New York ISO
Independent System Operator

# Standard Error Analysis

Bianca Prinsloo

**ICS Meeting #309:**
November 5, 2025

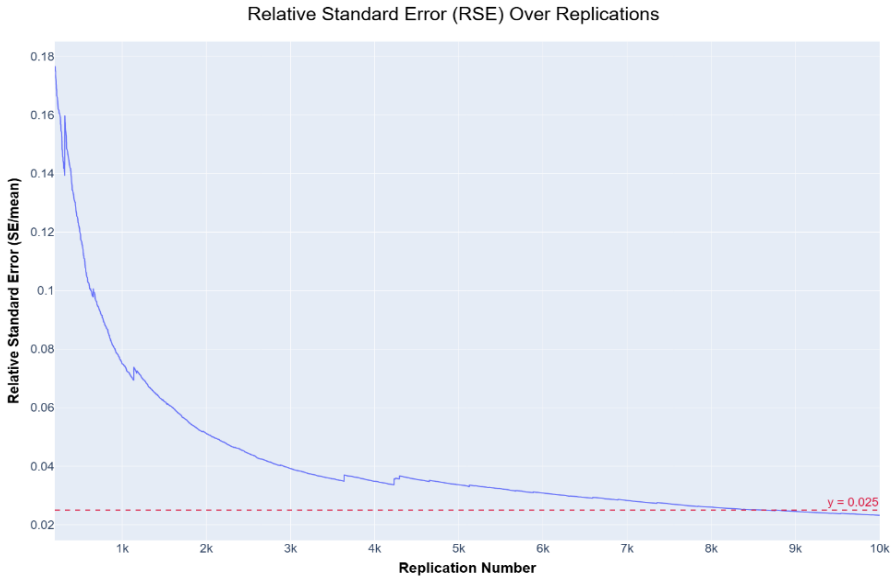
# Background

- Under NYSRC Policy 5-19, Section 3.8, the standard error of the Installed Reserve Margin (IRM) should be 0.025 of the mean loss of load expectation (LOLE), consistent with a 95% confidence level.
- With increased modeling changes in recent years of the IRM study, the number of replications to achieve targeted standard error has been increasing.
  - At the conclusion of the Preliminary Base Case (PBC), the standard error requirement was achieved at 1,865 replications and therefore replications were reduced to 2,000.
  - At the conclusion of the Final Base Case (FBC) parametric results, a standard error requirement was achieved at 8,750 replications.
- In order to complete the 2026-2027 IRM Final Base Case Tan45 for this ICS meeting, NYISO produced the Tan45 at 3,000 replications which achieved a 94% confidence level.
- Conducting a Tan45 case at 8,750 replications could not be completed in advance of this ICS meeting.
  - The NYISO is unable to guaranty that it could complete Tan45 cases with 8,750 replications in a timely manner to meet NYSRC deliverables in November and December.
- A resolution is needed in order to complete the 2026-2027 IRM study on time and meet NYSRC deadlines.

New York ISO

# Final Base Case Standard Error

- **Significant number of replications is needed to meet the target standard error of 0.025, indicating inefficient convergence due to the high variance introduced by the heavy-tailed LOLE distribution.**
  - This plateauing effect is typical in heavy-tailed distributions, where a few extreme values inflate the variance.
- **The presence of certain LOLE spikes within the tail of the distribution requires many more replications to achieve small reductions in the relative standard error.**
  - This suggests that relying on standard error as a convergence metric may be misleading or inefficient in this context.




New York ISO

# Final Base Case Distribution of LOLE

- **The distribution of LOLE in the final base case is heavy-tailed (see Figure 2 in the Appendix), with most replications showing low LOLE values.**
  - Of the 10,000 replications tested, approximately 9,000 replications have LOLE values less than 0.25.
    - The average LOLE of all replications that fall less than to 0.25 LOLE is 0.05, indicating LOLE values of <0.1 in a significant number of replications.
  - A small number of replications exhibit LOLE values greater than 3, representing rare but extreme outcomes.
  - These extreme values inflate the variance (see Figures 2 and 3 in the Appendix), making it unstable and causing the convergence rate to slow down.
- **Due to the high variance and skewness exhibited by the final base case, calculating a standard error may not be the most appropriate measure of uncertainty.**
- **Alternative metrics or robust statistical techniques might be better suited for characterizing uncertainty in such heavy-tailed distributions.**
  - Short term: To accommodate the time constraints of the 2026-2027 IRM study, the NYISO proposes to temporarily lower the confidence level, allowing the analysis to converge within a reasonable number of replications.
  - Long term: Alternative metric(s) will be evaluated as part of the parametric improvement study, aiming for a more stable and representative measure of uncertainty.

