KEIT
INDUSTRIAL ANALYTICS
Measure | Trust | Improve

Calibrating the IRmadillo for Ethanol Production

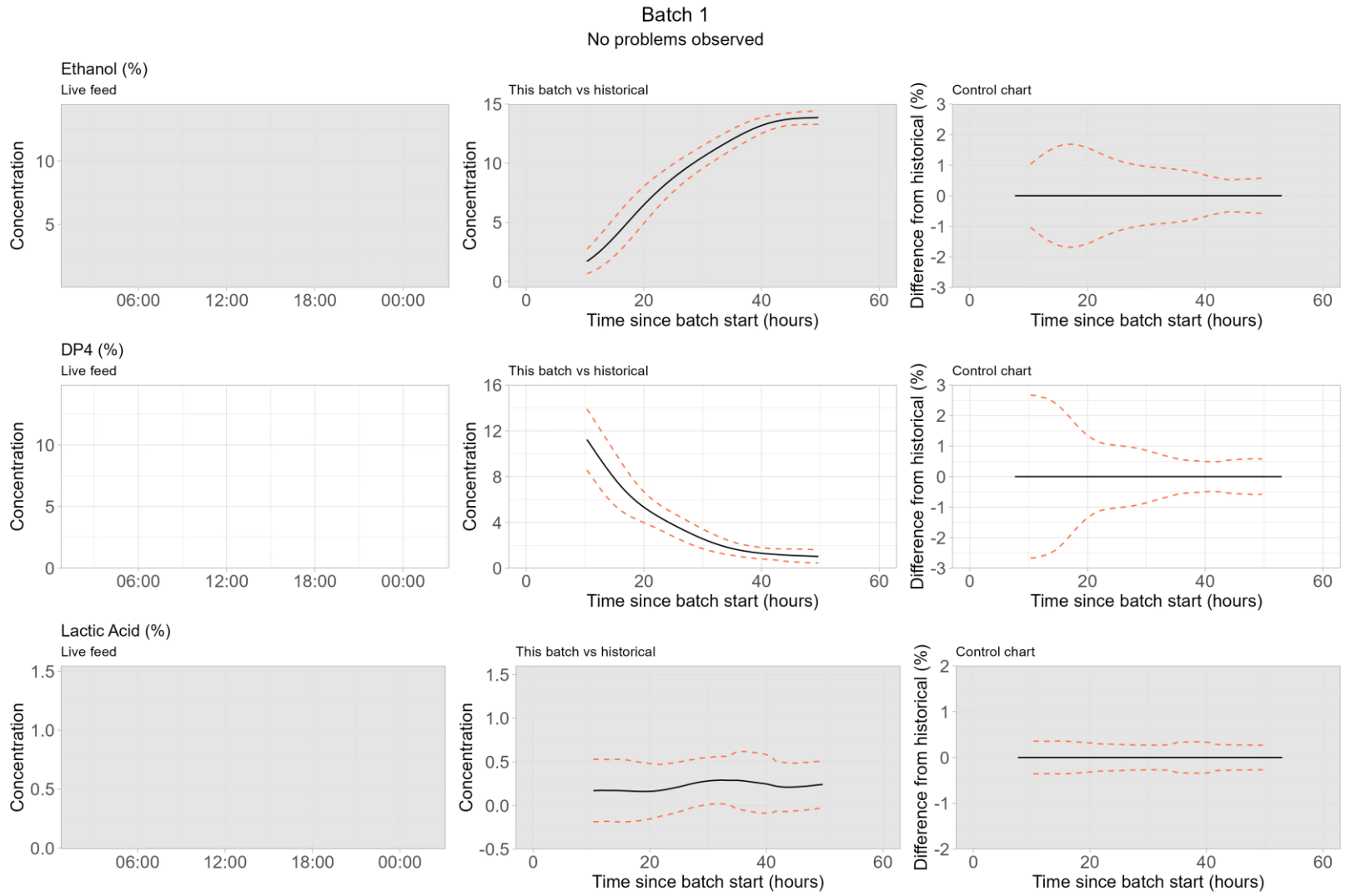
Why are we doing this?

What is calibration?

What is KEIT's process?

