

Data Centers

Treatment as Intermittent Sources and Use
of Monte Carlo Simulation for
NAAQS Modeling Evaluations in Colorado

Nathaniel Collett, Dispersion Modeling Section

September 2025



COLORADO
Department of Public
Health & Environment

Intermittent Sources – EPA Guidance

- 2010 EPA memo¹ address the complications of intermittent sources in the context of the new (at the time) 1-hr NO₂ standard.
- The hierarchy of modeling methods (in order of most to least conservative) based on this EPA guidance is:
 - Model intermittent sources as operating continuously with an emission rate based on maximum short-term PTE.
 - Model intermittent sources as operating continuously with an emission rate based on an annualized rate (e.g. PTE * 500 hr-op/8760 hr-total).
 - Omit intermittent sources from modeling.
- Memo was written in the context of fairly small emission units (e.g. fire pump emergency engine) at major source facilities that became the controlling source for NAAQS compliance. Given the probabilistic form of the new 1-hr NO₂ NAAQS it was unlikely that the actual operations of the intermittent source would consistently align with the model hours resulting in high impacts.

¹Fox, Tyler; Additional Clarification Regarding Application of Appendix W Modeling Guidance for the 1-hour NO₂ ,National Ambient Air Quality Standard; U.S. Environmental Protection Agency; Office of Air Quality and Planning; Research Triangle Park, North Carolina, 2010.

Data Centers as Intermittent Sources

- Given the nature of data centers, emergency generators make up all or a large majority of total facility emissions, DMS does not believe it appropriate to omit the emergency engines when demonstrating NAAQS compliance.
- Similarly, EPA's guidance around intermittent sources appears to be geared more towards single, or at most a handful of, infrequently operating sources given that annualizing the emissions over the year does not differentiate between timing of source operations.
- Recent presentations/comments by EPA (e.g. 2025 RSL Modelers' Workshop) support that the 2010 memo is not intended for use with data center type facilities. Further guidance on intermittent sources is expected to be released eventually but how long that will take is unknown.