New York ISO

# Runtime for Tan45

- **Each point for the Tan45 process requires anywhere from 40-60 hours to finish all of the iterations**
  - For a 3,000 replication Tan45, the NYISO can run all 12 points simultaneously
    - Overall, the process can take anywhere from 60-80 hours to complete without interruption
  - For an 8,750 replication Tan45, this exceeds the NYISO's core limit, restricting the number of points that can be run simultaneously.
    - This increases the process time significantly to anywhere from 140-200 hours to complete without interruption
- **The Tan45 using 3,000 replications that is presented at today's ICS meeting required a week to complete due to multiple interruptions.**
  - Increasing the replications will increase the time needed to run a case, making the case more prone to interruptions
- **Increased replications beyond 3,000 increases the risk of failure to provide key results in time for NYSRC deliverables.**
- **NYISO can parallelize the MARS simulations among several users to speed up time to completion, however, significant risk exists for job failures.**
  - NYISO cannot guarantee its ability to complete an additional Tan45 beyond 3,000 replications prior to the November 14th Executive Committee meeting at which Final Base Case acceptance is expected to occur.

New York ISO

# Impact Analysis

- To evaluate the potential impact on the IRM of lowering the confidence level (CI), the number of replications for the Preliminary Base Case was reduced to meet at least a 90% confidence level instead of 95%.
- NYISO concluded that reducing the confidence level to 90% for the Preliminary Base Case—which requires fewer replications—produced no impact on the IRM.
  - See Figure 1 in the Appendix for additional details.
- NYISO also concluded that increasing replications for the Final Base Case (parametric results) from 2,000 to 8,750 to meet the confidence level of 95% produced no impact on the IRM.
  - See Figure 1 in the Appendix for additional details.
- NYISO conducted such analysis through parametric results. NYISO was unable to complete an impact analysis for the Tan45 results due to computational and time constraints.
- The findings of this analysis support that using a lower confidence level is a practical and reasonable short-term solution for the 2026-2027 IRM Final Base Case, where achieving 95% confidence is computationally infeasible due to the heavy-tailed nature of the LOLE distribution for the case.

New York ISO

# Summary

- NYSRC Policy 5-19 requires the IRM study to meet a 0.025 standard error/95% confidence level for the final base case prior to the Tan45.
- For the 2026-2027 IRM study, achieving this requirement would increase the number of replications to 8,750 introducing significant risk to the ability to complete the Tan45 process the study on time.
    - Parallelizing the jobs across users can reduce time to completion, however, Amazon Web Services (AWS) constraints including core limits and outages further increase the processing time.
- In order to complete the 2026-2027 IRM study on time, NYISO produced a Tan45 at 3,000 replications which achieves a 94% confidence level. The results of this Tan45 are provided at this meeting.
- Increasing replications to meet the target standard error is leading to significant challenges and volatility indicating that this metric may not be well suited for the Final Base Case due to the heavy-tailed LOLE distribution.
- Additionally, NYISO-conducted analysis indicates that increasing the number of replication beyond 2,000 (i.e., the value identified for the Preliminary Base Case) is unlikely to materially affect the IRM.
- NYISO recommends considering an exception for the standard error/confidence level if the number of replications are excessive, to facilitate timely completion of the IRM study process and meet NYSRC deadlines.

New York ISO

# Next Steps

- **Pending a recommendation from ICS, the NYISO has completed the Final Base Case Tan45 at 3,000 replications which will achieve at least 90% confidence level**
  - 3,000 replications meets a 92% confidence level prior to the Tan45 and 94% after the Tan45
  - The Tan45 results posted for today's meeting were completed with the 3,000 replications
- **Update NYSRC Policy 5-19 to allow for this short-term solution.**
  - To be initially discussed at this meeting with the intent to finalize at the 11/11/2025 ICS meeting
  - This timeline is intended to facilitate review by the Executive Committee at its 11/14/2025 meeting
- **Explore options for longer-term solutions as part of the Parametric Improvement Whitepaper in early 2026.**