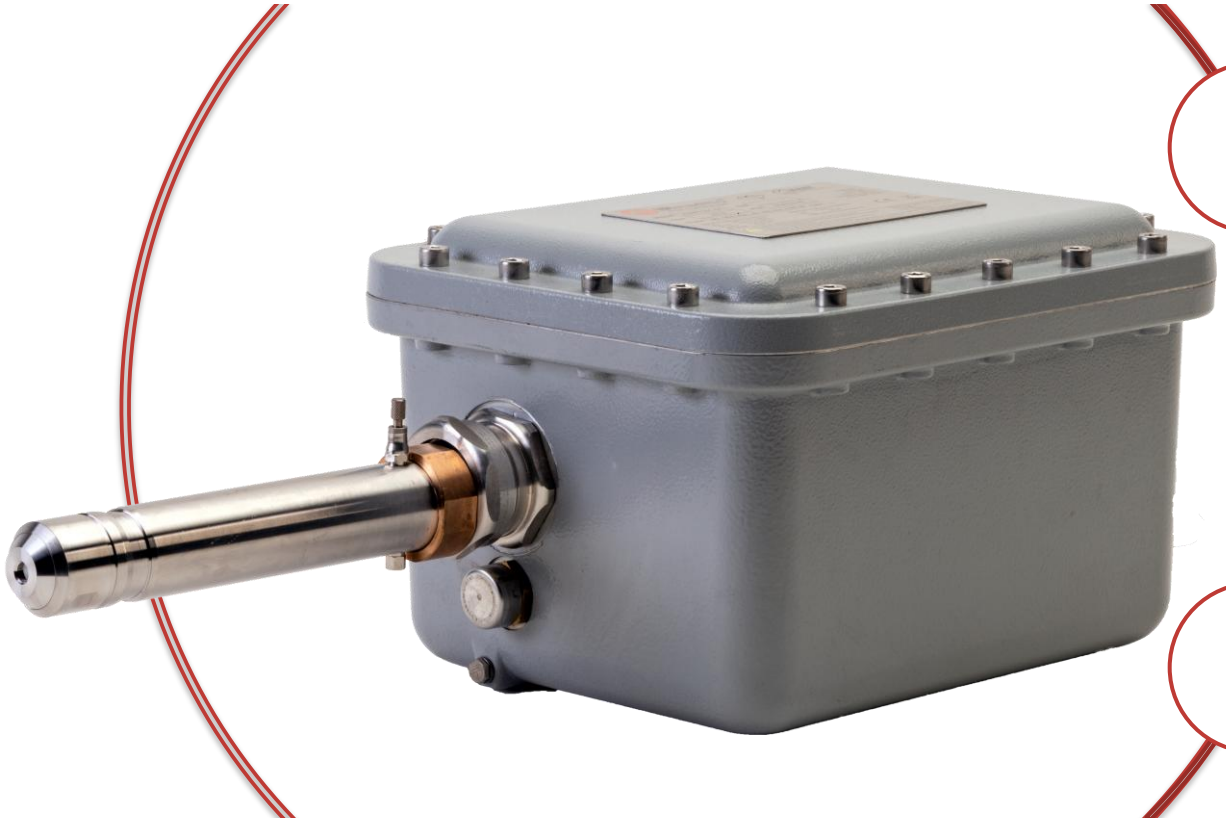
What we need to get the best model

Why are we doing this?



Video available at www.keit.co.uk/irmadillo-higher-ethanol-yields

What is Calibration?



- Chemometric modelling turns the IRmadillo from a Spectrometer into a concentration measuring instrument
- Once calibrated, the IRmadillo can output concentration values for unseen samples
- Multiple calibration models can be run simultaneously on the same IRmadillo

What is Calibration?

