-term resource procurement.

In addition to the stated goals of the resource efficiency evaluation, it is worth considering the consequences of false positives and false negatives with this test. This can help inform decisions as to the costs and benefits of including imperfect and volatile measures of available capacity and flexibility.

Looked at from this perspective, the failure of a resource sufficiency evaluation triggers an operational penalty meant to provide incentives for better planning in the future. A false positive (when an entity passes the test when it should have failed) threatens to weaken those incentives, although to a degree that is very difficult to quantify. A false negative (when an entity fails the test when it actually should have passed) by contrast creates immediate and more easily quantifiable consequences in terms of higher costs and potentially reduced reliability through the artificial limitation of the EIM transfers.

In examining the different elements of the sufficiency evaluation, it is therefore important to also consider that we want to enable entities that procure and assign resources sufficient to cover a reasonable expectation of their local needs to fully participate in the Western EIM market, but the goal is not to sanction entities who fail to cover a requirement simply because it is overly volatile and difficult to predict. Further, transactions through the EIM should not be penalized relative to comparable transactions that occur on a marginally longer timeframe such as HASP. Such trade-offs do not appear to be in conflict with the high-level objectives of the RSE. We believe that the benefits of the EIM to participants will be jeopardized by overly punitive consequences that ensue because of artificially stringent tests, which could reduce the benefits of EIM participation and at times even endanger reliability.

3. Changes Regarding Net Load Uncertainty and Intertie Deviation Requirements

3.1 Statistical Issues

Before turning to a discussion of the CAISO's proposal, we begin with a brief explanation of the fundamental statistical issues contributing to the problems with the CAISO's current implementation of the net load uncertainty and intertie deviation requirements.

Our comments focus on two statistical issues with the CAISO methodology for calculating the uncertainty adjustment. First, the current look-back methods for calculating the intertie deviation uncertainty requirement and to some extent the net load uncertainty requirement do not appropriately account for the statistical issues associated with estimating the variance and tails of

an unknown distribution and rely upon on inadequate sample size tor these calculations. Second, the current methodology adds the uncertainty requirement calculated for net load uncertainty to the intertie deviation uncertainty requirement. This methodology results in an uncertainty adder that covers these uncertainties at a far higher than intended level of uncertainty, resulting in more resource sufficiency evaluation failures than intended.

3.1.1. Sample Size Issues. In general, the choice of sample size in an estimation procedure involves a tradeoff between statistical precision (larger samples reduce imprecision of estimates) and bias (larger samples introduce bias if sample size is enlarged by using older data, whose distributions may differ from more recent information due to nonstationarities of load and variable renewable distributions within the year and between years[7]). The ISO's compromises between these two objectives of precision and avoiding bias are not unreasonable, but the limited data available necessarily lead to statistical imprecision and volatility in the results, especially in estimating tails of distribution. We now discuss those issues.

The sample size currently used to calculate the intertie deviation uncertainty requirement is far too small to provide a reasonably accurate estimate of the variability of intertie deviations and this is likely also the case for the sample used to calculate the net load uncertainty requirement for weekends. By chance, just one or two unusual values in a data set will have outsized influence on the estimated variance. This has the consequence that the requirements will often cover these uncertainties at a higher or lower level of probability than intended. These errors will at times be in the direction of covering these uncertainties at a lower level than intended, potentially allowing EIM entities to pass the test when they should not. On the other hand, those errors will also at times be in the direction of covering these uncertainties at a higher, potentially much higher, level than intended,--causing EIM entities to fail the test when they should not. If EIM entities consistently schedule enough capacity to pass a properly defined resource sufficiency evaluation, the anomalous outcomes impacting the setting of the requirements will cause EIM entities to fail the test when they should not. The data that the Department of Market Monitoring (DMM) has begun posting would allow analysis of the frequency of such anomalous outcomes, but we are not aware of any such analyses to date.[8]

The underlying statistical issue is that the sample size used to calculate these uncertainty adjustments may be large enough to reliably estimate the mean of an unknown distribution but is much too small to estimate the variance or other measure of the shape of the tail of an unknown distribution. For example, if we look back over the same hour in the last 90 days to estimate the 2.5% tail for interchange under-deliveries, that tail would be defined by 3 data points.[9] With only 90

[7] Sample sizes have also been enlarged by the CAISO's use of several data points within a given hour associated with multiple intervals, but the strong autocorrelations within the hour mean that the degrees of freedom added by this expansion are, in effect, much fewer than if the samples were truly independent.

[8] The DMM has long identified the existence of statistical problems, but has refrained from tracking the frequency of anomalous outcomes because that would require DMM to develop what it thinks is the correct requirement, and then calculate the frequency of anomalies relative to the assumed "correct" outcome. Such an effort would require definition of "correct", which is the effort of the future market design process.

[9] We mention 90 days to illustrate the impact of sample error for the intertie deviation calculations. However, to calculate the net load uncertainty-based flexible ramping requirement for a given hour, the logic instead relies on the

data points included in the sample, individual samples would from time to time include more than 3 data points drawn from the 2.5% tail, resulting in uncertainty estimates that are larger than the true distribution of net load uncertainty. The magnitude of the estimation error in turn depends on how tight the true distribution of outcomes is in the tail. If the tail is thicker than a normal distribution (which is the case for many distributions of outcomes in electricity markets) variations in the sample could result in large anomalies from time to time in the uncertainty calculation. This could result in inappropriate failures every time the resource sufficiency evaluation is based on an anomalous value for either the net load uncertainty or intertie deviation uncertainty requirements. While these anomalous values would be sporadic, it should be a concern if they account for a material proportion of sufficiency evaluation failures.[10]

The impact of small sample size on the variability of estimates of the tail of a distribution is well illustrated by a graphic DMM prepared in 2015 to illustrate this point. The graphic portrays the variations in the CAISO flexible ramping required estimated using the BARR tool, which looked back over a limited number of intervals. The graphic shows clearly how the estimate swung back from the cap to the floor as small numbers of data points entered and exited the sampling period.[11] The number of data points included in the sample has been increased since 2015, so this figure does not the current variability of the uncertainty requirement, but our point still applies: the necessary compromise between statistical precision (from enlarging samples) and avoiding bias (from avoiding enlarging samples to include less relevant data) results in samples that are too small to yield stable and precise estimates of tails of distributions and, thus, the requirements based on those tables.

---

latest 40 days for weekdays and the latest 20 weekends, and uses data from multiple intervals within each hour to make the sample much bigger. Currently, deviations are calculated for each of the 12 RTD intervals within an hour, allowing use of 320 data points for weekday hours. But in that case, the effective degrees of freedom would be much smaller than 320 due to correlations of intervals in the same hour, and the autocorrelation from day-to-day, if any. There are, in reality, only 40 completely independent observations. (Quantifying the actual effective degrees of freedom is unfortunately difficult, but we anticipate that it would be closer to 40 than 320 in that case.) Meanwhile, the number of data points would only be 160 for weekend hours, with only 20 actually independent observations.

[10]Statistical theory gives some insight on the large uncertainty associated with estimates of standard deviations when samples are small. For a normal distribution, the standard error of SD (the estimate of the standard deviation) itself is well approximated by $SD/(2(n-1))^{0.5}$, for a sample size n greater than say 20, assuming independent samples (e.g., S. Ahn and J. Fessler, "Standard Errors of Mean, Variance, and Standard Deviation Estimators", University of Michigan, Dept. Elect. Eng., 2003, https://web.eecs.umich.edu/~fessler/papers/files/tr/stderr.pdf) . So, for a 90-day sample, the standard deviation will be off by 15% or more in about 5% of the time (+/- two standard deviations = +/- $2* SD/(2(n-1))^{0.5}$. If samples are positively correlated (days with greater than average values tend to be followed by additional days that are above normal), then the error is much greater. For, instance, the effective sample size N* for calculating standard errors of the mean is $N^*=N(1-r)/(1+r)$, where r is the autocorrelation of the data (C. Bretherton, Statistical Sampling Uncertainty, Lecture 3, U. of Minnesota, 2015, https://atmos.uw.edu/~breth/classes/AS552/lect/lect03.pdf.). So, with an autocorrelation of 0.5, the effective sample size when there are 90 days is actually only 30 days, which roughly doubles the standard error. Further error can be introduced if the distribution is non-normal with heavy tails or if the distribution is nonstationary over the 90 days (e.g., the standard deviation changes from month to month). Although we have information on estimated autocorrelations or tail characteristics, for instance, we strongly suspect that the estimated standard deviations are unstable and lead to the problems discussed in the text.

[11] For instance, comparing the hour-by-hour requirements for April 5, 2015, with March 31, 2015 in the figure shows large differences although the sampling window used to construct the histograms differed by only 6 days. The instability over hours within a day also is at least in part an artifact of sample error.

