

## CEC SDG Glossary of Terms

### Notes

- Terms are listed alphabetically
- Acronyms are either identified by underlined words (e.g. Co-ordinating European Council), or refer to a separate full-text entry
- Each **bold and underlined** entry in the Description column refers to another entry in this document. These are not hyperlinked

Term	Description
<b>Accuracy</b>	A measure of how far an individual measurement, or average of several measurements, is from the <b><u>true value</u></b> .
<b>Actual value</b>	The actual quantitative value for the prepared sample (the actual value only exists for fundamental physical or chemical properties such as density, concentration, temperature, etc.)
<b>ATC-ERC</b>	<u>Additive Technical Committee - European Registration Centre</u> . The <b><u>data depository</u></b> for all <u>CEC</u> engine test results.
<b>Back-to-back</b>	Consecutive determinations.
<b>Bias</b>	The bias in a test method is the difference between the <b><u>true value</u></b> and the <b><u>actual value</u></b> .
<b>Bias Limits</b>	The bias limits determine how far the <b><u>EWMA</u></b> trend line can deviate from the <b><u>target</u></b> before an action is required. Typically this is set to one <b><u>reproducibility standard deviation (RSD)</u></b> from the <b><u>target</u></b> .
<b>Candidate Test</b>	A test performed on a fluid of unknown performance, usually as part of a fluid development programme.
<b>CEC</b>	<u>Co-ordinating European Council</u> (for the development of performance tests for fuels, lubricants and other fluids).
<b>CEC-TMS</b>	The <u>CEC</u> on-line <b><u>test monitoring system (TMS)</u></b> used for the majority of non-engine tests.
<b>Cochran's Test</b>	An outlier test which looks for unusually large differences across sets of two or more test results on the same sample at the same laboratory. If the standard deviation across such a set is large relative to comparable sets, then one or more of the results may be identified as an <b><u>outlier</u></b> or <b><u>straggler</u></b> .
<b>Confidence Interval</b>	The range of plausible values for calculated statistics such as the <b><u>mean</u></b> , <b><u>repeatability</u></b> or <b><u>reproducibility</u></b> . Typically this would express 95% or 99% confidence that the true value lies in the stated range.
<b>Control Chart</b>	Within <u>CEC</u> this is a chart used to plot <b><u>reference test</u></b> data against time to determine if a laboratory (or stand) is performing as expected. Such a chart may additionally display the <b><u>target</u></b> performance level, <b><u>control limits</u></b> , <b><u>warning limits</u></b> , <b><u>bias limits</u></b> or the <b><u>EWMA</u></b> trend line. Control charts may also be used to monitor industry trends across laboratories.
<b>Control Limits</b>	Control limits are lines on a statistical <b><u>control chart</u></b> some distance from the <b><u>target</u></b> which are used to detect results of testing which deviate substantially from the <b><u>target</u></b> . Results outside the control limits are deemed to be non-compliant and a laboratory (or stand) must be re-referenced before <b><u>candidate testing</u></b> can resume. Within <u>CEC</u> these are set at $\pm k \times \text{RSD}$ where k is a multiplier for the <b><u>RSD</u></b> (usually 1.8 but may be between 1.6 and 2.0).