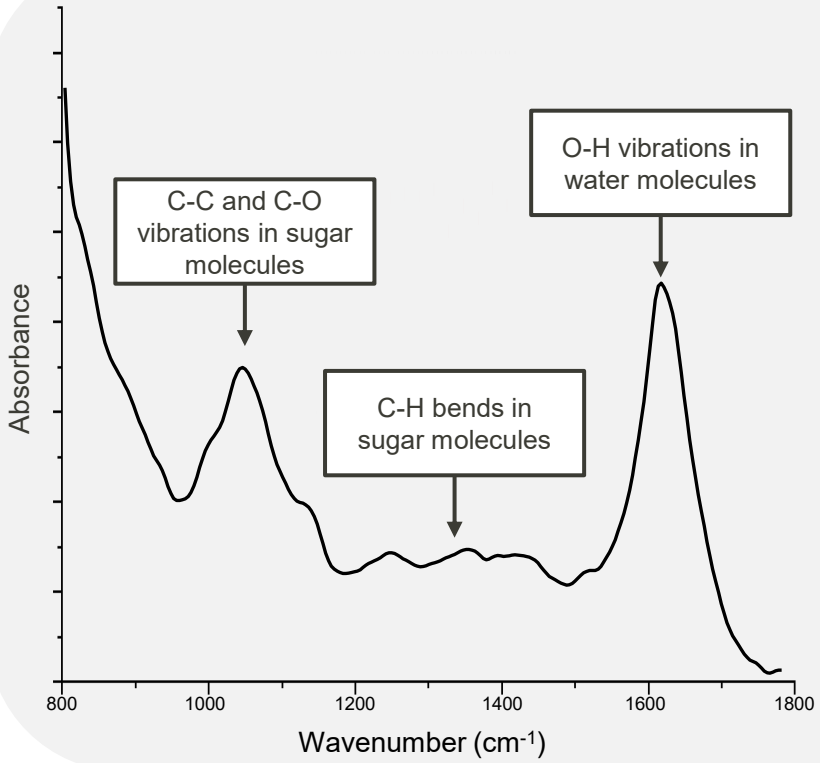
Spectrum of sample (mixed) species

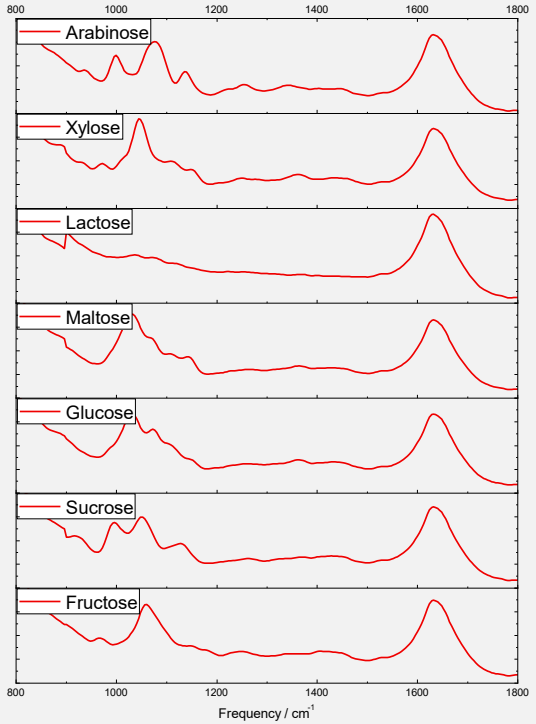
+

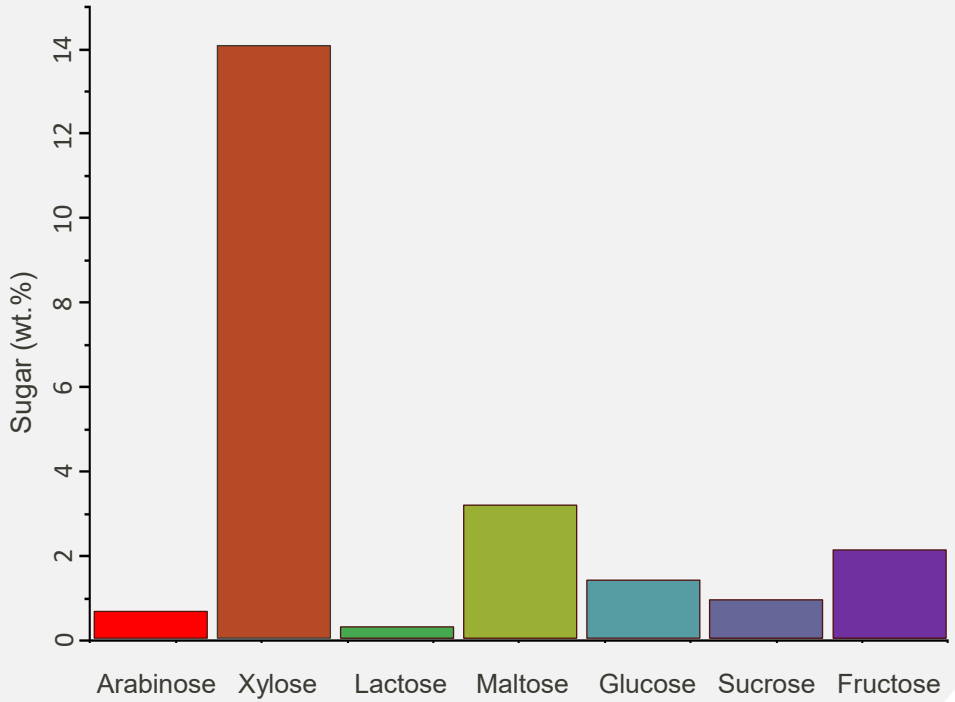
Calibrations for each individual species