New York ISO

# Questions?



New York ISO

# Appendix



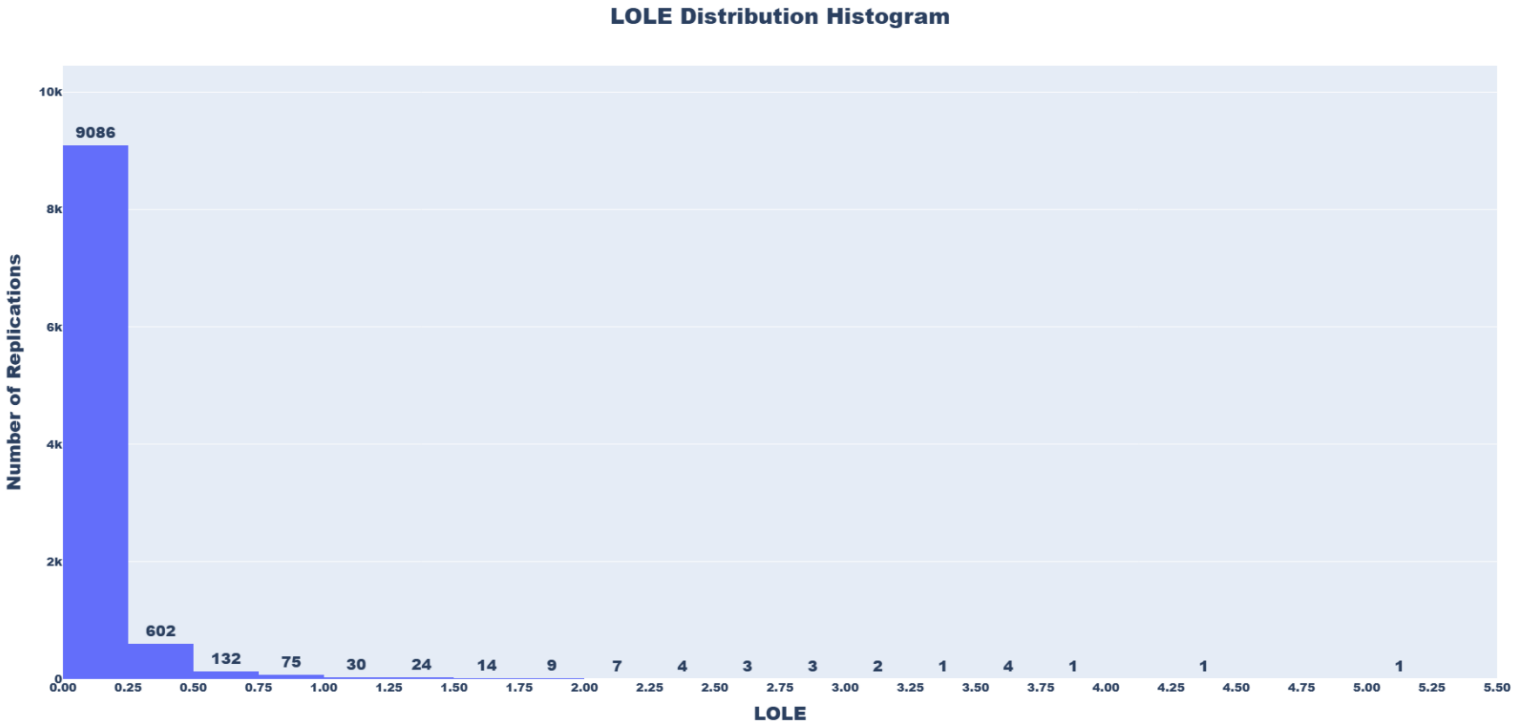
New York ISO

# Figure 1: IRM Impact Analysis from Different Replications/Standard Error

| Case | CI | N Replications | Standard Error at N replications | IRM | Delta from 95% CI |
|---|---|---|---|---|---|
| **2026-2027 Preliminary Base Case** | | | | | |
| 2026-2027 PBC Parametric | 95% | 2,000 | 0.024 | 25.741 | - |
| 2026-2027 PBC Parametric | 93% | 1,000 | 0.035 | 25.741 | ***0.00*** |
| **2026-2027 Final Base Case** | | | | | |
| 2026-2027 FBC Parametric | 95% | 8,750 | 0.025 | 27.490 | - |
| 2026-2027 FBC Parametric | 92% | 3,000 | 0.039 | 27.468 | ***-0.02*** |
| 2026-2027 FBC Parametric | 90% | 2,000 | 0.051 | 27.490 | ***0.00*** |