Data Centers as Intermittent Sources - Annualization



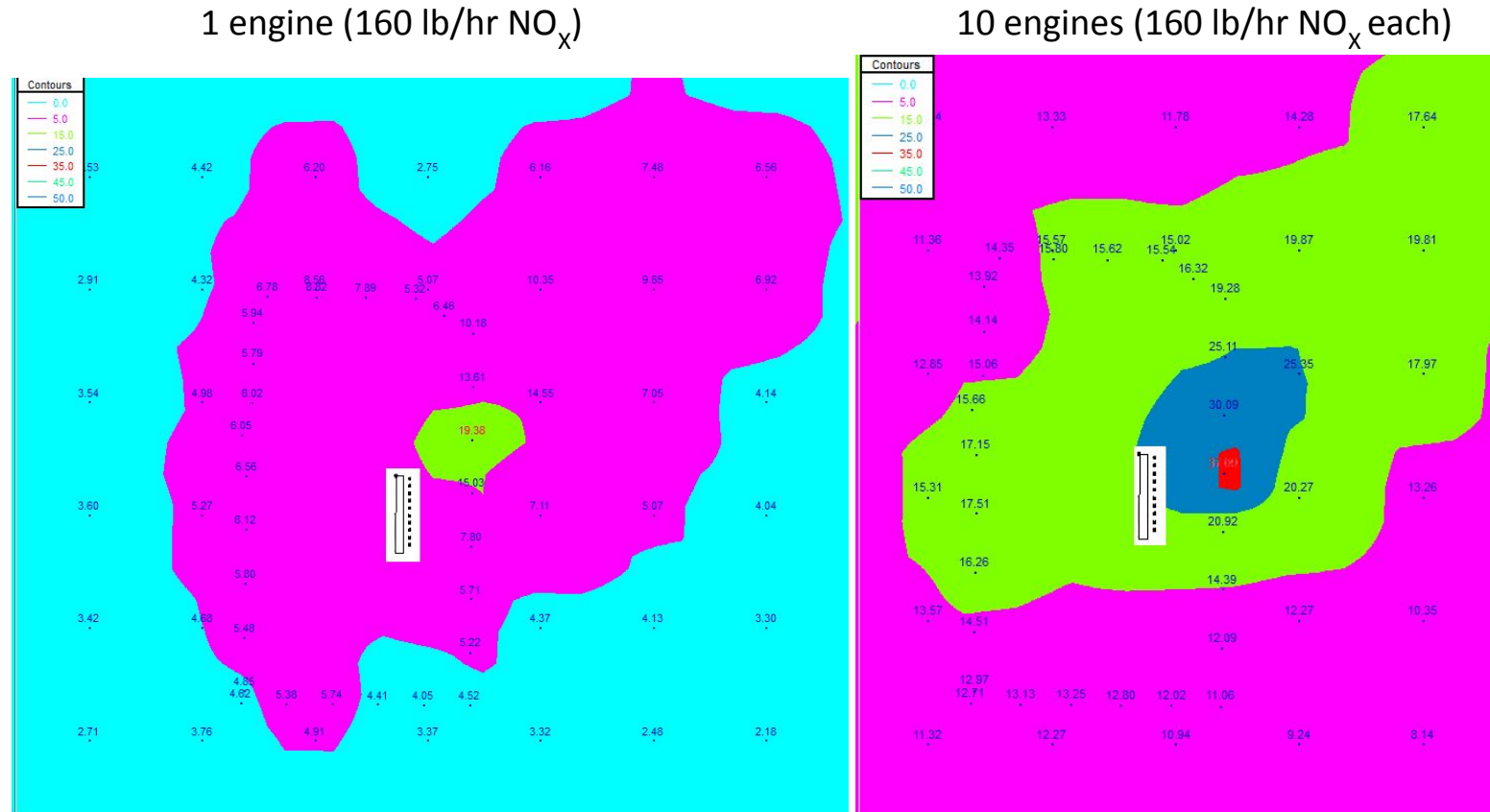
DIA Meteorology & includes Denver area Seasonal NO2 background

- Main concern with annualization is the removal of an evaluation of concurrency from the model. Allowed for non-concurrent operation
- Ten engines operating individually for 50 hours a year each (500 hours total) annualizes the same as ten engines operating concurrently for 50 hours a year (still 500 total engine hours)
- Latter case has higher single hour PTE so even more favorable meteorological conditions may still result in modeled exceedances.



Data Centers as Intermittent Sources – Exceedance %

- 1 engine at a time has modeled exceedances of the 1-hr NO₂ NAAQS for 19.4% of the possible hours in the year.
- 10 engines at a time has modeled exceedances of the 1-hr NO₂ NAAQS for 37.1% of the possible hours in the year.
- Based on emission rates of the hypothetical source the proportion of hours with poor outcomes is likely high enough that neither would pass under Monte Carlo



