KEIT INDUSTRIAL ANALYTICS

IRmadillo™

DOC1118 Repeatability and Reproducibility Study – Guidelines for Ethanol Industry

www.keit.co.uk
support@keit.co.uk

TABLE OF CONTENTS

1. WHAT IS A REPEATABILITY AND REPRODUCIBILITY STUDY?

Keit builds chemometric models based on HPLC data. It is important to understand the uncertainty in the reference data because it may limit the accuracy of the chemometric models.

The accuracy of the IRmadillo measurements will never exceed the accuracy of your HPLC measurements.

The repeatability and reproducibility (R&R) study is meant to capture the totality of all uncertainty introduced throughout the sampling and HPLC measurement process.

A well-maintained HPLC can give reliable results when measuring the same sample – but the process of sample preparation (including drip and filtering) introduces its own uncertainty, so that each sample presented to the HPLC is slightly different. Both of these sources of uncertainty (sample preparation and HPLC measurement) will affect the reliability of your reference measurements, which in turn will affect the reliability of chemometric models.

1.1. Repeatability

What spread of reported values is possible for a given sample?

Repeatability is the difference between repeated test results. This can be assessed by having a single operator prepare and measure the same sample multiple times. Since HPLC reliability often depends on concentration, the test should ideally be conducted across the concentration range to assess performance at different levels.

Repeatability can be evaluated by calculating the standard deviation and the coefficient of variation of the reported values for a given species.

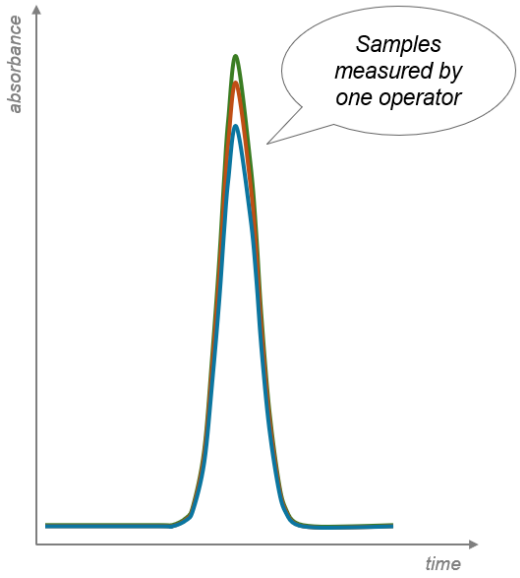

Figure 1 – Schematic of multiple measurements of the same sample by one operator. Note the differences in peak height.

1.2. Reproducibility

How much variation is there between operators?

Samples may be collected and analyzed by different operators/technicians over the course of the IRmadillo calibration process. Everyone has a slightly different way of working (which is entirely natural, and in no way a criticism of individual operators/technicians!). It is important to understand the impact of inter-operator variability to reduce sources of uncertainty. A reproducibility study should ideally include multiple operators, each of whom measures the same sample multiple times.

How much variation is there between different HPLC instruments?

The uncertainty of the HPLC data is embedded into the chemometric models. If different HPLC instruments are used to monitor a process, then the impact of inter-HPLC variability can also be measured.

Reproducibility can be evaluated using ANOVA (analysis of variance). This allows us to compare whether several groups statistically differ across an independent variable, i.e. it helps to check whether variations observed in the data are due to actual differences in the analysis technique, or are due to random chance.

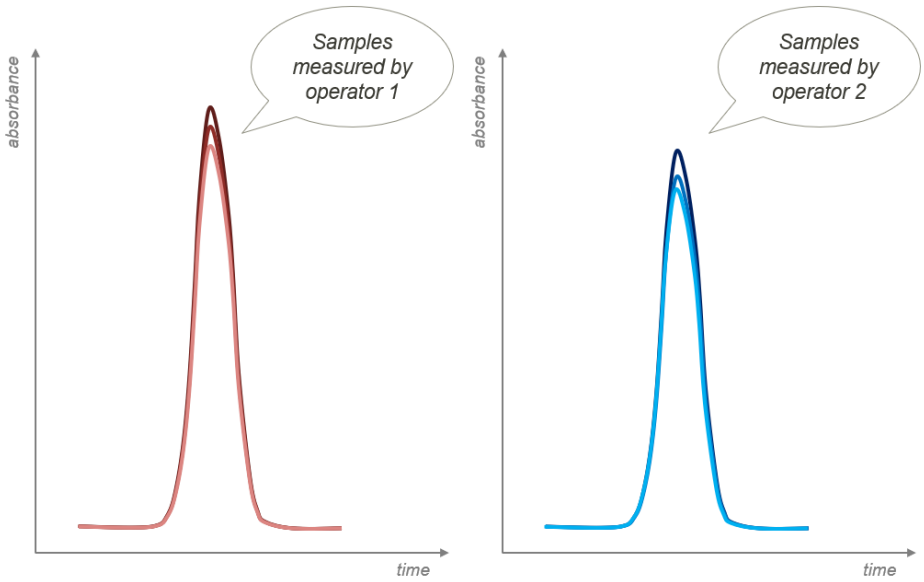

Figure 2 – Example of multiple measurements of the same sample by two operators.