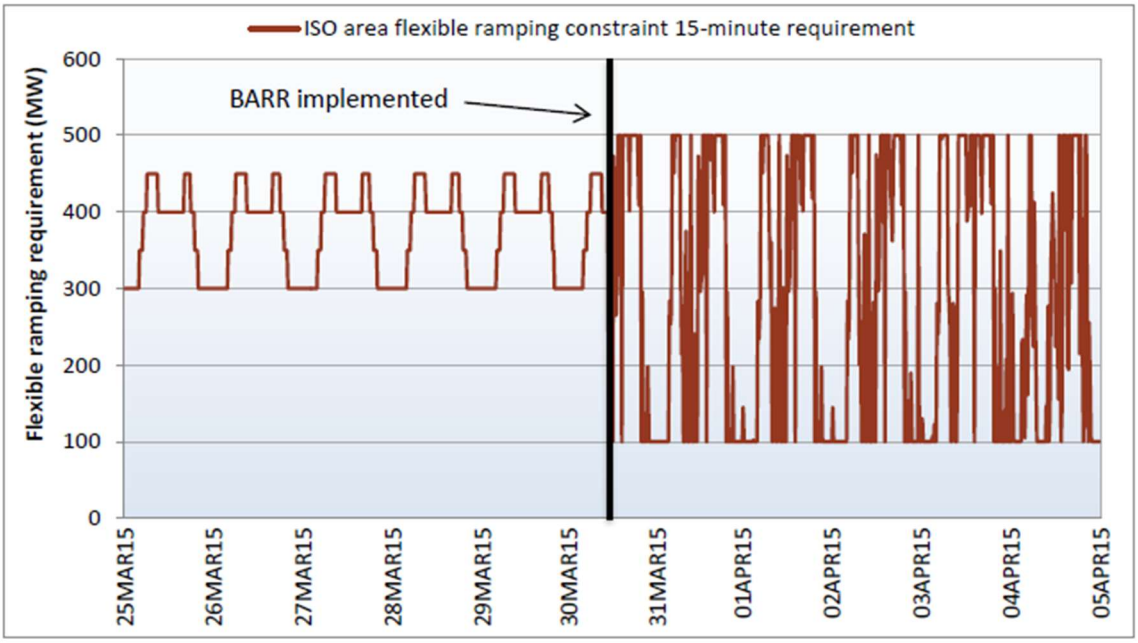

Figure 1. CAISO Flexiramp Requirements 2015. (Source: California ISO, Department of Market Monitoring, Q2 2015 Report on Market Issues and Performance, August 17, 2015, Figure 2.4, p. 44)

We see indications of similar variability of the CAISO's intertie deviation uncertainty measure in Figures 15 to 36 in the "Analysis of the Intertie Deviation Adder used in the Capacity Test."[12] We understand from the CAISO that 320 data points are used to estimate the net load uncertainty requirement for weekdays and 160 for weekends.[13] If they were independent samples, 320 data points would be a reasonable sample size, but since there are 12 data points for each historical hour covering the 12 RTD intervals, there are actually only 40 fully independent data points for weekday hours and 20 fully independent data points for weekend hours. As a result, the effective degrees of freedom in the statistical analysis may be much closer to 40 than 320. We understand that it is not simple to increase the sample size. For instance, looking back even 8 weeks as the CAISO does today has the impact that the sunrise and sunset are in a different hour during some periods of the year and load conditions can also be very different 8 weeks in the past, e.g., June vs April. Alternatively, using data from the same season in prior years can understate variability when there are significant year to year changes in the resource mix.

Stakeholders could get a sense of whether the small sample size used to determine the net load uncertainty and intertie deviation uncertainty requirements is leading to anomalous outcomes by examining the stability of the uncertainty adjustments over the year and examining whether a material number of failures are associated with anomalously large uncertainty values. It is some-

[12] California ISO, "Analysis of the Intertie Deviation Adder used in the Capacity Test," October 6, 2021

[13] See Footnote 9, infra.

what difficult to infer the extent to which this is happening from the CAISO October Intertie Deviation report.[14] There also does not appear to be any reporting of the frequency of anomalous values used to calculate the net load uncertainty requirement, nor of the extent to which these anomalous values are resulting in sufficiency evaluation failures,[15] but the data now being posted by the CAISO's Department of Market Monitoring would enable this kind of analysis.

3.1.2. Combining Uncertainties. The current methodology for calculating the uncertainty requirement for the resource sufficiency capacity test calculates an uncertainty adjustment for net load uncertainty and a separate uncertainty adjustment for intertie deviation uncertainty, then adds the megawatt values together. This is not a statistically valid approach. The standard deviation of the sum of two distributions is not equal to the sum of the standard deviations of the two distributions except in the extreme case in which the distributions are perfectly correlated.[16] Hence by calculating the 2.5% tail of one distribution and then adding this value to the estimated 2.5% tail of another distribution, we are calculating a value that can be much less likely than 1/40. If the two distributions were normal and uncorrelated with the same variance, we would be calculating an uncertainty requirement to protect against a much rarer probability event, i.e., a less than 0.3% tail.[17]

The data can be analyzed to determine how much of a problem the resulting over-conservatism is; if it turns out that the two variables are nearly perfectly correlated, then the CAISO's procedure is valid (i.e., there is one significant principal component). On the other hand, if the correlation is well below 1, then the procedure suggested in the following paragraph would avoid defining requirements that are too large and have much smaller than desired probabilities of exceedance.

This flaw in the methodology for calculating the uncertainty adjustment could be corrected by calculating the covariance between the two distributions but this would be complicated, and would also be subject to the small sample error discussed in Section 3.1.2. A simpler and more defensible approach if the two values were drawn from the same time period would be to add the

[14] California ISO, "Analysis of the Intertie Deviation Adder Used in the Capacity Test," October 6, 2021.

[15] As noted above, such an effort requires definition of a "correct" outcome, which is the subject of future processes.

[16] If two variables X and Y are statistically independent, then the variance of the sum X+Y is the sum of the variances of the two variables, where variance is defined the average squared deviation from the mean. This means that the standard deviation (the square root of the variance) of X+Y is $(var(X)+var(Y))^{0.5}$, where var(X) is the variance of variable X. For instance, if X and Y have equal standard deviations, then then standard deviation of their sum is 1.41 times the standard deviation of each variable. Summing the standard deviations would grossly overestimate the uncertainty of their sum (resulting in 2 times the standard deviation in this case). On the other hand, if X and Y are perfectly correlated (r=1), then indeed the standard deviation of X+Y is the sum of the standard deviations of the two variables, which is what the present procedure implicitly assumes.

[17] The exact probability depends on the relative sizes of the two standard deviations, their actual correlations, and whether they are normal distributions or not. Continuing the example of the previous footnote, if the standard deviations of X and Y are both 1.0, assuming they are perfectly correlated results in a standard deviation of 2.0 for X+Y. The 97.5th percentile of that sum is then 1.96 standard deviations (=3.92) above the mean. However, if in reality X and Y are uncorrelated (so X+Y has a standard deviation of 1.414 = SQRT(2) rather than 2), 3.92 above the mean would not be the 97.5th percentile—it is instead the 99.72nd percentile (roughly a 1/350 chance rather than 1/40 chance of exceedance).

values for each time period and calculate the 2.5% tail of the sum of the values.[18] A solution would be more complex to develop if the tail is calculated by drawing data points from different hours and days for the two components of the uncertainty requirement.[19]

3.1.3. Conclusions about Statistical Issues. These considerations imply that the current methodology for determining the net load uncertainty and intertie deviation uncertainty requirements utilizes clearly flawed statistical methods. The impact on EIM entities of the potential for anomalous uncertainty requirements is magnified by the current design, which looks back in time over similar periods of the day in calculating these requirements but does not look back over similar market conditions. Thus, high requirements for upward uncertainty could be calculated based on outcomes on days with high levels of intermittent resource output or high levels of intertie imports in base schedules, and then applied to days with very low levels of intermittent resource output and low levels of base schedules for intertie imports. The concerns of EIM entities with outcomes of this type were discussed in a prior stakeholder process and are one of the reasons the CAISO is developing the quantile regression methodology.

However, while there clearly are problems with the current statistical methodology for determining the net load uncertainty and intertie deviation requirements, we have not been able to assess how important these issues are in practice, but this should be possible over time with the data now being posted by the CAISO Department of Market Monitoring.

It is not reasonable that the estimate of net load forecast uncertainty or intertie deviation uncertainty used in the resource sufficiency capacity or flexibility tests would change dramatically from day to day, if no information specific to a given day's resource or demand conditions is considered. But such changes can occur under the present procedure due to sample error merely by as an artifact of replacing 1 day in the 90-day sample, if that one day happens to be relatively extreme. There can be dramatic changes from one day to the next in the estimates if the estimates are based on small number of data points, so that a few data points leaving or entering the sample can lead to big changes in the estimated value.

---

[18] In actuality, the data is not drawn from the same look-back period for the two uncertainty measures, but analysis of data for coincident periods could identify the correlation that is likely to apply in a given day and interval, and could be used to adjust the estimated confidence interval of the sum of the variables. In particular, CAISO staff have explained that the FRP requirement is derived from a different period used to derive the intertie deviation requirement. FRP is based on last 40 weekdays or last 20 weekend for requirements for week or weekend days, respectively. This is a moving window with a couple of days delay. In contrast, the intertie deviation relies on a static window of 90 days from the past and has a latency of about 15 days. For instance, January requirements are calculated based on the 90 days from September 15 through December 15. The 15-day latency to January 1 is to give staff the opportunity to filter out outliers (a process that is defined through the CAISO's Business Practices Manual).

[19] One possible approach would be to add a value based on a much lower point in the distribution for one uncertainty to the 2.5% tail for the other uncertainty. For example, the CAISO could add the value of the 25% tail for interchange uncertainty to the 2.5% tail for net load forecast uncertainty. These values are arbitrary, but the CAISO could calculate the historical performance of a combination of the 2.5% tail estimate for net load uncertainty and a lower threshold for intertie deviations and examine its performance in covering the actual total variations in net load and intertie schedules. The CAISO could also study the impact of calculating a combined uncertainty metric using data for the same hour and compare that to the sum of the values used historically. This might help the CAISO and stakeholders understand the magnitude of the impacts of the current approach.