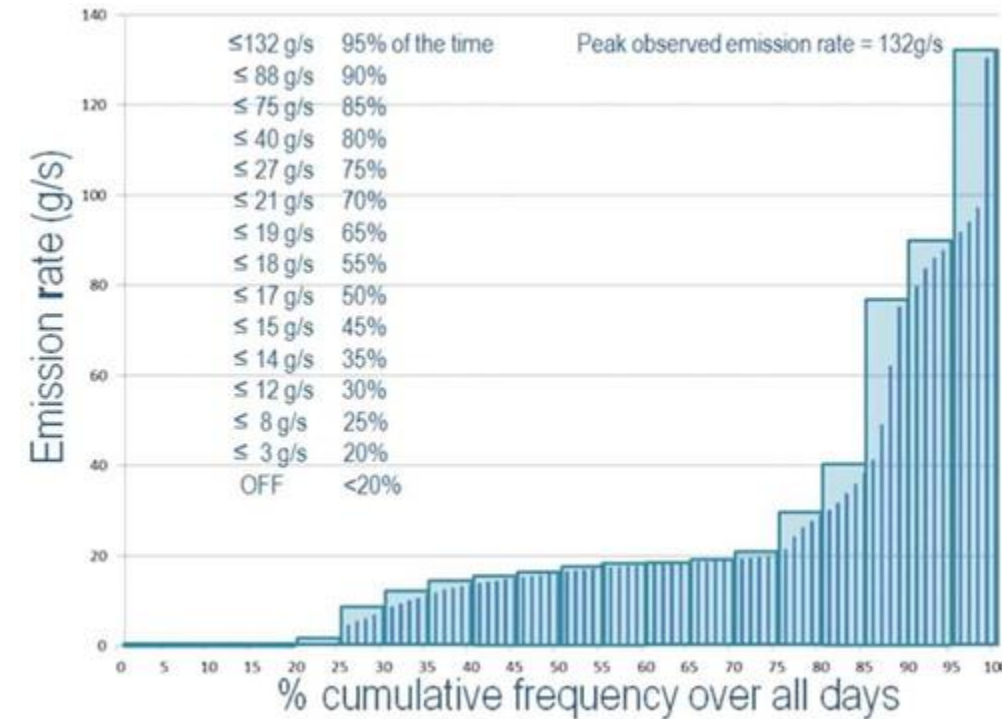
DIA Meteorology & includes Denver area Seasonal NO₂ background

Intermittent Sources – DMS as Reviewing Authority

- Using a Monte Carlo approach to evaluate the likelihood of a NAAQS exceedance based on the engine PTE and operational schedule would be more robust for multi-engine situations like data centers.
- For minor source projects where CDPHE is the reviewing authority DMS is currently allowing for the use of Monte Carlo to demonstrate NAAQS compliance of intermittent sources.
 - Inputs to Monte Carlo evaluation are incorporated as permit conditions (e.g. facility may operate 6 engines at any one time for up to 50 hours in a calendar year).
- The hierarchy of modeling methods (again in order of most to least conservative) would then be:
 - Model intermittent sources as operating continuously with an emission rate based on maximum short-term PTE.
 - *Model intermittent sources using Monte Carlo (placement has been true for DMS reviewed projects so far).*
 - Model intermittent sources as operating continuously with an emission rate based on an annualized rate. (may still be acceptable for single/non-concurrent engine operation)
 - Omit intermittent sources. (Refer to Minor Source Guidelines, typically not data centers)

Monte Carlo – General Procedure

- Develop emission bins and number of operating days (or hours) per year for each emission bin.
- Run AERMOD and write POSTFILE.
- Loop through the output POSTFILE concentrations:
 - Sample the distribution according to the defined modes and random selection for each emissions bin.
 - Calculate daily maximum concentrations.
 - Compute H8H or 98th percentile for each sample from daily maximums.
 - Repeat the loop many times.
- Extract 50th percentile, average, or median from all samples correlating to the form of the standard.



Paine, R. *et al.* (2014) 'Emissions Variability Processor (EMVAP): Design, evaluation, and application', *Journal of the Air & Waste Management Association*, 64(12), pp. 1390–1402.

Monte Carlo and Existing Code Routines

- There are currently two more established code routines for use with Monte Carlo on AERMOD Outputs, one from the State of Washington and the other from EPRI.
 - State of Washington distills impacts to daily-high before random sampling, allows for use of EMISFACT keyword, and runs years independently
 - EPRI consists of EMDIST/EMVAP/EMPOST, allows for unitized emission rates, can process up to 5 POSTFILES as input to EMVAP
- DMS does not prefer one code routine over the other, though we suggest applicants intending to use Monte Carlo contact/email the Division to discuss their proposed approach as the input constraints will be converted to permit conditions.
- Monte Carlo is only effective in demonstrating modeling compliance when both of the following are true:
 - Operation of the source is relatively infrequent.
 - Hours contributing to an exceedance are not common (correlated with engine emission rates and count).

