

EPA Method 8000D initial calibration criteria, explained.

Method 8000D sets the calibration rules behind every SW-846 chromatographic method. Passing r^2 is only one of its tests. The model-independent checks in Section 11.5.4 are where most curves are actually decided.

SUMMARY

EPA SW-846 Method 8000D defines the initial calibration requirements shared by the SW-846 chromatographic methods. This paper sets out those requirements by calibration model, explains the model-independent percent error and RSE checks in Section 11.5.4, and applies every check to one worked data set. It shows how a curve can pass the correlation requirement and still fail the method, and closes with an eight-step ICAL review checklist.

What Method 8000D governs.

SW-846 Method 8000D, *Determinative Chromatographic Separations*, is the method EPA uses to set the calibration and quality control requirements shared by the SW-846 chromatographic methods, such as 8081, 8082, 8260, and 8270 ^[1]. Most environmental laboratories never run 8000D by itself. They run a determinative method that refers back to it.

That relationship matters when a curve is reviewed. Where the determinative method specifies its own quality control criteria, those criteria take precedence over the general guidance in 8000D (Sec. 9.1). Where it does not, 8000D supplies the default. A reviewer who knows 8000D knows the baseline that every SW-846 calibration is measured against.

The foundation of every ICAL.

Before any model is chosen, Section 11.4 sets the structure of the initial calibration (ICAL).

01 **At least five concentrations.**

Standards for each target analyte and surrogate are prepared at a minimum of five levels within the working range of the instrument (Sec. 11.4.1.1).

02 **The lowest standard sets the LLOQ.**

The lowest calibration standard generally establishes the lower limit of quantitation (Sec. 11.4.1.2), and at least one standard must sit at or below the quantitation level the project needs (Sec. 11.4.1.4).

03 **No extrapolation.**

Reporting results above or below the calibrated range is not appropriate and may lead to significant quantitative error, regardless of the model (Sec. 11.4).

04 **No zero point.**

A concentration of zero is not included as a calibration point (Sec. 11.5).

Criteria by calibration model.

Method 8000D allows three families of calibration model. Each has its own minimum number of levels and its own primary acceptance criterion.

MODEL	MINIMUM LEVELS	PRIMARY CRITERION	SECTION
Average calibration or response factor	5	RSD of the factors $\leq 20\%$	11.5.1
Linear least squares regression	5	$r \geq 0.995$, equivalent to $r^2 \geq 0.99$. Weighting by $1/x$ or $1/x^2$ is permitted.	11.5.2
Quadratic (second order)	6	$r^2 \geq 0.99$. No more than four parameters in any non-linear model.	11.5.3
Cubic (third order)	7	$r^2 \geq 0.99$. Third order is the highest polynomial permitted.	11.5.3

Two details in Section 11.5.2 are often misread. First, regression may be forced through the origin, but forcing it may not be used as a rationale for reporting results below the calibrated range. Second, the method names weighting, by $1/x$ or $1/x^2$, as a way to improve the fit at low concentrations. A lab that always runs unweighted regression is not using a tool the method expects it to consider.

The checks that apply to every model.

Section 11.5.4 adds checks that apply regardless of which model was chosen. They measure the error the curve actually introduces, in the units the lab reports.

CHECK	REQUIREMENT	SECTION
Percent error, every standard	Calculated versus expected amount $\leq 30\%$ at every calibration level	11.5.4.1
Percent error, lowest standard	$\leq 50\%$ may be acceptable for the lowest point, for some data uses	11.5.4.1
Relative standard error	Same limit as the method's RSD limit. Where none is defined: $\leq 20\%$ for good performing compounds, $\leq 30\%$ for poor performing compounds.	11.5.4.2

These checks are the reason a correlation coefficient alone does not settle an 8000D calibration. As [Insights 01](#) showed, r^2 is dominated by the high end of the range, while percent error and RSE weight every level by its own concentration ^[2]. A curve can satisfy Section 11.5.2.2 and still fail Section 11.5.4.