If there is a pattern of the estimated net load uncertainty requirement or the intertie deviation requirement jumping up and down from day to day, that would suggest that the CAISO is not accurately estimating the underlying relationships. While we identify this statistical issue, we have had have the time to analyze the net load uncertainty and intertie deviation requirements that are actually being used by the CAISO (which are available on OASIS and by DMM), so we were unable to assess in this opinion whether this statistical issue is in practice having a material impact on the requirements estimated by the CAISO, nor can we assess if these anomalous values are contributing to a material number of spurious resource sufficiency evaluation failures.

The data now being posted by the Department of Market Monitoring should facilitate analyses that would better inform discussion of these issues. We have reviewed some of that data, but it only covers the October-December period, so it does not yet provide insight into the level of variability in requirements during other seasons. Over time, graphics could be generated showing the daily values of the requirements used for each hour over a 90 day or longer period (separately for weekdays and weekends). This kind of graphic would help stakeholders, and the MSC, understand the practical impact of these statistical issues. This kind of review could also be carried out prospectively when the CAISO's quantile regression method is able to generate uncertainty requirement calculations.

A higher variability of the net load and intertie deviation uncertainty estimates used by the CAISO to apply the tests makes it difficult for balancing areas to predict their requirement and make sufficient resources available to pass the test. Hence, balancing areas not only may fail the test because the requirement is too high for them to meet, they may fail even when they could have had enough capacity available to meet the requirement had they known the requirement further in advance.[20] While the current design gives the balancing area 10 minutes or so to adjust the resources available to meet the test, this may not be enough time when there are large anomalous values for the uncertainty requirements and meeting the requirement would require them to have committed another unit. The impact of anomalous requirements in causing spurious sufficiency evaluation failures would very likely be greater if off-line units were not counted as part of the available supply, as we discuss below.

3.2 CAISO Proposal

The CAISO proposes in the interim period until the quantile regression method is implemented to no longer include either the net load uncertainty requirements or the intertie deviation uncertainty requirement in the resource sufficiency capacity test, but to continue to include the net load uncertainty requirement in the flexibility test. This essentially reverts to the design prior to summer 2021 for the net load uncertainty requirement. The intertie deviation requirement was included in the capacity test prior to summer 2021.

This design change would avoid inappropriate failures of the capacity test due to anomalous intertie deviation or net load uncertainty requirements arising from the statistical issues discussed above. Conversely, however, the change would result in a capacity test that does not include any

[20] We note that this impact is related to the flaws in flexiramp pricing that we discuss in Section 4.1 that have the effect that there is no compensation for EIM entities providing excess flexible capacity since the price is uniformly zero.

requirement for capacity to meet either net load forecast uncertainty or to meet intertie deviation uncertainty.

The CAISO proposal does not involve any changes to the flexibility test, which would continue to include a net load uncertainty requirement. On the one hand, EIM market participants could therefore continue to fail the flexibility test when they should not because of anomalously high net load uncertainty requirements due to the issues discussed in Section 3.1. On the other hand, the flexibility test alone will not necessarily be effective in deterring leaning as high levels of EIM transfer imports could allow an EIM entity to pass the flexibility test even if the EIM entity does not actually control enough capacity to provide upward flexibility. This possibility is currently constrained by the capacity test, but the capacity test will be somewhat looser once it is applied with the net load uncertainty requirement and intertie deviation uncertainty requirements both removed.

Stakeholders have pointed out that the mere fact that addition of the uncertainty requirement to the resource sufficiency capacity test has resulted in more failures does not necessarily imply that there is anything wrong with the test that needs to be fixed. We agree with this, but as discussed above, there are reasons to expect that the current design could lead to EIM entities failing the resource sufficiency evaluation when they should not. However, we do not know if this is the case in practice, hence our suggestion that the CAISO and stakeholders make use of the historical data compiled and posted by the Department of Market Monitoring portraying actual day-to-day variations in the requirements to assess the extent to which failures are associated with anomalous requirements. There is more than one way to define anomalous outcomes, but one approach would be to use a threshold such as a value that is 50% higher than the value 7 calendar days before.

Moreover, as discussed further in Section 4 below, continuing flaws in the implementation of the flexiramp requirement in the market have the effect that leaning will not be deterred by the EIM market, as the shadow price of flexiramp in the EIM is almost uniformly zero, even when it should not be. This has the consequence that failing to have enough flexiramp to meet an EIM entities requirement has no cost, as the requirement can be met by purchasing flexiramp from other balancing areas at a zero price.

3.3 Future Methodology

While the quantile regression method proposed by the CAISO would be able to incorporate information specific to the day the estimate is applied to, such as the forecasted wind or solar output in each hour, this methodology will still tend to produce highly variable estimates if the estimates are based on a small sample.[21] There would also be the same statistical issue we discussed

[21] Our experience with quantile regression is that estimates of coefficients of the model can be highly unstable for high values of "tau" (the probability of exceedance being assumed) when samples are small. For 90 data points, a 97.5th percentile quantile regression will have only about 3 points above the line separating high and low values; if this line is being estimated as a function of several variables (such as forecast wind, solar, and load), it will be highly unstable. An alternative approach which might be somewhat more stable is to estimate quantile regressions for smaller values of tau (such as 75th and 90th percentiles) and then extrapolate to the 97.5th percentile using a normal

above if the net load and intertie deviation uncertainties were estimated separately, and the estimated measures of uncertainty added.

We share the concern of many stakeholders regarding how long it will be until CAISO will be able to implement the quantile method for estimating net load uncertainty and an improved method for estimating the intertie deviation requirement. We also share the concern of many stakeholders that the CAISO should not spend several more years developing the quantile regression methodology without discussion of the methodology with stakeholders so there can be assurance that the methodology does not have fundamental flaws that will prevent its successful implementation. The CAISO has made commendable efforts to communicate the details of its quantile regression methodology through publishing a technical memo and making presentations,[22] but until recently we perceive that stakeholders have not appreciated the importance of this technical change and its implications not only for flexiramp requirements but also for the tests in the RSE. Moreover, the CAISO presentations are very general and do not discuss the time periods that were used to estimate the requirements. In addition, while there are statistics comparing closeness and the amount the estimate exceeds what is necessary under the histogram and current method, there are no data showing the overall performance of the quantile method uncertainty estimates or the frequency of anomalous estimates.

We therefore recommend that the CAISO provide information on the quantile models being developed and their performance both so there can be discussion of the approach and the testing. This will help stakeholders assess if the quantile regression methodology is likely to yield acceptable results and so stakeholders can assess how much longer it will likely be before this methodology would likely be ready for a successful implementation. There is also a need for the CAISO to discuss what method it will use to combine estimates of net load uncertainty and intertie deviation uncertainty in light of the statistical issues discussed above. We understand that this information is to be part of the analysis to be undertaken during Phase 1b of the initiative.

In general, in choosing a longer-run solution, there is a tradeoff between the need to avoid the small sample problems that cause high variability in the tests, while still allowing the test to be responsive to the most recent and relevant system conditions for assessing risks. We recognize that there is no easy answer, and that there is unlikely to be a statistical procedure that is stable and predictable, while making risk estimates reflective of changing conditions. Below we suggest some possible approaches that might help mitigate the small-sample problem without artificially stabilizing estimates and making them unresponsive when risks are in reality changing. It would be useful in discussions with stakeholders about alternative approaches if data is made available that would facilitate assessment of their performance.

---

distribution assumption for the shape of the tail. This will underestimate that more extreme value, however, if the actual distributions have heavier tails than a normal distribution.

[22] E.g., CAISO, Flexible Ramping Product Refinements Initiative, Appendix C - Quantile Regression Approach, May 8, 2020, www.caiso.com/InitiativeDocuments/AppendixC-QuantileRegressionApproach-FlexibleRampingProductRequirements.pdf

If anomalous values for the uncertainty adjustments appear to be causing a material portion of EIM entity sufficiency evaluation failures, we agree with the view of a number of stakeholders that it would be desirable for the CAISO to have short-term, easy to implement changes to mitigate these impacts while long run improvements are being developed, particularly since the time frame for implementing those long run improvements is uncertain. [23] However, we recognize that this is much easier to propose than for the CAISO to implement.

To deal with the small sample error problem, the CAISO could as an interim solution apply the current methodologies with a much longer look-back period so as to greatly increase the number of data points used to estimate the tail of these distributions.[24] The CAISO could also combine the values for net load uncertainty and intertie deviation uncertainty and calculate a single estimate of the tail of the distribution. One shortcoming of such a methodology is that it would be slow to react to changes in market conditions, such as rising levels of intermittent resource output or continuing changes in the distribution of weather outcomes.[25] Such an approach would not address the problem of drawing data points from time periods with different market conditions, such as high or low intermittent resource output. In addition, even simple modifications would require CAISO resources to implement and would not be appropriate to divert resources to such an effort if the development of the quantile regression methodology is nearly complete and the design has performed sufficiently well in testing to provide some assurance that it will meet these needs.

Depending on the timeline for implementing a quantile regression methodology for estimating uncertainty requirements, it might be desirable to make use of the historical data being posted by the Department of Market Monitoring to assess the type of short-term interim changes that might be able to reduce the variability of the estimated requirements under the current methodology. The objective would be to enable these requirements to be applied in some fashion until the quantile regression or other end-state methodology can be implemented.