Monte Carlo Reviews by DMS

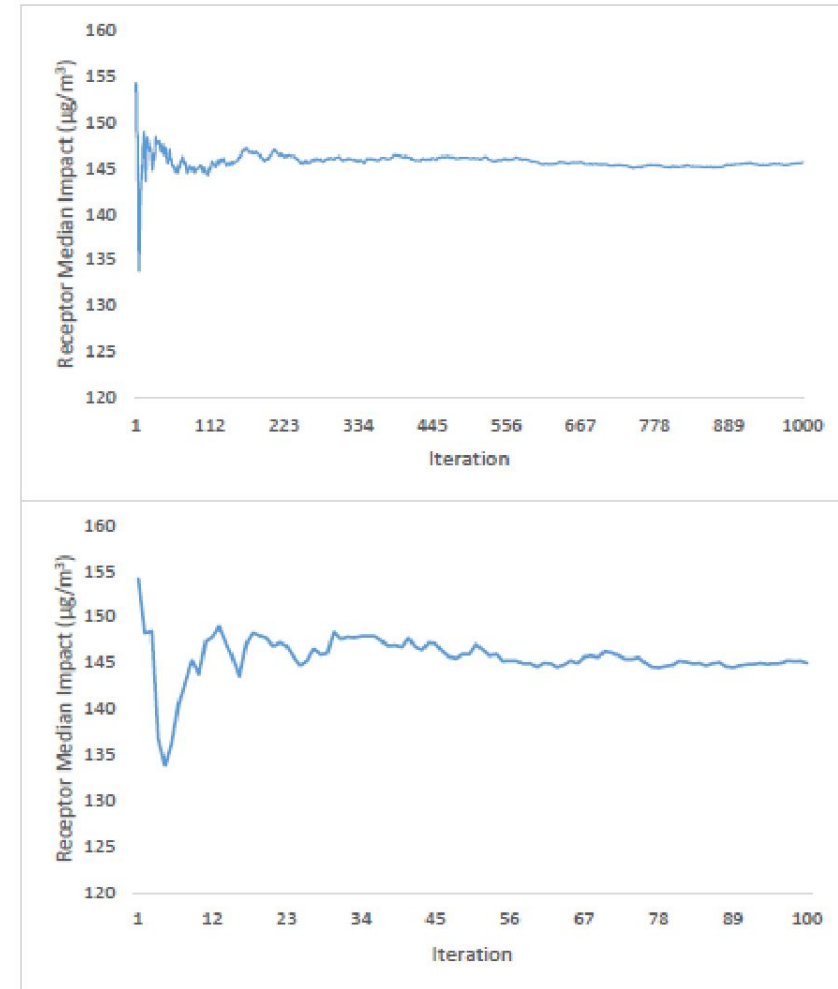
- The Monte Carlo submissions that DMS has reviewed to date have not exactly aligned with either of these code routines. The most common submission has been random sampling based on an R script that most closely matches that of the State of Washington with the exception that the modeled impacts are not distilled to daily averages before extracting the random sample. As a result the associated permit condition remained based on hours of operation per year versus number of days operating per year.
- The next slides summarize additional checks that DMS typically performs on the modeling submissions in addition to the usual review process for an AERMOD submission.

Monte Carlo – Form of the Standard

- Technically, the form of the 1-hr NO₂ standard is the average impact at the receptor and most Monte Carlo scripts are returning the median.
- For sufficiently large datasets the median approaches the mean.
 - DMS has been calculating the mean result based on the submitted results as well and so far there hasn't been a situation where the median passes but the mean does not (e.g. median just below 188, mean just above 188). In the event this does occur at some point, additional discussion on what the Division would accept may be necessary.