=

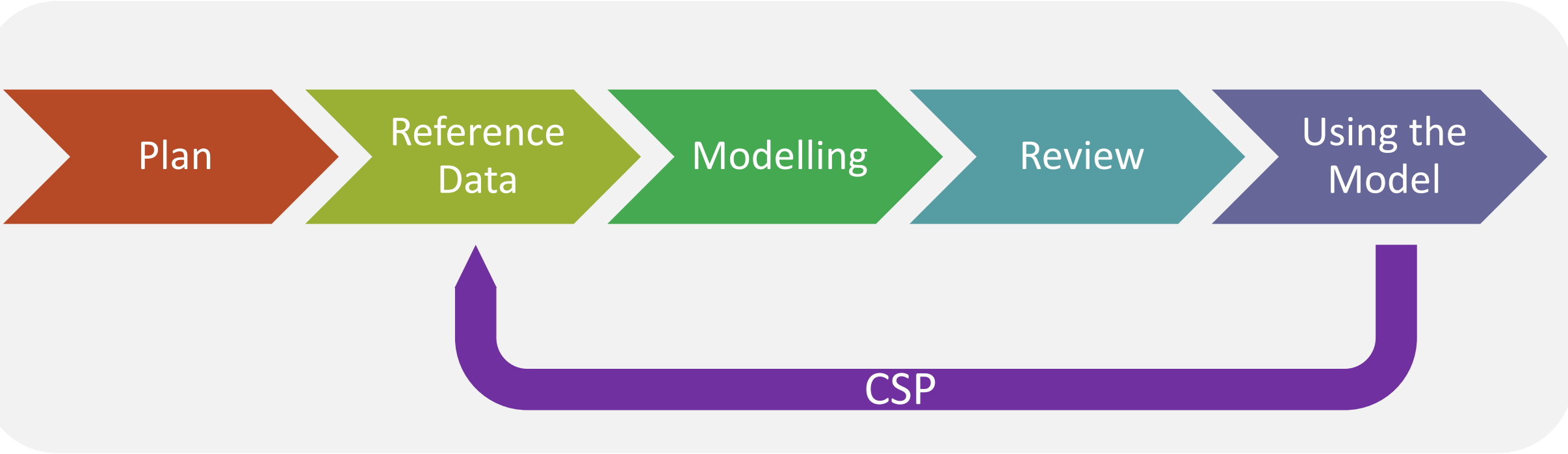
Concentrations of individual species in sample