4. Load Conformance and Capacity Test Requirements

There is a concern of some EIM participants that the level of CAISO EIM transfer imports reflects undue leaning on the pool and that this leaning is facilitated by the use of load conformance adjustments by CAISO operators. This has led to a suggestion that load conformance adjustments be included in the load used to calculate the capacity requirement for the resource sufficiency capacity test. We do not agree with this recommendation and also have a different view of the implications of the EIM transfer data. In this section, we first discuss the EIM transfer data, then discuss implications of such a change in the resource sufficiency evaluation. We then

[23] Alternative approaches might be considered for estimating thresholds in the short-term, such as using some kind of long-run average cut by hour of day and maybe season, where the average could be a percentage of load or a MW number.

[24] A weighted procedure could be combined with using a larger sample, in which more recent values are weighted more heavily. The combined effect would be to lessen that the impact of dropping off the oldest value as the procedure rolls forward.

[25] This shortcoming might be somewhat mitigated by the weighting procedure proposed in the previous footnote.

discuss the concern with the potential for EIM transfers to enable EIM entities to pass the resource sufficiency flexibility test.

4.1 Significance of EIM Imports

We have heard a comment several times that EIM imports are supposed to be driven by economics and that the high level of EIM imports by the CAISO over the past year reflect a lack of capacity within the CAISO rather than economics. However, we believe that this pattern may at least in part have been driven by economic considerations, albeit economic considerations unfortunately created by market design and implementation flaws. These flaws involve the way flexiramp targets and available flexiramp supply are currently determined for the EIM balancing areas outside the CAISO in the FMM and RTD.

DMM reports establish four facts that are very relevant to the level of CAISO EIM imports since the minimum requirement was implemented for the CAISO balancing area in FMM at the end of 2020. First, the implementation of the minimum locational requirement for the CAISO has often caused the shadow price of flexiramp in the CAISO in the FMM to fall to zero, although it is non-zero in significant number of FMM intervals and the shadow price is sometimes quite high.[26] Non-zero shadow prices of flexiramp will increase the cost of meeting CAISO load with CAISO resources whenever the dispatch of a CAISO resource for energy would reduce the supply of flexiramp by a similar amount. Second, the shadow price of flexiramp in the EIM region in the FMM has remained zero in almost every interval and is virtually always zero for individual EIM entities except when they fail a sufficiency evaluation.[27] This causes EIM imports to the CAISO to appear to be lower cost in the FMM than energy from a CAISO resource with the same energy offer price, which, all else being equal, would decrease EIM exports to CAISO, and decrease CAISO exports to the rest of the EIM in the FMM. Third, the shadow price of flexiramp in CAISO has remained zero in RTD because as DMM has reported, the CAISO decided not to apply the locational flexiramp requirement in RTD.[28] This has the effect that the CAISO resource that is not dispatched in FMM because it would reduce the supply of flexiramp within the CAISO balancing area will have a lower cost in RTD and is likely to be dispatched instead of the EIM resource scheduled in the FMM. Fourth, the shadow price of flexiramp in the EIM in RTD has also remained zero in almost every interval and is almost always zero for individual balancing areas except when they fail a sufficiency evaluation.[29]

The implication of these facts is that there will be times when the shadow price of ramp is positive and material in the CAISO in FMM but the shadow price of flexiramp for the EIM region outside the CAISO will almost always be zero in the FMM.

[26] California ISO, Department of Market Monitoring, Q3 Report on Market Issues and Performance, December 9, 2021. Figure 1.27, p. 31.

[27] Ibid..

[28] Ibid., p. 29.

[29] Ibid. p. 30.

When the flexiramp constraint is binding in the CAISO, this typically raises the cost of dispatching CAISO generation, as dispatching generation up to meet load will often incur the penalty price for reduced flexiramp procurement.[30] If the price of flexiramp is zero in the EIM region outside the CAISO, however, then there is no trade-off between energy and flexiramp in dispatching EIM resources located outside CAISO to meet load, and there will be no similar cost to using increased EIM transfers to free up flexiramp capacity. Hence, HASP and RTPD will dispatch EIM resources to meet CAISO load instead of dispatching CAISO resources with similar energy offer costs, because dispatching the EIM resource will not incur a penalty cost for reducing flexiramp supply, even when the EIM market is very tight on supply. If the flexiramp price in the CAISO is $50/MW, then the CAISO could dispatch EIM resources for energy that have a $50 higher incremental energy costs than CAISO resources. If the CAISO flexiramp price rose to $247, then the CAISO could dispatch EIM resources with a cost $247 higher than CAISO resources.

Hence, whenever the CAISO begins to trade-off energy and flexiramp in RTPD within the CAISO balancing area, the HASP and RTPD economic evaluation will tend to schedule imports from the rest of the EIM into the CAISO instead of dispatching CAISO generation to meet CAISO load, or to meet load elsewhere in the EIM. Importantly, this will be the case even if the EIM resource being dispatched is within the same constrained region as the CAISO and the dispatch of the EIM resource reduces the supply of ramp within the constrained region the same as would dispatch of a resource located within the CAISO.

Therefore, while the level of load conformance in HASP and RTPD is part of the reason for high EIM transfers being scheduled into the CAISO in HASP and RTPD over the past year, another reason for these transfers on high priced days is that the shadow price of flexiramp is much higher inside the CAISO than in EIM entities, making EIM resources located outside the CAISO appear to be lower cost than comparable internal CAISO resources. This higher cost of CAISO resources will also tend to increase the portion of load conformance adjustments that is met in HASP and FMM with EIM transfers, rather than met by scheduling hourly imports in HASP or scheduling CAISO resources, because EIM transfers will appear to be a cheaper way to meet the load conformance adjustments.

In addition, after RTPD has scheduled these EIM imports in FMM, when the shadow price of ramp inside the CAISO falls to zero in RTD, internal CAISO generation will be cheaper than it was in RTPD and would be dispatched up to displace EIM transfer imports scheduled in FMM. This would also induce a dispatch down of non-CAISO EIM generation in RTD, which would be accompanied by reductions in non-CAISO EIM prices.

Hence the zero cost of flexiramp in the EIM in FMM likely contributes to high CAISO EIM imports in FMM, while the zero cost of flexiramp in the CAISO in RTD likely contributes to lower RTD prices and lower EIM imports in RTD.

[30] If all resources providing flexiramp were constrained by their ramp rate, rather than their Pmax (capacity limit) in the amount of ramp they provide, then there would be no trade-off between energy and ramp. This would also be the case if the resources that were constrained by their Pmax were so high cost that they would not be the incremental source of energy.

This low price for EIM ramp in FMM, which implies the existence of a surplus of ramp even during tight market conditions, means that the flexiramp requirement is not serving the scarcity pricing role that was intended by the CAISO and expected by EIM entities. This has the consequence that RTPD will not see the need to commit generation outside the CAISO to maintain ramp when it is getting short and there will be no gradual rise in energy prices as supply tightens and the price of flexiramp gradually rises. Instead, the energy price will rise from the incremental cost of energy to the power balance violation price when no flexiramp capacity is actually available to meet load in FMM or RTD, while the flexiramp price will remain zero.

A striking example of this outcome was July 9, 2022, when the CAISO FMM price for flexiramp was at least $50 and ranged up to $247 in many intervals during the evening shortage period. The EIM region price for flexiramp was zero throughout this period, indicating plenty of capacity available throughout the EIM over the evening. The price of flexiramp was also zero for most EIM entities although there were a few price spikes for flexiramp in SRP, NV Energy, and other balancing areas that apparently occurred when these EIM entities failed a resource sufficiency evaluation.[31] These zero prices of flexiramp persisted through many FMM intervals in which the price of energy exceeded $500 across much of EIM east for several hours. However, the EIM regions exporting to the CAISO were almost certainly not long on ramp over this period, as their energy prices were very high. These outcomes reflect a major disconnect between the price of energy and of flexiramp.

Only the CAISO, BANC and SRP appear to have had non-zero flexiramp prices during the emergency hours on that day.[32] EIM imports that result on flexiramp shortfalls should result in non-zero flexiramp prices. The fact that flexiramp prices remained low while prices were spiking to $1000 in RTPD and RTD indicates the magnitude of the unresolved flaws in the flexiramp design.

While the CAISO price of flexiramp was high in RTD in some intervals during the evening of July 9, it was zero during many intervals. The intervals in which the RTD price was high appeared to be the intervals in which the CAISO failed the resource sufficiency evaluation.

We do not know exactly why the shadow price of ramp in the FMM for the EIM region outside the CAISO remained zero on July 9, or on nearly every other day, even when energy prices are very high. But we initially conjectured that at least on July 9 this may have been due to a surplus of flexiramp within a constrained-down region in the northern part of the Pacific Northwest. This region was constrained down in the energy dispatch, with prices much lower than the rest of the EIM,[33] so no capacity in that region could be dispatched to meet load in the rest of the EIM. The existence of these transmission constraints also means that capacity in this constrained down region could not actually provide flexiramp for the rest of the EIM, but we speculate that unloaded capacity in this region was counted as providing flexiramp to the overall EIM region, including EIM balancing areas within the constrained-up region outside the Pacific Northwest.

---

[31] Real-Time Daily Market Watch - July 9, 2021, http://www.caiso.com/Documents/Real-TimeDailyMarketWatchJul09-2021.html

[32] www.caiso.com/Documents/Real-TimeDailyMarketWatchJul09-2021.html

[33] Ibid.

This congestion pattern could have resulted in zero shadow prices for flexiramp for the EIM region, and most EIM entities over the high-priced evening period.