1.3. What does Keit need from you?

There are some important factors that should be considered for a successful R&R study, and consequently good chemometric models.

It is essential that Keit sees what **everyday looks like, not the best**. Often ethanol plants have multiple HPLC instruments to monitor their processes which may differ in performance. Analyzing data from the “good” HPLC and the “bad” HPLC allows Keit to see if the choice of HPLC impacts the measurements, and therefore the performance of the chemometric models.

Sample preparation is also a critical step. Extra care may be taken when taking samples for model *building*; but if a more relaxed effort is used when sampling for model *validation*, the models may be performing very well but *appear* to be performing badly when compared with slightly less accurate HPLC data. All steps count!

Figure 3 shows an example of how data can be reported to Keit for the R&R study. You can use a template provided by Keit. The following information is needed:

- The sample ID for easy identification.
- Sampling location (fermenter, distillation, etc.)
- The HPLC used to measure each sample.
- The units of each chemical concentration are included in the column headers only. Concentration data values must be numeric only. The decimal separator is a period or full stop “.” not a comma “,”.

Sample ID	Sampling Location	Sampling Time (hr)	Operator	HPLC	DP4+ (%w/v)	Maltotriose (%w/v)	Maltose (%w/v)
A1	Ferm 9	6	O1	1	11.442	5.371	4.94
A2	Ferm 9	6	O1	1	11.023	5.357	5.048
A3	Ferm 9	6	O1	1	11.164	5.34	5.063
A4	Ferm 9	6	O1	1	10.949	5.285	5.072
A5	Ferm 9	6	O1	1	11.003	5.228	5.055
A6	Ferm 9	6	O1	1	11.093	5.268	5.104
B1	Ferm 9	30	O1	1	0.691	0.597	0.31
B2	Ferm 9	30	O1	1	0.613	0.482	0.125
B3	Ferm 9	30	O1	1	0.664	0.557	0.287
B4	Ferm 9	30	O1	1	0.611	0.492	0.122
B5	Ferm 9	30	O1	1	0.619	0.485	0.122
B6	Ferm 9	30	O1	1	0.608	0.482	0.123
C1	Ferm 1	48	O1	1	0.512	0.16	0.209
C2	Ferm 1	48	O1	1	0.451	0.111	0.144
C3	Ferm 1	48	O1	1	0.477	0.149	0.224
C4	Ferm 1	48	O1	1	0.477	0.147	0.225
C5	Ferm 1	48	O1	1	0.483	0.154	0.227
C6	Ferm 1	48	O1	1	0.475	0.155	0.231

Sample ID	Sampling Location	Sampling Time (hr)	Operator	HPLC	DP4+ (%w/v)	Maltotriose (%w/v)	Maltose (%w/v)
X1	Ferm 2	24	O1	2	2.003	0.904	1.583
X2	Ferm 2	24	O1	2	1.867	0.92	1.375
X3	Ferm 2	24	O1	2	1.988	0.892	1.532
X4	Ferm 2	24	O1	2	1.834	0.922	1.321
X5	Ferm 2	24	O1	2	1.956	0.911	1.527
X6	Ferm 2	24	O1	2	1.954	0.887	1.495
X7	Ferm 2	24	O2	2	2.065	0.862	1.632
X8	Ferm 2	24	O2	2	1.953	0.889	1.488
X9	Ferm 2	24	O2	2	1.957	0.897	1.476
X10	Ferm 2	24	O2	2	1.953	0.902	1.51
X11	Ferm 2	24	O2	2	1.875	0.841	1.351
X12	Ferm 2	24	O2	2	1.864	0.773	1.267
X13	Ferm 2	24	O3	2	2.086	0.911	1.696
X14	Ferm 2	24	O3	2	1.962	0.896	1.486
X15	Ferm 2	24	O3	2	1.929	0.915	1.434
X16	Ferm 2	24	O3	2	1.968	0.91	1.431
X17	Ferm 2	24	O3	2	1.914	0.768	1.353
X18	Ferm 2	24	O3	2	1.913	0.842	1.417

Figure 3 – Example of how data for R&R study can be reported to Keit.