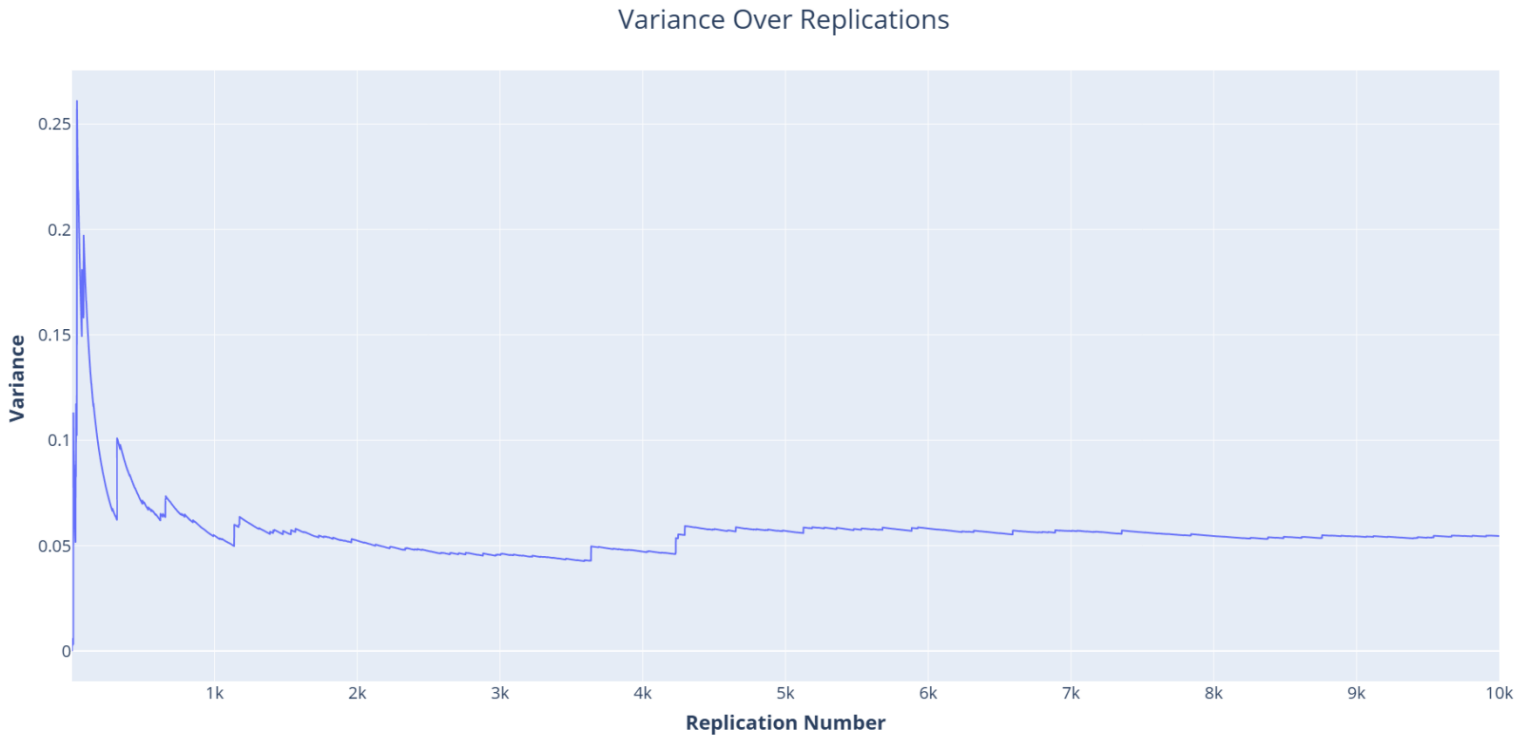
New York ISO

# Figure 2: 2026-2027 Final Base Case Heavy Tailed LOLE Distribution




New York ISO

# Figure 3: Variance Instability




lew York ISO

# Policy 5-19, Section 3.8

### 3.8 Standard Error

Another step in assuring a quality result is to determine whether the standard error is acceptable. The MARS model is run for a set number of iterations at increments of 250. Ideally, the standard error value remains less than 0.025 throughout the entire IRM Study. However, to provide a quality result, the ICS has determined that the desired standard error value for the mean Loss of Load Expectation (LOLE) at the 95% confidence level shall be less than or equal to 0.025 at the final iteration at three critical points; a) the beginning of the IRM Study; b) at the conclusion of the Preliminary Base Case prior to the Tan 45 process; and c) at the conclusion of the Final Base Case prior to the Tan 45 process. If the standard error is not met at these critical points, the number of iterations is increased by 250 iterations until the condition is met. The MARS model is then returned to criteria as necessary and the results reported to ICS as part of the parametric analysis of the respective base cases. The Tan 45 process then proceeds as planned for either the preliminary or final Base Case.

By default, the standard error at the beginning of the IRM study is met as the study begins with the final Base Case from the prior year IRM study. However, if the number of iterations has become excessive with a corresponding standard error well below the 0.025 standard error value, it may be appropriate to decrease the number of iterations. In general, the more iterations used by MARS, the better the convergence and the better the confidence in the result. Unless the MARS runtime is seriously impacted, there is no technical reason to reduce the number of iterations necessary to achieve the standard error value.

New York ISO

# Our Mission and Vision

## Mission

Ensure power system reliability and competitive markets for New York in a clean energy future

## Vision

Working together with stakeholders to build the cleanest, most reliable electric system in the nation

New York ISO