However, recent discussions with the CAISO in response to questions the MSC has raised over the past few months regarding the energy and flexiramp prices on July 9 have identified additional elements of the current flexiramp design that appear to have contributed to these outcomes. Our current understanding is that while the lack of deliverability tests applied to balancing areas outside the CAISO may have contributed somewhat to these anomalous prices, it is likely that the low flexiramp prices on July 9 are largely due to two other factors. One factor appears to be the way the flexiramp design nests individual BAA requirements and flexiramp prices within the EIM. The other factor seems to be the way the CAISO minimum requirement was implemented. We are not going to go into these issues in this Opinion because they are still being explored by both the CAISO and the MSC and the details are off point. However, it is on point to recognize that flexiramp prices for the EIM region are being set at zero even when the EIM is short overall on ramp capacity, and very short within a constrained region.

These flexiramp prices have additional implications for RTD. In part because the CAISO locational requirement for flexiramp is not enforced in RTD, the CAISO shadow price of ramp consistently falls to zero in RTD, as does the price the rest of the EIM.[34] This lower shadow price of RTD ramp in CAISO makes it economic to dispatch CAISO capacity in RTD that was uneconomic in FMM, resulting in lower RTD prices within the constrained EIM region and lower EIM transfer imports into the CAISO in RTD, consistent with what happened over the past year, and in particular on July 9, 2021.[35]

We have previously noted that the pattern of the CAISO's load conformance adjustments--high in HASP and FMM, and low in RTD--has a similar effect of raising CAISO prices and EIM transfer imports in the FMM, and then causing them both to fall in RTD. We cannot assess the relative importance of flexiramp pricing and load conformance adjustments in producing this pattern with the data available to us, and the two processes probably interact. Nevertheless, it is important to recognize that this pricing pattern, and the level of EIM transfer imports into the CAISO, is in part driven by the anomalies in flexiramp prices. We think these observations, in particular the pattern observed on July 9, is yet another indication that the CAISO needs to focus on correcting flaws in the pricing of flexiramp, not only for the CAISO but throughout the EIM. The consequences of flaws with flexiramp go beyond the resource sufficiency evaluation. The zero price of non-CAISO EIM flexiramp in both FMM and RTD undermines incentives to offer flexible capacity beyond the minimum requirement.

[34] The typically zero price of flexiramp in RTD is also at least in part due to the way the flexiramp RTD requirement is set and applied in RTD. While these zero prices of flexiramp in RTD may be a correct outcome of the required design and requirements, it is not an appropriate outcome to have a zero shadow price when operators are apparently taking large out of market actions (such as load conformance adjustments) to increase the supply of ramp.

[35] California ISO, Department of Market Monitoring, Q3 Report on Market Issues and Performance, December 9, 2021, p. 30.

Furthermore, the zero price of flexiramp even when capacity is short also undermines the CAISO commitment to use flexiramp to provide a scarcity pricing signal in the EIM.[36] Correcting this problem should be a first step in any ISO initiative to discuss improvements to the current scarcity pricing design.

4.2 Load Conformance and the Resource Sufficiency Evaluation

At present, the resource sufficiency capacity test is carried out by calculating the available capacity relative to the net load forecast, not the net load value used in RTPD, which could include load conformance adjustments up or down.

Several stakeholders have argued that manual adjustments to the load forecasts, known as load conformance, should be included on the demand side of a resource sufficiency evaluation. DMM calculations show that the CAISO on average adjusts load upward using conformance in the HASP and FMM timeframes.[37] While it might seem logical to include any adjustments to the forecast to load in applying the resources sufficiency capacity test, this presumes these adjustments were made in order to reflect a more accurate forecast of the true demand.

But within the CAISO, this is usually not the case. It appears that load conformance is instead a tool applied by operators to compensate for the flaws in the flexiramp implementation and for the inconsistencies between the HASP and RTPD solutions and resource sufficiency requirements (discussed further below).

In practice, it is our understanding that CAISO operators use FMM and HASP load conformance adjustments to increase the amount of hourly imports scheduled in HASP and to commit additional units in HASP and RTPD, which both somewhat increases the supply of ramp and also helps the CAISO to pass the resource sufficiency capacity evaluation.[38] These adjustments do not reflect net load forecast error because they are much larger than the load conformance adjustments in RTD.[39]

Our understanding is that it is proposed by some EIM entities that the load conformance included in the RTPD load (either the actual value for that RTPD interval or the average value for that hour over some period) should be added to the actual load forecast in applying the resource sufficiency capacity test.

---

[36] California ISO, Flexible Ramping Product Refinements, Draft Final Proposal, May 8, 2020, Section 8, pp. 18-19.

[37] California ISO, Department of Market Monitoring, Q3 Report on Market Issues and Performance, December 9, 2021, Figure 1.53, p. 62.

[38] We do not know the relative importance of these factors in operator decision making, it may be that the size of the adjustments made to manage ramp obviate the need to consider adjustments to pass the resource sufficiency test.

[39] As documented in pp. 20-34 of G. Bautista-Alderete, "Resource Sufficiency Evaluation," Presentation at the CAISO Market Surveillance Committee, Nov. 19 2021, www.caiso.com/Documents/ResourceSufficiencyEvaluation-Presentation-Nov19-2021.pdf

We have two market and reliability concerns with such a proposed design change. We are concerned that such changes would prevent the CAISO operators from using market mechanisms, first, to enable the CAISO to pass the resource sufficiency capacity test (discussed further in Section 5) and, second, to ensure that there is sufficient ramp capability to balance variations in load in an operational time frame. While it is conceivable that CAISO operators could find ways to use non-market adjustments to achieve these goals, that is not assured. Such a non-market approach would have the potential for unintended market impacts that might prove difficult to fully anticipate and might also contribute to adverse reliability impacts.[40] We discuss these concerns in more detail below.

It is widely acknowledged that the flexiramp product is not working as intended and there are flaws in the flexiramp implementation that still need to be addressed.[41] As a result of the disfunction of the flexiramp implementation, CAISO operators appear to have come to rely on large load conformance adjustments in HASP and RTPD to enable them to have enough ramp to balance net load over the operating hour. These load conformance adjustments result in more units internal to the CAISO being committed, more HASP imports being scheduled (thus potentially creating ramp by backing down internal units), and more EIM transfer imports being scheduled (also potentially creating ramp by backing down internal units) than would otherwise be the case. A part (and only a part) of this increase in supply is translated into increased ramp that will be available to balance variations in net load in RTD. The supply scheduled in HASP and RTPD is by definition deliverable to meet load, as transmission constraints are evaluated on the HASP and RTPD schedules. However, since load conformance adjustments do not translate directly into additional ramp, operators may need to increase the load conformance by several megawatts in order to obtain a single additional megawatt of ramp.[42]

---

[40] The impact of out of market adjustments would generally be reflected in subsequent market runs which would tend to undo the impact of the out of market adjustments. This could probably be avoided to a degree by ongoing operator adjustments, but this would not only tie up the attention and time of operators during stressed system conditions but would be likely to lead to other unintended outcomes.

[41] These flaws include 1) the lack of an effective delivery test (hopefully to be addressed by the nodal delivery test); 2) the lack of minimum flexiramp requirement for the CAISO in RTD; 3) the need to implement different target levels for flexiramp requirements over different look-ahead horizons in RTPD (the Department of Market Monitoring has been pointing out the need for changes for several years); and now the additional issue of 4) the way EIM flexiramp prices are calculated. For discussions of these flaws, see CAISO, CAISO Energy Markets Price Performance Report, Prepared by Market Analysis and Forecasting, September 23, 2019, www.caiso.com/documents/finalreport-priceperformanceanalysis.pdf; J. Bushnell, S. Harvey, and B.F. Hobbs, Opinion on Flexible Ramping Product Refinements, Market Surveillance Committee of the CAISO, Sept. 9, 2020, www.caiso.com/Documents/MSC-OpinionFlexibleRampingProductEnhancements-Sep8_2020.pdf; CAISO Department of Market Monitoring, Comments on Issue Paper on Extending the Day-Ahead Market to EIM Entities, Nov. 22, 2019, www.caiso.com/initiativedocuments/dmmcomments-extendedday-aheadmarket-issuepaper.pdf.

[42] As we have discussed, the load conformance adjustment does not ensure that generation is dispatched so as to provide additional ramp. Some of the increased intertie supply will simply back units down their supply curve without creating additional ramp. Additional ramp is only created on on-line units when a unit whose available ramp is constrained by its upper limit is backed down. A portion of the flexiramp scheduled in RTPD may be behind transmission constraints while another portion will consist of phantom ramp that will not be available in RTPD. Similarly, because the locational constraint is not enforced for flexiramp in RTD, the flexiramp scheduled in RTD will even more often behind a transmission constraint, reducing the amount that is actually available to meet load in RTD.

We are at a point where the flexiramp product is not working to make additional ramp available, so CAISO operators apparently depend on their ability to make load conformance adjustments in HASP and RTPD in order to have enough ramp available to balance net load in RTD. Moreover, the operators likely need to make significantly larger load conformance adjustments in HASP and RTPD than the amount of ramp they need in RTD, because load conformance adjustments do not translate megawatt for megawatt into ramp.

In summary, the basic point is that the load conformance adjustments are not motivated by a desire to correct load forecasts. In fact, the power scheduled in FMM as a result of load conformance adjustments generally does not flow in RTD. There is therefore no leaning on the rest of the system as a result of using load conformance adjustments in HASP and RTPD; rather, the CAISO loads buy power at high prices in the FMM and sells it back in RTD.[43] This should be a concern to CAISO ratepayers, but it benefits non-CAISO EIM entities who export power to the CAISO in the FMM.