1.4. What will Keit provide you?

Keit will share all the results of the R&R study with you. Studies like this have often prompted ethanol plants to make changes and improve their measurement processes.

The R&R tests results will be plotted and shared in a report. R&R test results will show the HPLC measurements taken for each sample and/or each operator with error bars that correspond to 95% confidence level. Examples of a repeatability and reproducibility test result are shown in Figure 4 and Figure 5, respectively.

In Figure 4, the black data points represent HPLC measurements of the same sample measured by one operator, and the error bars correspond to two standard deviations. In this example, 3 samples collected at different times into the fermentation batch were analysed 6 times in the HPLC. This representation of the data allows to visually check the spread of repeated measurements.

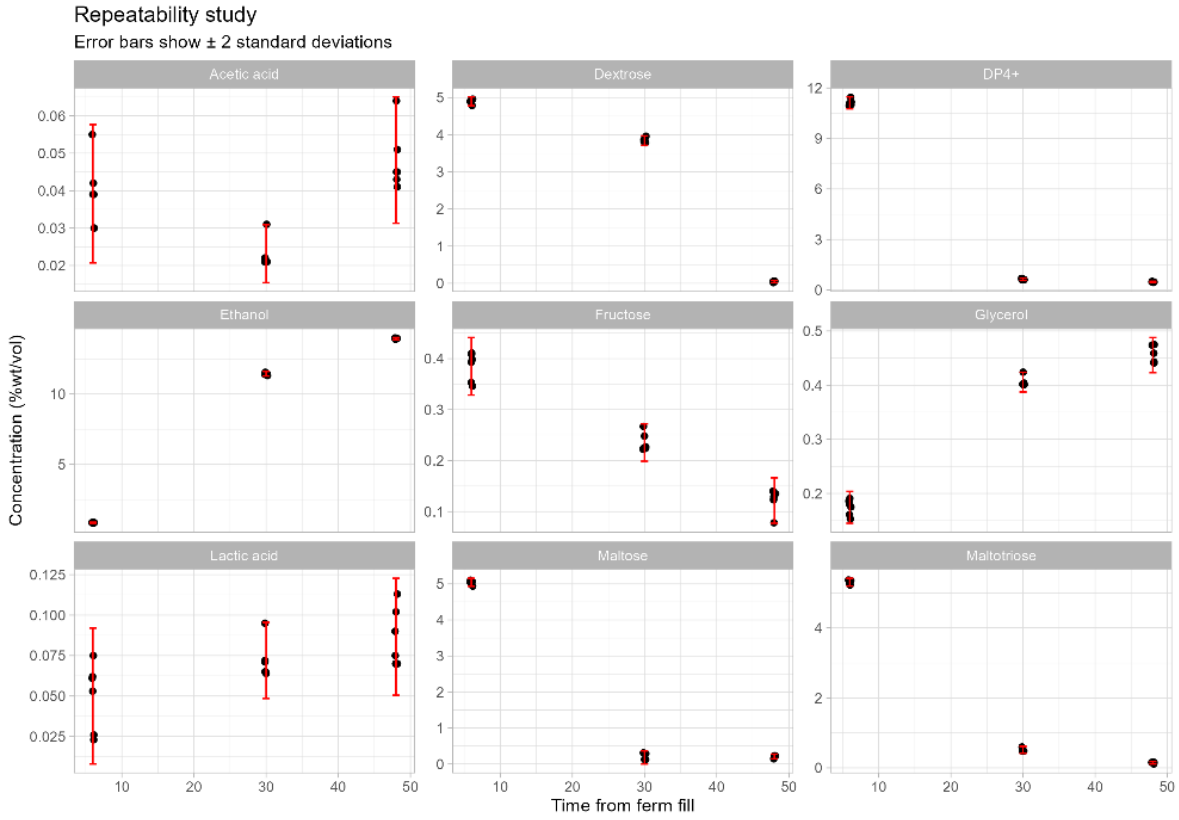

Figure 4 – Example of a repeatability test result.

In Figure 5, the black data points represent HPLC measurements of the same sample, and the error bars correspond to two standard deviations. In this test, 3 distinct operators were selected in which each measured the same sample 6 times in the HPLC. The visual representation of the data allows to check the spread of the data and how different the mean value is for each operator. These plots are complemented with the results of the ANOVA test.