PERCENT ERROR, PER SECTION 11.5.4.1

$$\% \text{ Error} = 100 \times (x'_i - x_i) / x_i$$

x'_i is the concentration of calibration standard i calculated from the curve, and x_i its expected concentration. RSE pools these errors across all levels, adjusted for the number of terms in the model.

One data set, every check.

The data set from [Insights 01](#) is an eight-level calibration from 0.5 to 100 ppb. Here it is fit four ways and run against every 8000D check: the primary criterion for each model, percent error at every level, and RSE against the default 20% limit.

MODEL	PRIMARY CRITERION	ERROR AT 0.5 PPB, ≤ 30% AT ALL LEVELS	RSE ≤ 20%	RESULT
Average response factor	RSD 3.5% PASS	+6.2% PASS	3.5% PASS	PASS
Linear, unweighted	r^2 0.99960 PASS	-36.7% FAIL	17.6% PASS	CONDITIONAL
Linear, weighted 1/x	r^2 0.99954 PASS	-0.5% PASS	2.8% PASS	PASS
Linear, weighted 1/x²	r^2 0.99945 PASS	+0.8% PASS	2.8% PASS	PASS

Illustrative data set, identical to Insights 01. Every value is computed from the same eight responses. Conditional means the curve passes only if the 50% allowance for the lowest standard is acceptable for the data use.

0.99960

R^2 OF THE UNWEIGHTED FIT, WHICH PASSES SECTION 11.5.2.2

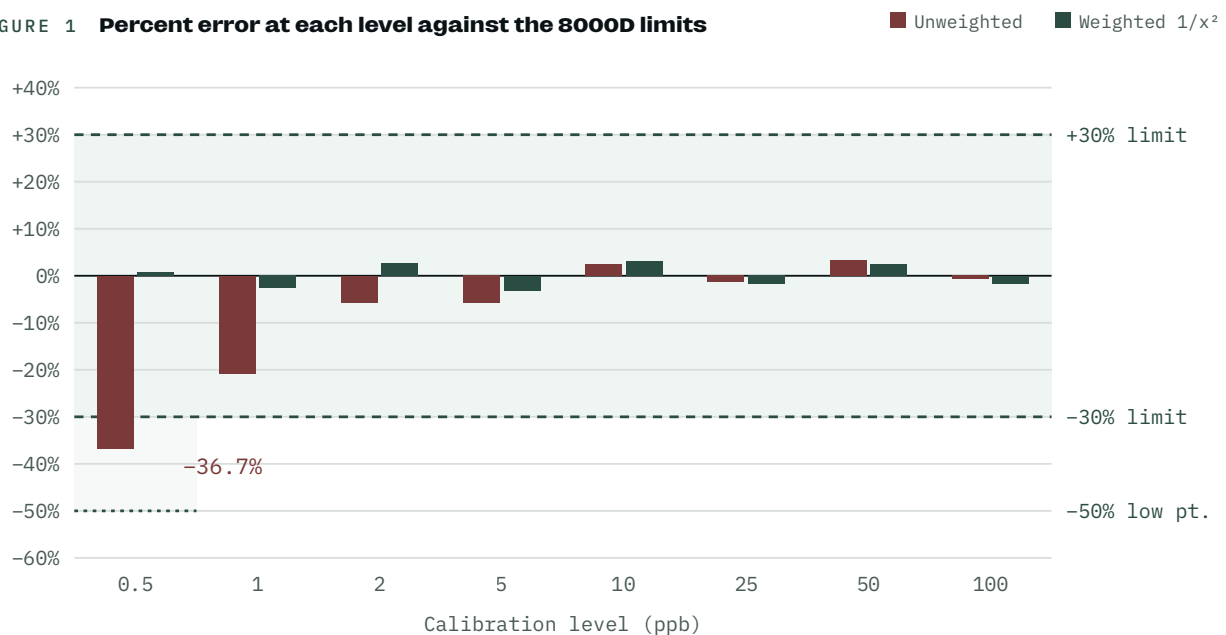
-36.7%

PERCENT ERROR AT THE LOWEST STANDARD, OUTSIDE THE 30% LIMIT

17.6%

RSE OF THE SAME FIT, INSIDE THE 20% LIMIT

FIGURE 1 Percent error at each level against the 8000D limits



The unweighted linear fit has the highest r^2 of the three regressions and clears the 0.99 requirement easily. Its RSE of 17.6% also passes. It fails only one check: the lowest standard back-calculates to 0.32 ppb, a percent error of -36.7%, outside the 30% limit.