4.3 Flexibility Sufficiency Test

There has been discussion of a concern that an EIM entity, including the CAISO, could use EIM transfer imports to pass the flexibility test by backing down its generation. This view appears to be correct to us. Our understanding is that this way of passing the flexibility test should ideally be constrained by the capacity test. However, the capacity test will be somewhat weakened if the intertie deviation and uncertainty requirements are removed from the capacity test.

If there is a concern that EIM transfer imports allowing EIM entities to pass the resource sufficiency flexibility test when they should not, the solution should be to change the flexibility test to exclude the flexibility made available by EIM transfer imports.[44] This is more appropriate than including load conformance adjustments in these tests as most of the load conformance adjustments do not result in EIM transfer imports, and there are other mechanisms EIM entities can use to increase EIM transfer imports to pass the flexibility test. For example, EIM entities that control generation can raise the offer prices of their resources in order to attract EIM transfer imports. This recalls the concern with mitigation of resource offers a few years ago that caused EIM imports to change into exports.

We recognize that there are complexities in adjusting the resource sufficiency flexibility test to account for the impact of EIM transfer imports on the supply of ramp, and the resource sufficiency capacity test tends to address the potential issue. However, if there were a concern with EIM transfer imports being used to enable EIM entities to pass the flexibility sufficiency test, that concern should be addressed with changes to the flexibility test.

---

[43] See California ISO, Department of Market Monitoring, Q3 Report on Market Issues and Performance, December 9, 2021, Figure 1.12, p. 18, which shows the difference between FMM and RTD prices over hours 18-21. Figure 1.53, p. 62, shows the large differences in load conformance adjustments between the HASP, FMM and RTD. These data are illustrative of the increased cost to load but the actual increase in cost would require detailed calculations based on the change in cleared supply and prices at each location between FMM and RTD.

[44] This is not as simple as just adjusting flexibility by the amount of transfers, because in their absence the BAA might have committed internal resources instead.

5. Structure of HASP, STUC, FMM, RTD and Resource Sufficiency Evaluations

We first discuss issues with the structure of HASP in relation to the resource sufficiency capacity test, then turn to a discussion of issues with the structure of RTD, and to a lesser extent FMM and STUC, in relation to the resource sufficiency flexibility test.

5.1 Structure of HASP

It has been pointed out in the course of sufficiency evaluation discussions over the past year that the structure of the CAISO HASP creates some challenges for CAISO operators in passing the resource sufficiency capacity test. The HASP clears hourly intertie schedules in combination with EIM transfers, and will substitute lower cost EIM transfers for hourly imports if the EIM transfers are lower cost than the hourly imports.

However, when the resource sufficiency evaluation would subsequently be applied, the hourly transactions that were not scheduled in HASP would no longer be included in supply, and the EIM transactions notionally scheduled in HASP would also not be included in the supply for the resource sufficiency capacity test. This misalignment between the constraints enforced in HASP and the resource sufficiency capacity test creates the potential for the CAISO to fail the resource sufficiency capacity test when it could have passed by scheduling additional hourly intertie transactions in HASP.

A similar issue arising from the HASP design that stakeholders have noted is that there is a potential for the HASP to schedule hourly exports that are supported by EIM transfer imports. Those EIM transfer imports are not included in the CAISO supply for the resource sufficiency capacity test, so the CAISO could schedule HASP exports supported by EIM transfer imports, then fail the resource sufficiency capacity test and potentially be unable to schedule sufficient EIM transfer imports to support the HASP exports. This kind of scheduling of HASP exports supported by EIM transfer imports could undermine the EIM with EIM balancing areas meeting their sufficiency evaluations with hourly imports from the CAISO that are in fact supported by EIM transfers that deplete EIM flexiramp supplies, with the result that the targeted level of unloaded capacity would not be available in FMM .[45] The capacity sufficiency test plays an important role in making sure this kind of export scheduling supported by EIM transfers does not occur by incenting operators to take steps to ensure that there is sufficiency capacity to pass the capacity sufficiency test, with a critical step likely being load conformance adjustments for the CAISO.

One mechanism that we understand CAISO operators may use to avoid resource sufficiency evaluation failures is to make load conformance adjustments in HASP to ensure that excess supply is scheduled in HASP, enabling the CAISO to pass the resource sufficiency capacity test despite the scheduling of exports in HASP. Hence, these issues with the HASP design may be one of the reasons that CAISO operators make load conformance adjustments in HASP, so that the CAISO can pass the resource sufficiency evaluation.

[45] This would be much less likely if the cost of EIM transfer imports reflected the correct shadow price of flexiramp.

If load conformance adjustments were included in load for the resource sufficiency evaluation, CAISO operators would not be able to use load conformance adjustments to enable the CAISO to pass the resource sufficiency evaluation if the resource sufficiency capacity test were changed to be based on load plus load conformance adjustments. With such a change, the capacity required to pass the test would be increased by the amount of the load conformance while supply would only be increased by the amount of additional hourly interchange scheduled. This could create a risk that whenever there was a lot of cheap non-CAISO EIM supply, the CAISO might not schedule enough HASP imports to pass the capacity sufficiency test even though there was enough supply available in HASP to meet the test.[46] This could cause a radical increase in the cost of meeting load and perhaps even an emergency in cases where the scheduled EIM imports would then be unavailable in the binding RTPD intervals because of a sufficiency evaluation failure.

CAISO operators could perhaps avoid failing the resource sufficiency evaluation as a result of the HASP design by scheduling more hourly imports out-of-market (i.e., not modeled in HASP) to ensure that the CAISO could pass the resource sufficiency evaluation. But then the actual 15-minute prices would be lower, making some hourly block transactions uneconomic at FMM prices. Increasing inconsistency between HASP and FMM prices in this manner would likely adversely impact offers and supply in HASP in the future.

We do not have the ability to assess how often the CAISO will be in the position that if it does not schedule HASP imports and instead relies on lower cost EIM imports in HASP, the CAISO will fail the resource sufficiency capacity test. However, given the high use of load conformance adjustments in the evening peak hours, it appears to be that this should be a concern. In addition to the direct effects of increased sufficiency evaluation failures in the real-time market due to such a change in the treatment of load conformance adjustments, this change and the potential for EIM imports to be blocked out in real-time would impact day-ahead market load and virtual supply bids in ways that are difficult to assess This makes it even more difficult to predict the full market and reliability impact of changes that would cause the CAISO to fail the resource sufficiency evaluation when CAISO operators utilized load conformance adjustments.

In addition, it appears that there could be situations in which HASP does not clear day-ahead market transactions on one CAISO interface because real-time prices are higher than was expected day-ahead in that external region, and HASP balances those hourly import reductions with EIM imports on another interface. If changes in the treatment of load conformance adjustments were implemented, the increased potential for the CAISO to fail the resource sufficiency evaluation could cause the CAISO to be unable to balance load in real-time because the day-ahead market imports would not have been scheduled in HASP, and the EIM transfer imports would be capped in RTPD and RTD. It would then be too late when the CAISO fails the resource sufficiency capacity test to commit generation that was not committed in the day-ahead market because the day-ahead market imports were cheaper. The potential for this to happen already exists, but increased sufficiency evaluation failures caused by including load conformance

[46] This would be most serious if CAISO notionally relied on EIM imports in HASP rather than scheduling hourly block RA imports. This will not occur if RA imports are offered at zero, but that is not efficient either.

adjustments in load for the test would increase the potential for adverse reliability impacts in such situations.

One possible solution to these inconsistencies might be for the CAISO to structure HASP to not consider EIM transfers in its solution, and then enforce the capacity sufficiency test requirements in HASP, enabling the CAISO to pass the resource sufficiency evaluation without the need for load conformance adjustments. However, if the CAISO took this approach, the EIM transfers excluded from HASP would be available in the FMM, causing FMM prices to be systematically lower than HASP prices when CAISO was an importer of EIM transfers. This could discourage the submission of HASP intertie offers because of the risk of high losses. This risk does not exist with the use of load conformance adjustments to pass the resource sufficiency evaluation because similar load conformance adjustments can be and are used in the HASP and FMM.

Conversely, when load is high outside the CAISO and at the same time HASP forecasts EIM transfer exports from the CAISO, the current HASP will take account of those EIM transfer exports and potentially commit additional units, keep units on-line, or schedule hourly imports to support those EIM export transfers. If HASP did not take account of those EIM transfer exports, FMM prices would be higher across the EIM when CAISO is exporting EIM transfers.

Finally, as we have noted in prior MSC opinions, there is a conflict between the absolute uncertainty requirement currently applied in the resource sufficiency evaluation and the demand curve applied in HASP and the FMM.[47] This conflict provides a further reason for CAISO operators to apply load conformance adjustments to avoid failing the resource sufficiency evaluation as a result of the HASP economic evaluation clearing less than the amount of flexiramp capacity required to pass the resource sufficiency evaluation.

Some of this conflict is a result of features of the resource sufficiency evaluation that are made less important by CAISO operators use of load conformance adjustments. The CAISO should in any case try to address these issues in the long run, rather than relying on operator load conformance adjustments to manage these problems, but these issues would need to be addressed in the short run if the change advocated by some stakeholders to add load conformance adjustments to the actual load forecast in the resource sufficiency capacity test is made.[48]

These conflicts between HASP economics and the sufficiency evaluation are likely magnified, if not directly caused by, the flaws in flexiramp pricing within the EIM discussed above. When HASP is scheduling EIM transfer imports today, there will be no cost to depleting flexiramp supply in the EIM because the shadow price of flexiramp in the EIM is zero, while the scheduling of hourly transactions will take account the source BA's need to need to pass the resource sufficiency evaluation. This discrepancy would tend to cause HASP to schedule EIM transfer imports because they appear cheap, even when they are not.