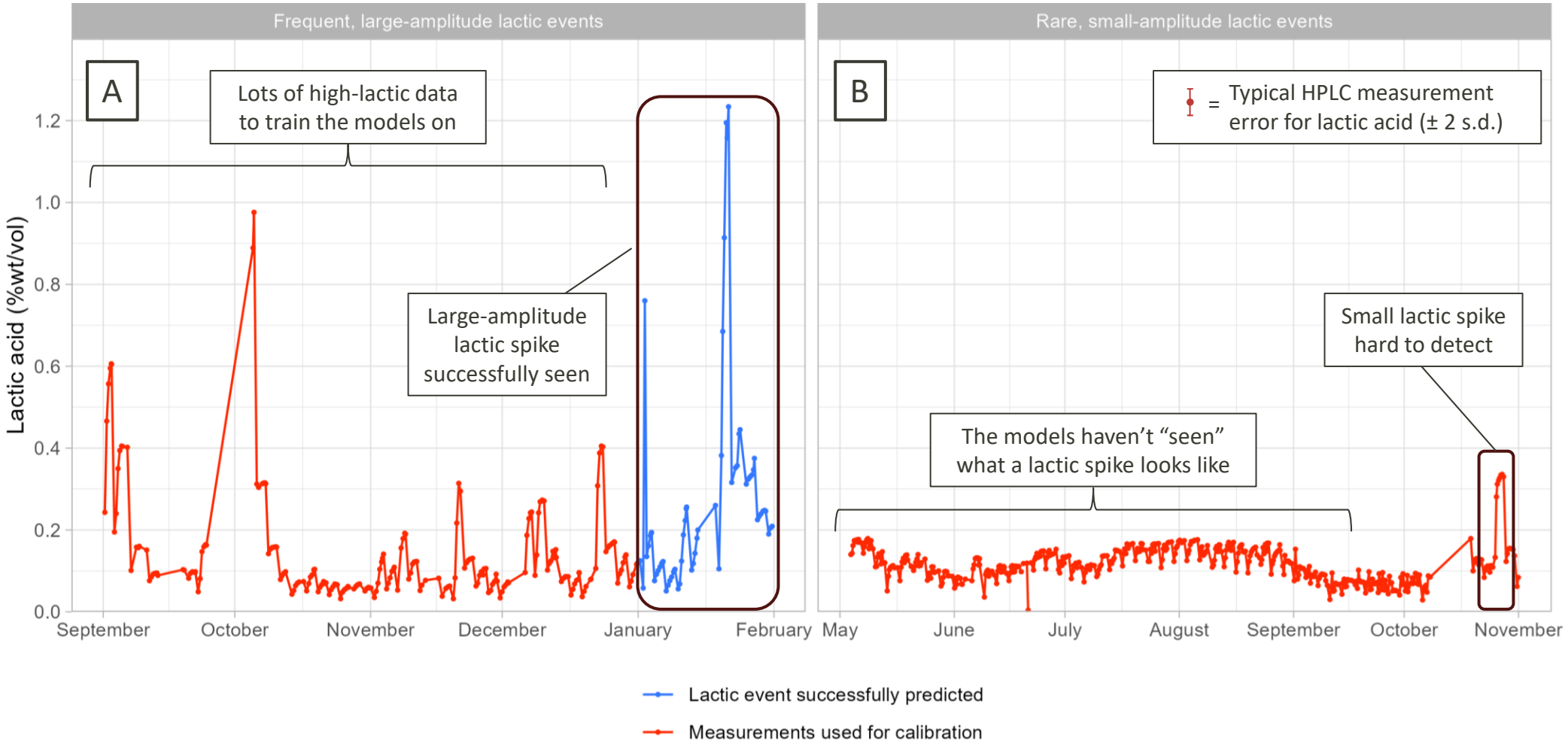


The Calibration process

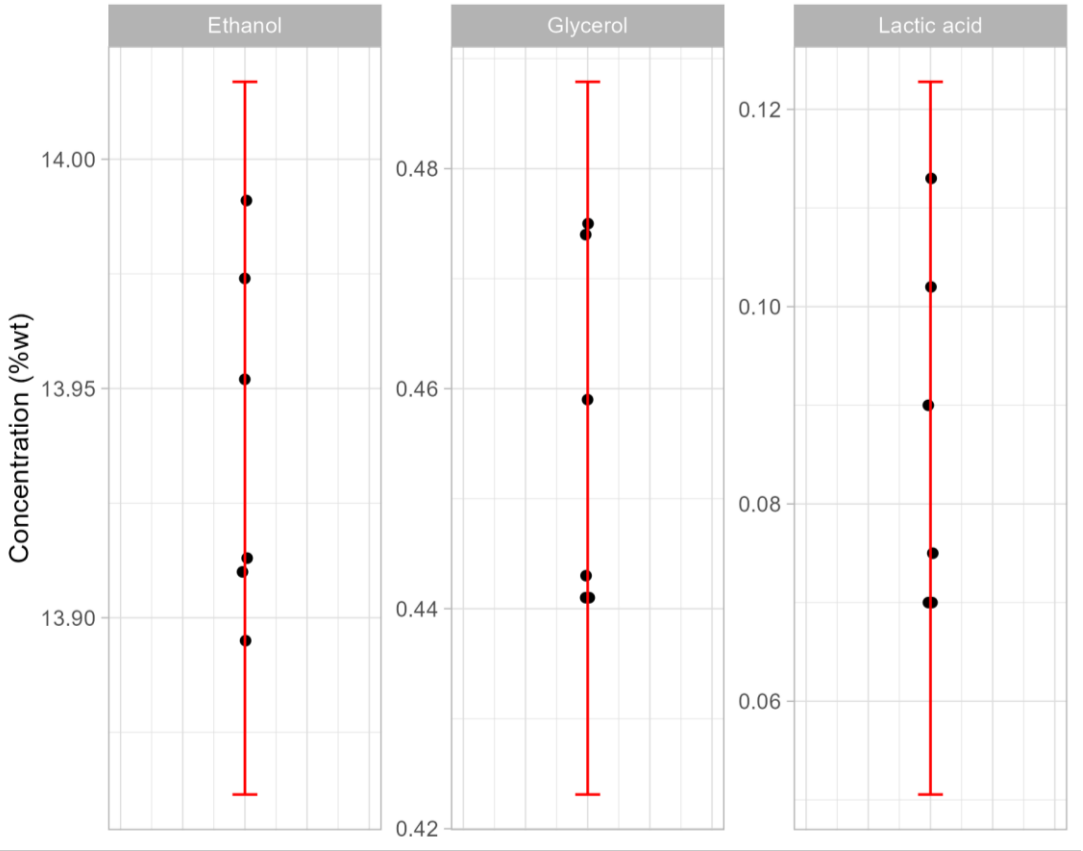


- This the standard KEIT Calibration process that is applied when building new models.
- If this is a repeat purchase for the same application or the installation is using *global models* then the process may vary.