<b>d.f.</b>	See <b><u>degrees of freedom</u></b>
<b>Data Depository</b>	A centralized web-based electronic database for the storage, visualization and retrieval of <b><u>test monitoring</u></b> data. This is the <b><u>ATC-ERC</u></b> database for engine tests and the <b><u>CEC-TMS</u></b> database for all other tests that have <b><u>test monitoring systems</u></b> . All <b><u>reference test</u></b> data should be uploaded. The ATC-ERC database also stores <b><u>candidate test</u></b> results.
<b>Data Dictionary</b>	Defines which data fields should be stored in the <b><u>data depository</u></b> and how that data should be stored e.g. how many decimal places, how many characters, which units, etc.
<b>Degrees of Freedom</b>	When calculating a statistic (e.g. the <b><u>mean</u></b> or <b><u>variance</u></b> ), the degrees of freedom is the number of results that we need to know before we can work out the remaining results from the data we already have. A higher number of degrees of freedom produces greater confidence in the estimated value and conversely lower degrees of freedom produces less confidence in the estimated value. The <b><u>d.f.</u></b> is calculated from the number of repeat tests on a fluid and the number of stands generating the data.
<b>Discrimination</b>	The ability of the test to show differences between fluids of known different performance.
<b>Discrimination Chart</b>	Plots the results for two or more fluids on the same graph to show the discrimination between the fluids. Also used to compare different <b><u>reference fluid</u></b> batches.
<b>Discriminating Power</b>	Is the probability of measuring a significant difference between two fluids (or between a product and a specification) when a genuine difference does exist.
<b>EWMA</b>	<u>Exponentially weighted moving average</u> . Also known as the trend line. It is used to help the user to detect shifts in the <b><u>mean</u></b> value of test results and is usually displayed on <b><u>control charts</u></b> .
<b>Grand Mean</b>	Calculated by taking the <b><u>mean</u></b> of the <b><u>laboratory mean</u></b> values for a particular fluid. The grand mean (or <b><u>industry mean</u></b> ) across many stands and laboratories is thus calculated in accordance with ISO 5725. This is done in order to prevent any one laboratory (or stand) having an undue influence on the statistics by virtue of having a greater number of test data points than other laboratories (or stands).
<b>Grubbs' Test</b>	An outlier test which is used to detect whether the <b><u>laboratory mean</u></b> value for a particular fluid differs markedly from other laboratories. If the difference between the <b><u>laboratory mean</u></b> and the <b><u>grand mean</u></b> is statistically <b><u>significant</u></b> , then one or more of its results may be identified as an <b><u>outlier</u></b> or <b><u>straggler</u></b> .
<b>Industry Mean</b>	See <b><u>grand mean</u></b>
<b>KSD</b>	This is an alternative method to <b><u>Grubbs' test</u></b> for detecting <b><u>outliers</u></b> and <b><u>stragglers</u></b> . It is used mainly for <b><u>test monitoring</u></b> data. K refers to the multiplier of the <b><u>Reproducibility Standard Deviation</u></b> used to determine outliers. K is typically set to 3.
<b>Laboratory Bias</b>	The difference between the expectation of the test results from a particular laboratory and the <b><u>true value</u></b> .
<b>Laboratory Mean</b>	The <b><u>mean</u></b> value of all the results for a particular fluid recorded at a single laboratory (or stand).
<b>LBL</b>	Lower <b><u>bias limit</u></b> - a line on a <b><u>control chart</u></b> .

<b>LCL</b>	Lower <b>control limit</b> - a line on a <b>control chart</b> .
<b>Liaison Officer</b>	Point of contact between groups. Within CEC this term is usually only applied to members of the <b>statistical development group</b> .
<b>LO</b>	See <b>liaison officer</b>
<b>LWL</b>	Lower <b>warning limit</b> - a line on a <b>control chart</b> .
<b>Mean</b>	A mean value is obtained by summing all values and dividing by the number of values.
<b>Median</b>	The middle value in a series of values when ordered by size.
<b>Mode</b>	The value observed the most times in a data set.
<b>Outlier</b>	An outlier is an unusual data point which, when examined using <b>Grubbs' test</b> , <b>Cochran's test</b> or <b>KSD</b> , is significant at <b>p</b> < 1%. Not all identified <b>outliers</b> are invalid (e.g. arising from deviations from the test method or changes in the test sample); it is left to the discretion of the <b>SDG LO</b> and the <b>WG</b> to determine, using statistical and engineering judgement, whether such data should be removed from further calculations of the <b>mean</b> , <b>repeatability</b> , <b>reproducibility</b> , etc.
<b>p (p-value)</b>	A measure of the statistical significance of a calculated result, such as the difference in performance between two fluids. Results are typically declared as <b>significant</b> if <b>p</b> < 5% and highly <b>significant</b> if <b>p</b> < 1%. In <b>CEC</b> , <b>p</b> -values are also used to determine whether test results are <b>outliers</b> or <b>stragglers</b> .
<b>Pilot Study</b>	Initial study at a small number of laboratories, prior to a full <b>round robin</b> , to determine whether operators can follow the test procedure, to check sample distribution and handling procedures, and to obtain rough estimates of sample <b>means</b> , <b>repeatability</b> and <b>reproducibility</b> .
<b>Precision</b>	A measure of how closely an individual measurement is to other measurements on the same fluid. <b>Repeatability</b> and <b>reproducibility</b> are measures of precision.
<b>Precision Statement</b>	A statement regarding the <b>precision</b> of the test method stating how the data was obtained, descriptions of the fluids, numbers of data points, numbers of laboratories (or stands), and the <b>mean</b> , <b>repeatability</b> and <b>reproducibility</b> of the test for each reference fluid. It is usually produced from the most recent <b>round robin</b> or from <b>TMS</b> data.
<b>Q<sub>r</sub></b>	The achieved <b>repeatability</b> divided by the <b>repeatability target</b> . For tests that are in control <b>Q<sub>r</sub></b> will be ≤ 1.
<b>Q<sub>R</sub></b>	The achieved <b>reproducibility</b> divided by the <b>reproducibility target</b> . For tests that are in control <b>Q<sub>R</sub></b> will be ≤ 1
<b>r</b>	See <b>Repeatability</b>
<b>R</b>	See <b>Reproducibility</b>
<b>r-Target</b>	See <b>repeatability target</b>
<b>R-Target</b>	See <b>reproducibility target</b>
<b>Reference Fluid</b>	An oil or fuel of known performance that is used periodically in <b>test monitoring</b> to check that a laboratory (or stand) is capable of obtaining acceptable results.