---

[47] J. Bushnell, S.M. Harvey, and B.F. Hobbs, Opinion on Market Enhancements for Summer 2021 Readiness, Market Surveillance Committee of the CAISO, March 8, 2021, Section 3.3, pp. 10-11.

[48] Part of the issue with these market outcomes is the failure to implement the flexiramp product as intended. If large EIM imports drove up the flexiramp price and that price was reflected in the energy price, the higher cost of EIM imports would tend to reduce undue reliance on EIM imports in HASP.

5.2 Structure of RTD, STUC, and FMM

Although this has not been discussed in the stakeholder process to date, our review of DMM data in its 3rd Quarter Report on Market Issues and Performance has caused us to consider whether there is a similar inconsistency between the structure of RTD and the resource sufficiency flexibility test that has the potential to cause EIM entities to fail the flexibility test because of the way they are dispatched in RTD, not because they have not provided enough ramp. With the net load uncertainty requirement included in both the capacity test and the flexibility test over the summer of 2021, instances of EIM entities passing the resource sufficiency capacity test but failing the sufficiency evaluation should logically (assuming no as yet undetected flaws in applying the resource sufficiency flexibility test) be a result of differences in capacity counted as available and flexible.

It appears to us that the source of such differences should either be capacity that is not counted towards meeting the flexibility requirement because it is offline with too long a start-up time, or if too much of the unloaded capacity is slow ramping. It is important to recognize that the latter situation could be a result either of the EIM entities' resource mix, or of the EIM dispatch in RTD. The EIM RTD dispatch does not include constraints that require that EIM balancing areas be able to pass the flexibility sufficiency test. As a result, the dispatch would not dispatch low cost, fast ramping generation down, nor slow ramping capacity up to offset this output in order to pass the resource sufficiency evaluation. Hence, a balancing area could fail the resource sufficiency flexibility test because RTD did not dispatch its generation to pass it. If the shadow price of flexiramp were significant in RTD, that would cause RTD to dispatch resources out of merit somewhere to create more ramp--but the RTD objective function would not create ramp in the particular balancing areas that might need it to pass the resource sufficiency flexibility test. Moreover, we know that in the summer of 2021, the price of flexiramp was zero in RTD, so the RTD dispatch would not incur any costs in order to create more unloaded flexiramp capacity.

There are potentially similar issues with the commitment of generation in STUC or FMM to enable a balancing area to pass the resource sufficiency flexibility test. Because the shadow prices of flexiramp in the EIM is consistently zero in FMM, STUC and FMM would not commit short-start generation to meet flexibility needs across the EIM, much less to enable particular balancing areas to pass the flexibility test. In this case, EIM entities might be able to manage this to a degree by self-committing generation not committed in STUC.

We therefore see a potential for the structure of RTD to have led to some of the observed flexibility test failures by BAAs that pass the capacity test. It does not appear to us that there is any good way for an individual balancing area to avoid this outcome other than to try to raise the offer prices of fast ramping capacity relative to slow ramping capacity to keep it unloaded. This is a very ad hoc process that could unnecessarily raise the cost of meeting load and lead to other undesirable consequences. Moreover, if the problems with flexiramp delivery and pricing were fixed, it would not be efficient to apply a balancing area-specific flexibility requirement in RTD if the shadow price of flexiramp across the EIM were zero.

These comments further highlight the linkages between the sufficiency evaluation, load conformance adjustments, and the flawed calculation of the EIM flexiramp price in FMM.

6. Counting Off-line Capacity in Capacity Test

The short-term unit commitment (STUC) process allows for the day-of commitment of units with longer start times utilizing a multi-hour lookahead algorithm. Some resources that would be available on a longer timeframe may therefore be unavailable in the FMM and RTD timeframe if the STUC process anticipated they would be unneeded to meet net load. There are two considerations in deciding how to account for units that can be committed in STUC in applying the resource sufficiency capacity test. On the one hand, as several EIM participants have pointed out, when there is capacity that cannot come on-line in time to provide the needed capacity in the period being tested, that capacity is unable to meet capacity needs in that period. On the other hand, it would be inefficient to commit thermal generation that is unneeded when intermittent resource output is high, and its commitment would either require curtailing zero emission generation or decommitting other resources.

An important function of the EIM is to make full use of intermittent resource output across the EIM footprint. Thus, requiring the commitment of excess thermal generation to meet a sufficiency evaluation during periods of high levels of intermittent resource output would be counter to this important goal of EIM.

Figure 2 in the CAISO Intertie deviation analysis report shows a fair number of sufficiency evaluation failures among EIM entities occur during the middle of the day when there is typically surplus capacity--not shortages of on-line capacity.[49] It would be informative to understand whether the sufficiency evaluation failures of EIM entities in non-ramping hours were failure in tight hours or hours with excess energy. If EIM entities are already failing the sufficiency evaluation in hours with surplus capacity across the WECC, we would not want to make changes that result in more spurious failures of the sufficiency evaluations by not counting off-line capacity that is off-line precisely because it is not needed.

The conflict between these considerations is that while capacity that cannot come on-line in 15 or 30 minutes does not need to be on-line--and perhaps should not be on-line--when capacity is in *surplus* within their region, this capacity should be on-line when capacity is *tight* within the region. However, the resource sufficiency evaluation does not test whether additional capacity is needed given the overall EIM supply demand balance, it just tests balancing area resources against the load forecast and uncertainty requirements. If the resource sufficiency capacity test is intended to test the adequacy of short-term unit commitment decisions, it may make sense to exclude off-line capacity from the test when the supply of capacity is tight. But it is not sensible to exclude off-line capacity from the test and cause EIM entities to fail the resource sufficiency evaluation because they did not commit the capacity when at the same time there is surplus capacity across the EIM, and the capacity was not committed in STUC because it was not needed.

[49] California ISO, "Analysis of the Intertie Deviation Adder used in the Capacity Test," October 6, 2021, p. 6.

The EIM might be able to balance these dual considerations in applying the resource sufficiency capacity test if it were able to only apply the sufficiency evaluations when capacity is needed. For example, the EIM might only apply the sufficiency evaluations when the shadow price of ramp within the balancing area was positive, indicating that less than the full target amount of flexible capacity was available. However, that is not a workable alternative at present because of the flaws in the calculation of the available flexible capacity discussed above that cause the shadow price of flexiramp to always be zero (except when a balancing area fails the resource sufficiency evaluation).

On balance, we recommend continuing to include resources with start times within the STUC time horizon as a test that the balancing area has the necessary resources to meet its load and simply did not commit them, perhaps because they were not needed given system conditions. We realize that this approach has the weakness that a balancing area might be able in some circumstances save commitment costs by not committing generation in uncertain conditions even though the capacity might be needed, but absent better information about the cause of failures during non-ramping hours under the current design, we would not recommend an approach that would likely result in even more spurious failures during what appear to be surplus conditions.

A number of stakeholders have asked for information on the impact of inter-temporal constraints such as start times on the available capacity during high load hours on high load days. That kind of analysis would be useful to understand if the treatment of units not committed in STUC had a material impact on the outcome of the resource sufficiency evaluation on these days. Such a metric would provide an indicator of how urgent it is for the EIM to find a better resolution of these issues. In addition, before making a change that would exclude capacity not committed in STUC from the resource sufficiency capacity test, stakeholders also need the CAISO to assess the impact of such a change on the ability of EIM entities to pass the resource sufficiency capacity test on days when the following conditions apply: the net load was not particularly high, the units would not have been needed to meet load, and their commitment at minimum load might have displaced zero emission output. The data now posted by DMM could perhaps inform this assessment with calculations of the number of sufficiency evaluation failures by hour of the day (as in Figure 2 of the intertie deviation report) and by calculations of how many resource sufficiency capacity test failures occurred when energy prices were below a threshold value.

7. Requirement for Transmission Tags at t-40

There is a concern with the inclusion of untagged intertie transactions in the supply used to apply the RSE's capacity test, reflecting a higher potential for these transactions to not flow in real-time than tagged transactions. The CAISO has presented data on the number of CAISO import transactions cleared in HASP that were not tagged at t-40 and were also not tagged at t-20 and therefore did not flow in real-time. This data is relevant to assessing how many intertie transactions that were included in the resource sufficiency capacity test but that did not actually flow in real-time. However, it is also important to understand how many transactions that were not tagged at t-40 were tagged at t-20 and did in fact flow. This data would be relevant to assessing how many transactions that did flow in real-time would be excluded from the resource sufficiency capacity test with the proposed change. Based on our conversations with the CAISO we understand that it is rarely the case that transactions that are not tagged at t-40 are tagged by t-20

and flow in real-time, but we have not reviewed any data on this point. If this is the case, modifying the sufficiency evaluation in the manner proposed should not cause the CAISO to fail the resource sufficiency capacity evaluation when it should not.

A related consideration is that there appear to be weak incentives for a market participant to tag an import transaction prior to t-40 rather than between t-40 and t-20 if the transaction is intended to flow.[50] If the market participant waits until sometime after t-40 to tag the transaction, the initial 15 minutes of flow would be priced in RTD, rather than FMM, and RTD prices are typically lower than FMM prices. In addition, the transaction would be at risk of being cut by the operators if the intertie were constrained, and perhaps also if operators were concerned that it would overload internal constraints. In practice, these considerations appear to be sufficient to incent market participants to tag transactions that will flow prior to t-40, in light of the CAISO's assessment that very few transactions that are not tagged at t-40 flow in real-time.