Variation requirements



How Keit can help: R&R studies

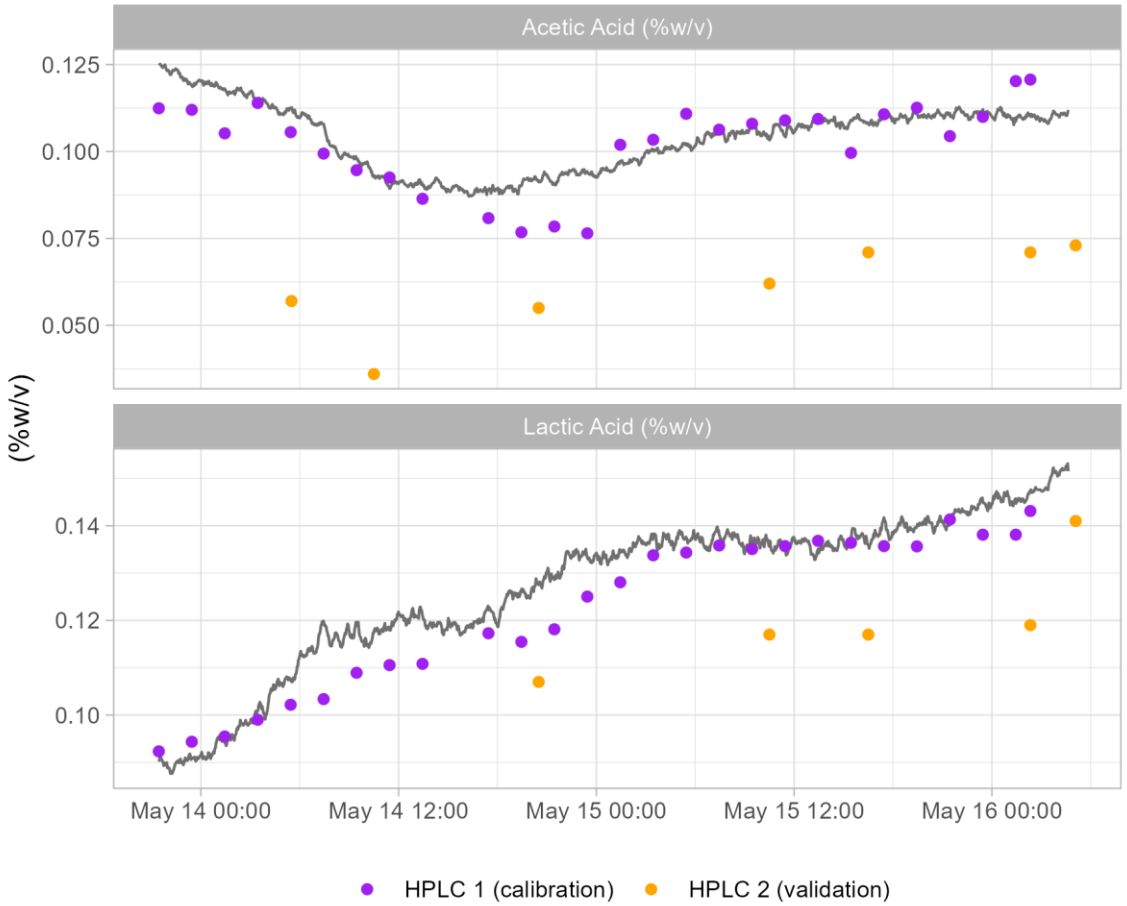


Keit can provide guidance on experiments and statistical techniques that measure and quantify reference data uncertainty.

- **Repeatability:** spread of measured values
- **Reproducibility:** inter-operator variability

These discussions have often prompted customers to improve their reference measurements.