<b>Reference Test</b>	A test performed on a <b>reference fluid</b> for checking the laboratory (or stand) is capable of running the test correctly.
<b>Repeatability</b>	An estimate of precision which indicates the difference (with a 95% chance) that two measurements obtained under <b>repeatability conditions</b> will lie. It is calculated as 2.8 x the <b>repeatability standard deviation (rSD)</b> .
<b>Repeatability Conditions</b>	Test results are obtained with the same method on aliquots of the same sample in the same laboratory by the same operator using the same equipment in a short period of time.
<b>Repeatability Standard Deviation</b>	A measure of the spread of data around the <b>mean</b> value when analysing data gathered under <b>repeatability conditions</b> . It is a measure of within laboratory (or stand) variability. <b>Repeatability</b> data from multiple locations can be pooled to give an overall measure across the industry.
<b>Repeatability Target</b>	The maximum <b>repeatability</b> that is generally acceptable taking into account the requirements of the test.
<b>Reproducibility</b>	An estimate of precision which indicates the difference (with a 95% chance) that two measurements obtained under <b>reproducibility conditions</b> will lie. It is calculated as 2.8 x the <b>reproducibility standard deviation (RSD)</b> .
<b>Reproducibility Conditions</b>	Test results are obtained with the same method on aliquots of the same sample in two or more laboratories (or stands).
<b>Reproducibility Standard Deviation</b>	A measure of the spread of data from a number of locations around the <b>mean</b> value. It incorporates both within-laboratory (or stand) variation and between-laboratory (or stand) variation.
<b>Reproducibility Target</b>	The maximum <b>reproducibility</b> that is generally acceptable taking into account the requirements of the test.
<b>Round Robin</b>	A programme of tests conducted at a number of laboratories using the same samples. Typically multiple samples are tested at multiple laboratories with each sample tested at least twice at each laboratory. The primary purposes are to establish (or check) the <b>repeatability</b> and <b>reproducibility</b> of a test method and to determine <b>target</b> values for <b>reference fluids</b> .
<b>rSD</b>	See <b>repeatability standard deviation</b>
<b>RSD</b>	See <b>reproducibility standard deviation</b>
<b>Run Rules</b>	Additional checks / constraints that can be applied to <b>reference test</b> results to ensure there is no systematic drift or variation occurring at a laboratory (or stand).
<b>SD</b>	See <b>Standard Deviation</b>
<b>SDG</b>	See <b>Statistical Development Group</b>
<b>SDG LO</b>	See <b>Statistical Development Group Liaison Officer</b>
<b>SG</b>	See <b>Surveillance group</b>
<b>Severity</b>	Severity is a measure of the deviation of a test result, or series of results, from the <b>grand mean</b> . A laboratory (or stand) is described as more severe, if its results show poorer performance of the fluid than expected