With the proposed change it will be important that market participants continue to tag transactions that expected to flow prior to t-40 as transactions tagged after t-40 would not be included in the resource sufficiency capacity evaluation, potentially causing the CAISO to fail the capacity evaluation despite actually having enough capacity in FMM. This failure would tend to hurt CAISO net buyers in the FMM, as EIM transfers could be capped at a level lower than cleared in HASP, leading to higher prices. If this were to occur, the higher RTD and FMM prices would tend to benefit, not hurt, the import supplier that waited until after t-40 to tag an hourly import transaction. Hence, if this change is implemented, the CAISO should monitor tagging behavior to identify any material increase in untagged transactions at t-40 that flow in real-time.

Finally, while we support implementing such a rule excluding untagged imports from the resource sufficiency capacity evaluation while the intertie deviation requirement is being redesigned, it appears that the two measures are duplicative, and there should not be a need to have them both in place at the same time.

8. Failures in Emergency Conditions

There are two sets of issues in the situation in which the system is in an emergency condition and shedding load, or in high risk of doing so. One is the issue regarding NERC rules concerning management of type 2 energy emergency alerts (EEA2) versus balancing area-specific rules for triggering a failure of the resource sufficiency capacity evaluation.

A second issue is NV Energy's concern with the concept of freezing EIM transfers when an entity enters an EAA2 situation.[51] Their concern is similar to the concern we have expressed in the past regarding whether freezing EIM transfers is the appropriate penalty for failing the resource sufficiency evaluation.

---

[50] Note that we are not referring to penalties for transactions that do not flow, which should be provided by the intertie deviation charges. We are referring to penalties for waiting until after t-40 to tag transactions that do flow.

[51] See NV Energy, Comments on Draft Final Proposal, Jan. 10, 2022, items 1, 6 and 8, https://stakeholdercenter.caiso.com/Comments/AllComments/5293f8fd-ab86-47ea-9168-3ed6dd8b8a68.

The latter issue is not part of this phase of the initiative, but will be important to address in the near future.

We believe that the long run design should not freeze EIM imports for BAs in emergency conditions. It would be inappropriate for a balancing area to shed load because it cannot schedule emergency transfers fast enough at a time when plenty of supply was available and could have been dispatched in the EIM-wide RTD. In the long run there should be financial penalties for failing the resource sufficiency evaluation rather than imposing a hard constraint on EIM transfers. A financial penalty-based approach would avoid the outcome in which a balancing area authority is barred from importing power via EIM transfers even when load is being curtailed and its willingness to pay is very high. The financial penalty could be high to make it costly but not prohibitive to fail the resource sufficiency evaluation when there is sufficient capacity to meet load across the EIM. To do otherwise would unduly diminish the benefits that could be realized from EIM.

9. Conclusions

In order to continue to improve the resource sufficiency evaluation, several items need to be addressed, some of which are interrelated. These include aspects of the EIM's design and implementation that impact the RSE, as well as elements of the RSE itself.

First, the CAISO needs to fix the pricing of flexiramp in the FMM and in RTD. The flaws in the current implementation of the flexiramp product in HASP, FMM and RTD that result in zero prices for flexiramp in regions in which it is actually very scarce are a critical issue that underlie several problems involving the resource sufficiency evaluation. The nodal deliverability test the CAISO is developing is intended to address the portion of this problem that arises because within BAA congestion is presently disregarded when procuring flexiramp, and we understand that this test is intended for implementation in fall 2022. However, the actual implementation date is uncertain until it is known how well it is likely to work. We believe that both the implementation timing, design, and performance of the nodal delivery test should be reviewed with stakeholders.

Depending on the outcome of this review it may be appropriate to consider alternative interim, or even long-run designs. This is particularly important for regions outside the CAISO in the near term in order to provide an appropriate price signal for scheduling EIM transfers in FMM. The CAISO also needs to review the concerns that caused it to not implement the minimum locational requirement for the CAISO in RTD and consider ways to address these concerns.

Second, it is important that the CAISO replace the current approach for calculating the resource sufficiency requirements with better methods. We understand that the quantile regression method is intended to replace the current methodology for calculating the net load uncertainty requirement, and we have been informed that the CAISO expects to implement the quantile regression method in conjunction with a nodal delivery test in fall 2022. However, it is not clear to us or stakeholders how these designs will likely work. We have pointed out above that the use of a small sample size can lead to anomalous results with the quantile regression methodology, just as it does with the current approach, so it is important to understand more about the proposed

methodology and assess how well it will perform for EIM entities as well as the CAISO.[52] It is our understanding that the CAISO plans to develop a new method for determining the intertie deviation requirement in a stakeholder process.

As part of the discussion of the calculation of these requirements we think it is important that the CAISO provide data portraying the day-to-day variability of the current measures so that stakeholders can assess how important the statistical issues we discuss in this opinion have been in practice. This is relevant both to assessing the potential for similar sample size impacts with a quantile regression methodology and to help understand if these sample size issues are the core problem leading to inappropriate sufficiency evaluation failures today, or if instead there are deeper problems.

Third, it is important that the CAISO address the inconsistency between the HASP evaluation and the resource sufficiency capacity test for the CAISO, and similarly that the CAISO also address the apparent inconsistency between the RTD dispatch and the resource sufficiency flexibility test for both the CAISO and other EIM entities. The solution to the problems created by these inconsistencies is related to the first problem, the pricing of flexiramp. The HASP and RTD will of course choose solutions that result in sufficiency evaluation failures if the price of ramp in FMM and RTD is zero for the EIM entities, because the software would incur zero costs to increase the supply of flexiramp through either redispatch or committing generation. It would be possible to address this inconsistency by including an additional set of constraints in HASP, FMM and RTD. It may be that this would be the best approach, but such a resolution would add complexity and solution time to these unit commitment and dispatch tools and has the potential for unintended impacts as discussed above.

Fourth, sufficiency evaluation penalty failures are a subject for Phase 2. We agree with stakeholders who have pointed out that capping EIM transfers at the current level can be an ineffective penalty for failing these tests if transfers were already high at the time of the sufficiency evaluation failure. On the other hand, extreme penalties for sufficiency evaluation failures are not appropriate until the issues identified under items two and three above are addressed so that high penalties are not imposed for spurious failures. We also agree with NV Energy's viewpoint[53] that the Western EIM FMM and RTD are potentially powerful tools for balancing load across the WECC to use in maintaining reliability in uncertain operating conditions and placing restrictions on their use during unexpected operating conditions has the potential for unacceptable outcomes. We do not think it would be an intended nor acceptable outcome to require an EIM entity to shed load that could have been met with EIM transfers when there was not enough time to schedule hourly base imports or even to arrange emergency imports.

Another consideration in developing an improved penalty design is that most or all of the problems identified under the third topic would be avoided if the RSE is included as constraints with relaxation penalties in the FMM and RTD unit commitment and dispatch, so there would not be

[52] The CAISO has made a good start on doing so in its January 2022 workshop on the day-ahead market enhancements (G. Bautista-Alderete and K. Zhao, "Day-Ahead Market Enhancement Analysis", www.caiso.com/Initiative-Documents/Day-AheadMarketEnhancementsAnalysisReport-Jan24-2022.pdf.

[53] NV Energy, op. cit..

inconsistencies between the market solution and the resource sufficiency evaluation requirements.

Fifth, we do not agree that load conformance adjustments by the CAISO or other entities should always be added to load in applying resource sufficiency evaluations. If there is an issue with a balancing area's load forecast consistently understating its real-time load, that outcome should be addressed. But there is no reason to artificially inflate the load used in the resource sufficiency evaluation if the currently used values are in fact unbiased. Moreover, while some of the pattern of CAISO EIM transfers in FMM and RTD that have occurred over the past year is likely due to the load conformance adjustments by CAISO operators, this pattern is also a result of the flawed pricing of flexiramp in FMM and RTD--particularly the zero shadow price of flexiramp in the EIM at times when the price of flexiramp is positive, and sometimes relatively high, in the CAISO.

As a general conclusion, the use of resource sufficiency evaluations to limit participation in the Western EIM's real-time dispatch is, in general, reasonable and analogous to the resource adequacy requirements historically imposed for participation in tight power pools. The presence of such a requirement for participation in the west-wide power real-time dispatch would enhance balancing area authority confidence that they would not be adversely impacted by undue leaning on the pool by other balancing areas and thereby increase their willingness to participate in the west-wide market. Wide participation is necessary for full realization of the potential benefits of the Western EIM market. However, the current penalty structure may fail to consistently deter undue leaning for several reasons. These include instances in which EIM transfers are capped at very high levels; when EIM transfers are inappropriately capped due to anomalous requirements for balancing capacity; or when the Western EIM real-time dispatch would help WECC balancing areas maintain reliability when system conditions are materially different than expected, potentially harming reliability by blocking transactions that could prevent load curtailment. At the same time, however, willingness of balancing areas to participate in the Western EIM requires assurance that balancing areas will not be required to shed load because of another balancing area's inadequate resources. Well-designed enhancements of the resource sufficiency evaluation process should encourage participation, increase cost-effective power exchanges across the west, and improve reliability by providing appropriate incentives to coordinate dispatch and at the same time discourage leaning.