Whether that curve is acceptable now depends on the 50% allowance for the lowest point, which 8000D grants only for some data uses. That is a judgment the lab has to make and document, and it is the kind of decision an auditor asks about. The weighted fits and the average response factor model pass every check with no exception needed, holding every standard within 3.5%.

A review that stops at r^2 would accept all three regressions as equivalent. A review that runs every 8000D check finds that one of them needs a documented exception and three do not.

When an analyte fails.

A failed criterion does not void the whole calibration. Under Section 11.5.6, the ICAL remains valid for every analyte that met its criteria. Data for a failing analyte may still be used for screening or estimated quantitation, where that is consistent with the project's objectives.

The method asks the analyst to inspect the model by refitting the calibration standards to it and comparing calculated with expected amounts. If any standard shows a percent error greater than 30%, corrective action may be needed before the analyte is reported quantitatively (Sec. 11.5.6.1). Whatever action is taken, the record should show what was evaluated and why the final model was accepted.

A Method 8000D ICAL review checklist.

01 **Confirm which criteria govern.**

Check the determinative method first. Its own criteria take precedence over 8000D where they exist.

02 **Count the levels.**

Five for average response factor or linear; six for quadratic, seven for cubic. No zero point.

03 **Confirm the low standard.**

It sets the LLOQ and must meet the project's quantitation needs.

04 **Apply the primary criterion for the model.**

$RSD \leq 20\%$ for average response factor; $r^2 \geq 0.99$ for regression.

05 **Back-calculate every standard.**

Percent error $\leq 30\%$ at every level, and record whether the 50% low-point allowance was used and why.

06 Calculate RSE.

Against the method limit, or $\leq 20\%$ for good performing compounds and $\leq 30\%$ for poor performers.

07 Evaluate weighting.

Record whether $1/x$ or $1/x^2$ was considered, as Section 11.5.2 suggests.

08 Document analyte-level outcomes.

Which analytes passed, which failed, and how failing data were qualified.

Where WLTR fits.

Steps four through seven multiply quickly. A semivolatile panel with dozens of analytes, each tested against several permitted models and every 8000D check, is more review than a spreadsheet supports on a routine schedule. Most labs evaluate the one model their data system applies by default.

WLTR is calibration data integrity infrastructure for regulated analytical laboratories. For each analyte, it derives nine calibration regression models (average response factor, plus linear and quadratic fits across four weightings each) and evaluates them against fifteen ICAL parameters, including RSE, computed orthogonally to the CDS. The methodology is covered by U.S. Patent 12,038,448 B2 ^[3].

CALIBRATION VERIFICATION

See how your ICAL review holds up against every Method 8000D check, not only the one your data system reports first.

[GET COMPLIANCE GAP ASSESSMENT](#)
[WLTR.TECH](#)

SOURCES

1. U.S. EPA, SW-846 Test Method 8000D, *Determinative Chromatographic Separations*, Revision 5, March 2018. Section numbers in this article refer to this document. www.epa.gov/hw-sw846/sw-846-test-method-8000d-determinative-chromatographic-separations
2. R. Burrows and J. Parr; "Evaluating the Goodness of Instrument Calibration for Chromatography Procedures," *LCGC Supplements* 38(11), November 2020. www.chromatographyonline.com/view/evaluating-the-goodness-of-instrument-calibration-for-chromatography-procedures
3. U.S. Patent 12,038,448 B2, inventor Walter "Chip" Hinchman Sr. patents.google.com/patent/US12038448B2

Walter Hinchman, MBA

WLTR, INC. • NEVADA C-CORP