	compared to other data from other laboratories (or stands) on the same fluid. It is described as milder if its results show better performance than expected compared to other data from other laboratories (or stands) on the same fluid. A test method can also be described as severe or mild.
<b>Significant</b>	An observed difference is statistically significant if the probability of seeing such an extreme value by chance, when in fact no underlying difference exists, is less than some predetermined value such as 5% or 1%. Significance is determined by the <b>p-value</b> .
<b>Site Precision</b>	An estimate of precision which indicates the difference (with a 95% chance) that two measurements obtained under <b>site precision conditions</b> would be expected to lie.
<b>Site Precision Conditions</b>	Test results are obtained with the same method on aliquots of the same sample in the same laboratory by one or more operators over an extended period of time.
<b>Standard Deviation</b>	A measure of the spread of data around the <b>mean</b> value. It is the square root of the <b>variance</b> .
<b>START</b>	The statistical program, written by <b>SDG</b> , for use by its members to consistently analyse <b>round robins</b> and <b>test monitoring</b> data to calculate estimates of the <b>mean</b> values and <b>precision</b> for each tested sample.
<b>Statistical Development Group</b>	A <b>CEC</b> working group whose members are responsible for the analysis of <b>CEC</b> data as well as the development of the statistical protocols and tools. Each <b>surveillance group</b> and <b>test development group</b> should include a member of the <u>Statistical Development Group</u> to provide advice and support.
<b>Statistical Development Group Liaison Officer</b>	A member of <b>SDG</b> assigned to any <b>surveillance group</b> or <b>test development group</b> to conduct all statistical evaluations and provide other statistical support and advice.
<b>Straggler</b>	A straggler is an unusual data point which, when examined using <b>Grubbs' test</b> , <b>Cochran's test</b> or <b>KSD</b> , is <b>significant</b> at $p < 5\%$ but not at $p < 1\%$ . <b>Stragglers</b> are usually retained for further calculations such as the <b>mean</b> , <b>repeatability</b> , <b>reproducibility</b> , etc.
<b>Surveillance Group</b>	A group consisting of representatives from institutions who either run or sponsor an approved <b>CEC</b> test. Their role is to monitor the on-going <b>precision</b> and <b>severity</b> of the test and to deal with any issues that may arise that affect the test.
<b>Target (for Control Charts)</b>	The target is a line on a <b>control chart</b> which is normally set to the <b>grand mean</b> (estimated from a <b>round robin</b> or <b>TMS</b> ). It is surrounded by <b>bias limits</b> , <b>warning limits</b> and <b>control limits</b> .
<b>TDG</b>	See <b>test development group</b>
<b>Test Development Group</b>	A group consisting of representatives from institutions who intend either to run or to sponsor a new <b>CEC</b> test. Their role is to develop the test to a point where it is acceptable to the <b>CEC</b> management board. <b>TDGs</b> usually become <b>SGs</b> once the test development is completed.
<b>Test Monitoring</b>	A process used to evaluate the ability of a laboratory (or stand) to obtain acceptable results by the stated test method. This is achieved by conducting tests at pre-defined intervals using <b>reference fluids</b> of known performance. The results are then uploaded to a <b>data depository</b> where <b>control charts</b> can be constructed.

<b>Test Monitoring System</b>	The <b>test monitoring</b> process and its associated <b>data depository</b> . The <b>TMS</b> for a particular test is defined in section 11 of the test procedure.
<b>Test Monitoring Targets</b>	See <b>r-target</b> and <b>R-target</b>
<b>Test Order</b>	For <b>round robins</b> it is not ideal to test the same sample <b>back-to-back</b> or test the samples in the same order at all laboratories. The <b>SDG LO</b> will determine a specific order in which fluids must be tested which will be, as far as possible, different at each laboratory.
<b>TMS</b>	See <b>Test Monitoring System</b>
<b>Transformation</b>	A function (such as logarithm) which is used to better analyse the actual measurements, e.g. to expedite precision analysis. This is usually done when the <b>standard deviation</b> is not constant for all samples but varies with the magnitude of the test result.
<b>True Value</b>	The unknown value that we try to estimate by testing. The value to which the <b>mean</b> of single results obtained by different laboratories tends as the number of tests increases towards infinity.
<b>UBL</b>	Upper <b>Bias limit</b> – a line on a <b>control chart</b> .
<b>UCL</b>	Upper <b>Control limit</b> – a line on a <b>control chart</b> .
<b>UWL</b>	Upper <b>Warning limit</b> – a line on a <b>control chart</b> .
<b>Variance</b>	A measure of deviation from a <b>mean</b> value. It is the sum of the squares of the deviations from the <b>mean</b> for all data points divided by the number of data points minus 1.
<b>Warning Limits</b>	Warning limits are lines on a statistical <b>control chart</b> some distance from the <b>target</b> which are used to detect results of testing which deviate somewhat from the <b>target</b> . Results outside the warning limits will usually trigger investigations to check that the test was run correctly but do not, on their own, constitute an invalid test. They may be referred to in any additional <b>run rules</b> that may be in place. Within <b>CEC</b> these are usually set to $0.75 \times k$ , where k is the factor used to calculate the <b>control limits</b> .
<b>WG</b>	See <b>working groups</b>
<b>Working Groups</b>	<b>Surveillance Groups</b> and <b>Test Development Groups</b> .
<b>z-Scale</b>	A standardised scale in which, for each appropriate <b>reference fluid</b> (or sample thereof), the <b>mean</b> is subtracted from each data point and the resulting value divided by the corresponding <b>standard deviation</b> . This allows different samples with different <b>mean</b> and/or <b>standard deviation</b> values to be compared on the same chart.