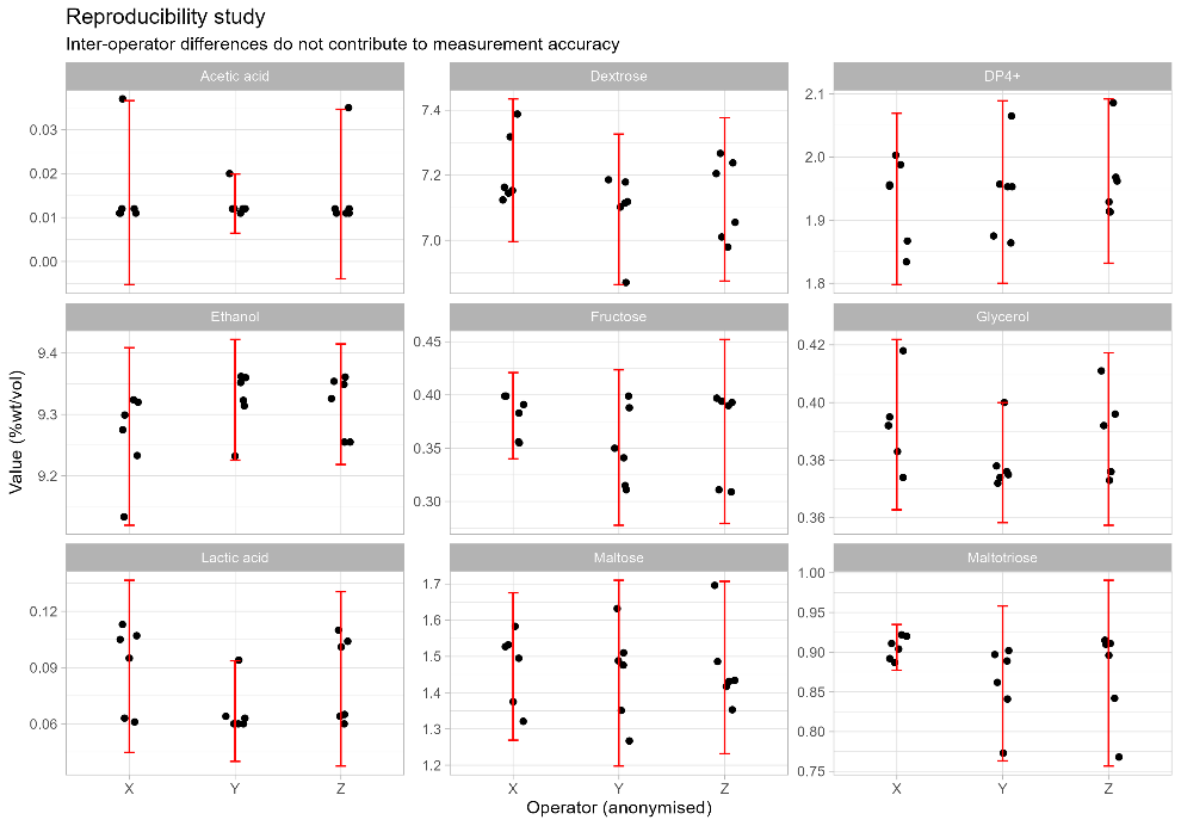

Figure 5 – Example of a reproducibility test result.

2. REPEATABILITY AND REPRODUCIBILITY STUDY - GUIDELINES

2.1.1. Repeatability: Testing how much spread there is in the HPLC data

Keit recommends analyzing 3 samples with different concentrations for each species so that a wide concentration range is analyzed. Each sample will be measured 6 times with the HPLC.

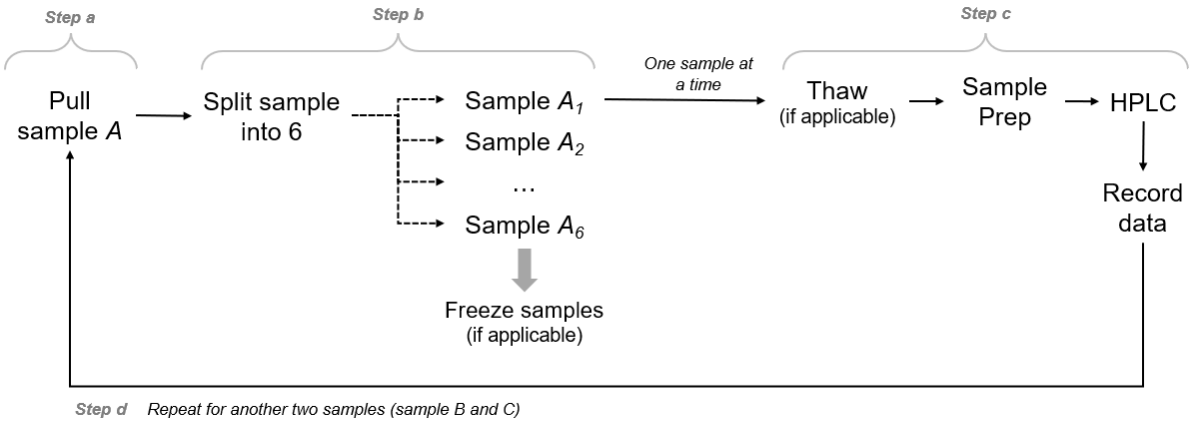

Figure 6 – Diagram of repeatability test step-by-step.