Monte Carlo – Solution Convergence

- DMS plots the median value of the results to visually evaluate if it appears the true expected average impact from the random schedule has been determined.
- 100 to 1000 iterations was recommended by the SC 1-hr NO₂/SO₂ workgroup.
- EPA used 100 iterations for the Monte Carlo guidance associated with SO₂ SIP modeling.
- The submissions reviewed by DMS have most often been 1000 iterations.



Monte Carlo – Randomness

Month	Occurrences	Percentage
01	2029	8.5%
02	1770	7.4%
03	2056	8.6%
04	1902	7.9%
05	2073	8.6%
06	1898	7.9%
07	2046	8.5%
08	2165	9.0%
09	2024	8.4%
10	2073	8.6%
11	1942	8.1%
12	2022	8.4%

Hour	Occurrences	Percentage
01	0	0.0%
02	0	0.0%
03	0	0.0%
04	0	0.0%
05	0	0.0%
06	0	0.0%
07	0	0.0%
08	2000	8.3%
09	1932	8.1%
10	1935	8.1%
11	2015	8.4%
12	2125	8.9%
13	1970	8.2%
14	2025	8.4%
15	1983	8.3%
16	2022	8.4%
17	1974	8.2%
18	2039	8.5%
19	1980	8.3%
20	0	0.0%
21	0	0.0%
22	0	0.0%
23	0	0.0%
24	0	0.0%

Day	Occurrences	Percentage
01	739	3.1%
02	774	3.2%
03	819	3.4%
04	781	3.3%
05	820	3.4%
06	845	3.5%
07	835	3.5%
08	787	3.3%
09	772	3.2%
10	767	3.2%
11	754	3.1%
12	780	3.3%
13	819	3.4%
14	802	3.3%
15	771	3.2%
16	779	3.2%
17	762	3.2%
18	753	3.1%
19	834	3.5%
20	770	3.2%
21	817	3.4%
22	810	3.4%
23	797	3.3%
24	806	3.4%
25	766	3.2%
26	781	3.3%
27	788	3.3%
28	772	3.2%
29	742	3.1%
30	699	2.9%
31	459	1.9%

- DMS calculates metrics for the frequency of datetimes in the supposed random sampling to identify if the dataset is unbiased.
- A bright-line threshold has not been established for the deviation, but DMS checks that the values appear consistent.
 - Some variation is expected for calendar differences.
 - Check that the random extract occurs in the represented operational period.

Monte Carlo – Verification of Results

- The POSTFILE from AERMOD is often output in a binary format (size constraints) so is not human readable.
- Re-running the R script could be useful to demonstrate that a different random extract still complies with the NAAQS but would not validate the results submitted by the applicant.
- DMS has been identifying the iteration corresponding to the reported modeled impacts (and another randomly selected iteration) and creating hourly emission files for the sources corresponding to the random datetimes for those iterations. AERMOD is run with the hourly emission files and the output is compared to the applicant reported results for those iterations.
 - Based on the reported receptor-by-receptor results for those verification runs agreeing with the applicant report results DMS concludes that the Monte Carlo R script is not significantly altering the impacts output by AERMOD.
 - If the full set of 1000 iterations were converted to hourly emission files and run through AERMOD (ignoring review time constraints) it is expected that the calculated median impact from the Monte Carlo script would match the result of running the 1000 iterations individually in AERMOD.

Conclusions

- Accurate representation of intermittent sources, most notably data centers, in AERMOD is an ongoing concern.
- DMS believes that the 2010 EPA memo on intermittent sources is not appropriate to use for facilities consisting mainly of intermittent sources (e.g. data centers) where multiple units are operated simultaneously.
- DMS has been accepting the use of Monte Carlo evaluation to reflect randomized operation based on AERMOD outputs and recommends pre-application discussion with the modeling team.
- DMS will consider any additional EPA guidance once released.

Thank you!

Work Email:

nathaniel.collett@state.co.us

DMS Email (for Modeling/Monte Carlo questions):

cdphe_apcd_modeling_files@state.co.us



COLORADO
Department of Public
Health & Environment