a. Pull a sample (sample *A*) at a given time *t* of your process (e.g. 10 hrs into the fermentation batch).
b. Split the sample into 6 (samples A_1, A_2, …, A_6). If applicable, freeze the samples to stop them from changing over time (i.e. stop the fermentation process).
c. Process one sample A_i at a time:
    i. Thaw sample A_i (if applicable).
    ii. Prepare sample A_i (drip, centrifuge, etc.).
    iii. Measure sample A_i using HPLC. Record the raw data for each species following your normal process.
d. Repeat steps *a* to *c* for another two samples collected at different sampling times, so that the species concentrations are different and ideally cover the wide range of concentrations seen in your process (e.g. 32 hrs and 52 hrs into the fermentation batch).

2.1.2. Reproducibility: Testing inter-HPLC variability

If multiple HPLC instruments are used in your plant, it can be important to understand the impact of the inter-HPLC variability. This is done by repeating the sampling process in section 2.1.1. Inject each sample A_i into the different HPLCs. Record the values output by the HPLCs and report those back to Keit.

2.1.3. Reproducibility: Testing inter-operator variability

Keit recommends using 3 different operators to perform the test. The number of operators may change depending on your resources. Each operator will measure the same sample 6 times using the HPLC.

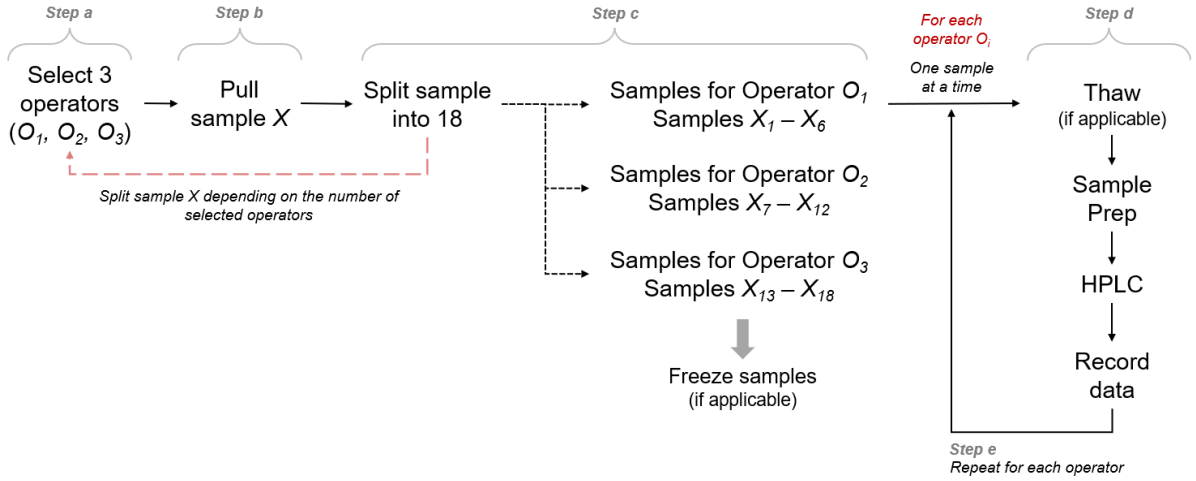

Figure 7 – Diagram of reproducibility test step-by-step.

a. Select 3 different operators to conduct the test (operators O_1, O_2, O_3).
b. Pull a sample (sample X) at a given time t of your process (e.g. 32 hrs into the fermentation batch).
c. Split sample X into multiple subsets so that each operator measures a subset of samples in the HPLC – if 3 operators are select to do the test, split the sample into 18 (samples X_1, X_2, …, X_{18}). If applicable, freeze the samples to stop them from changing over time (i.e. stop the fermentation process). This way, *all operators are measuring identical samples.*
d. For each operator O_i, process one sample X_i at a time:
    i. Thaw sample X_i (if applicable).
    ii. Prepare sample X_i (drip, centrifuge, etc.).
    iii. Measure sample X_i using HPLC. Record the raw data for each species following your normal process
e. Repeat step *d* for each operator.