Using two instruments



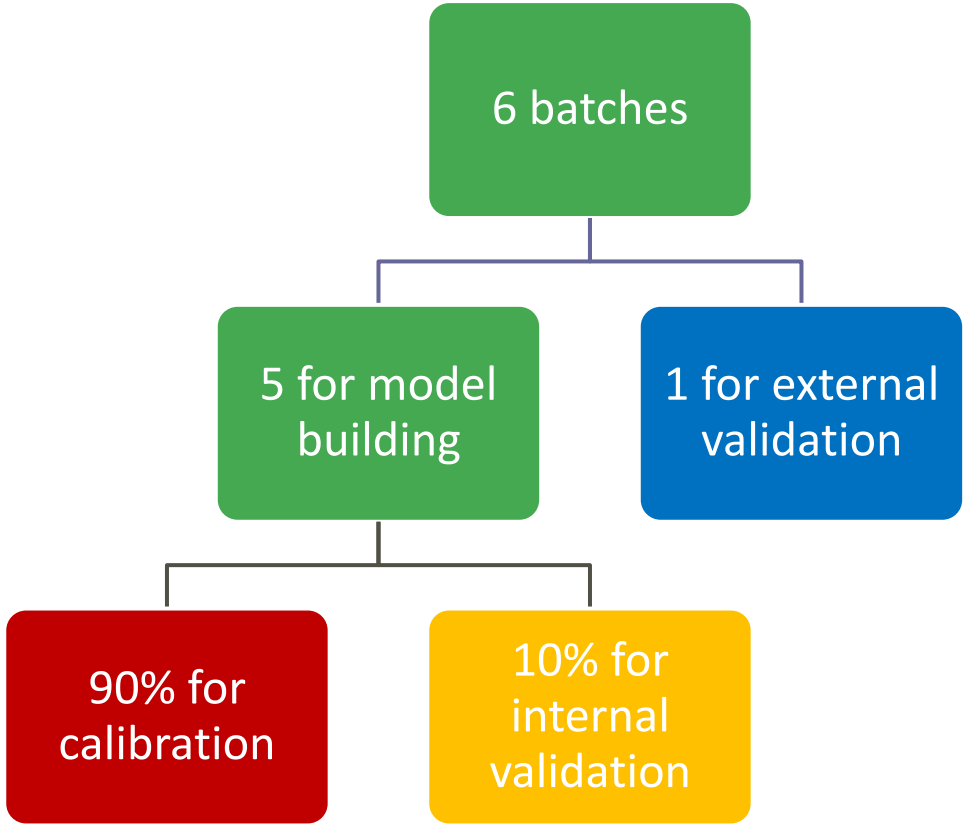
This customer used two HPLC instruments which were assumed to be equivalent.

HPLC 1 was used to build the models, but these were validated using routine samples run on HPLC 2.

They were unaware of a systematic bias in their measurements.

Building and Testing the Models

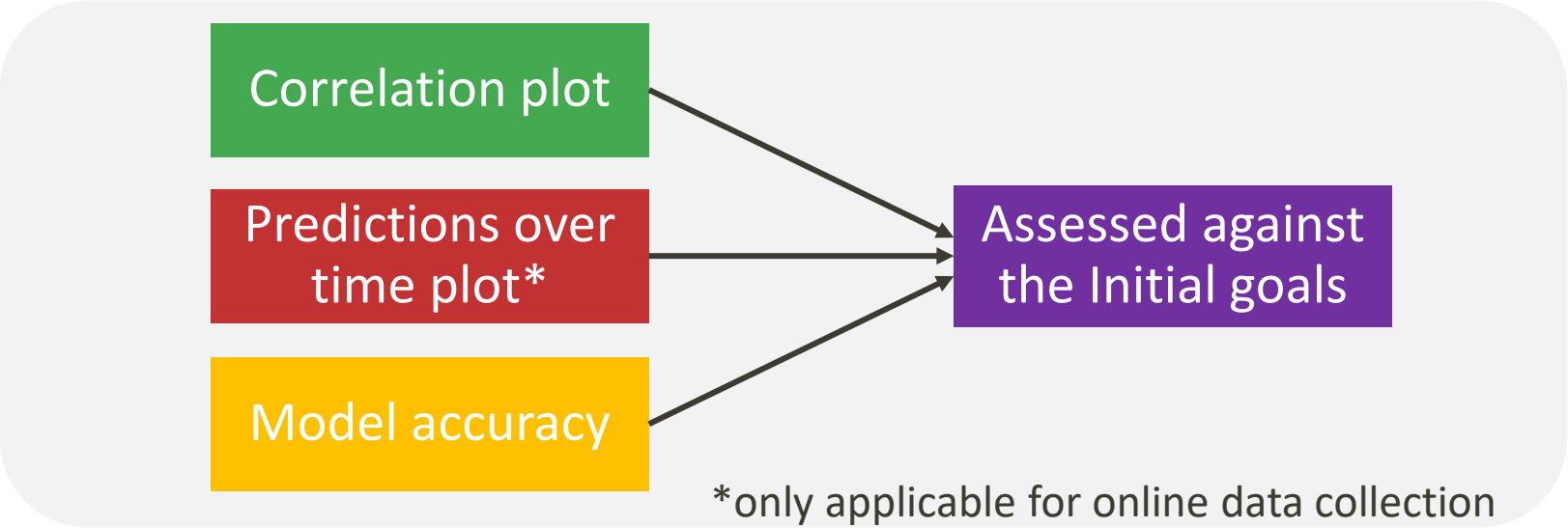
- Data is taken and split into 2 sets.
  - One set for external validation
  - One set for model building
    - This is then split to calibrate/assess the model
- Calibration work is queued at the point of receiving the data



An example of splitting 6 batches for a model

Reviewing the Models

- We present the results to you:



- Followed by a joint discussion over the next steps we can take

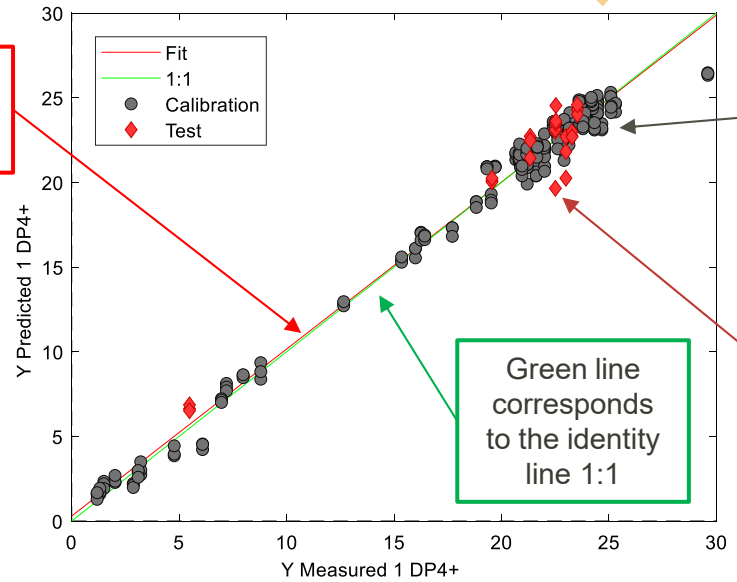
Presenting Results

Report Example – Correlation plot

Species being predicted (units)

DP4+ (%w/v)

Correlation plot



Red line corresponds to the regression model

Grey points are the calibration samples used to train the model

Red points are the internal validation samples used to test the model

Green line corresponds to the identity line 1:1

Typical information shared in the calibration report is described below. The content may change, and it may be adapted.

Accuracy and calibration range

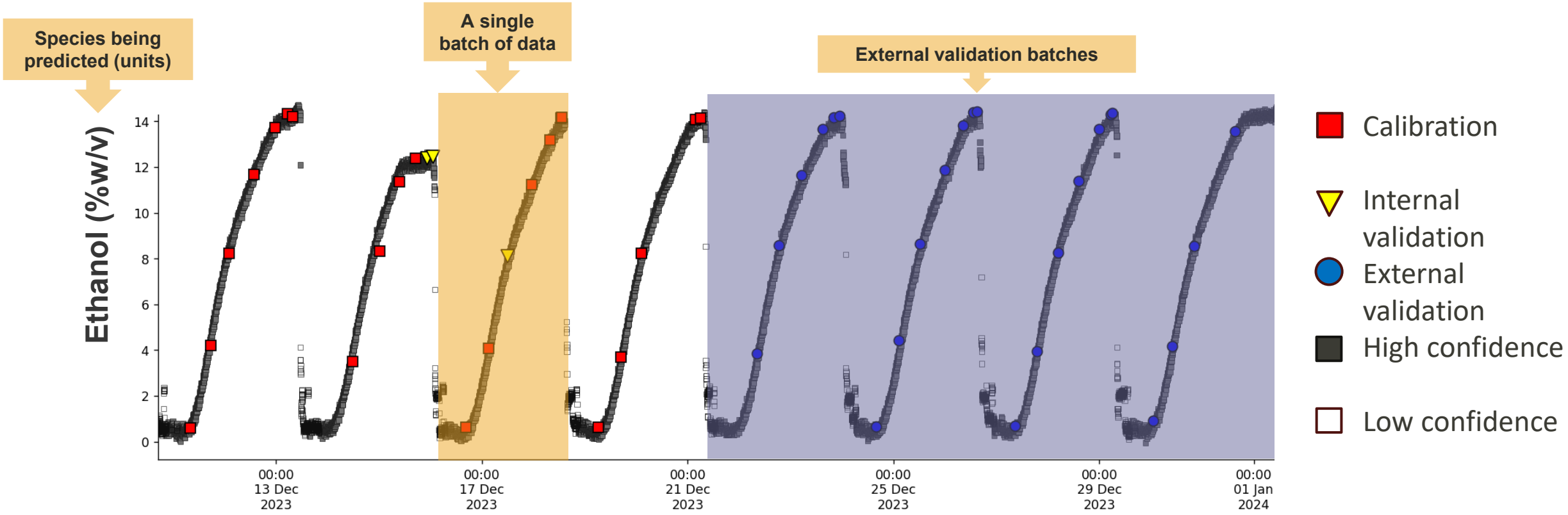
Parameter	Value
Accuracy	1.16 %w/v
Range	0 – 30 %w/v

- Data mainly clustered around the higher concentration range.

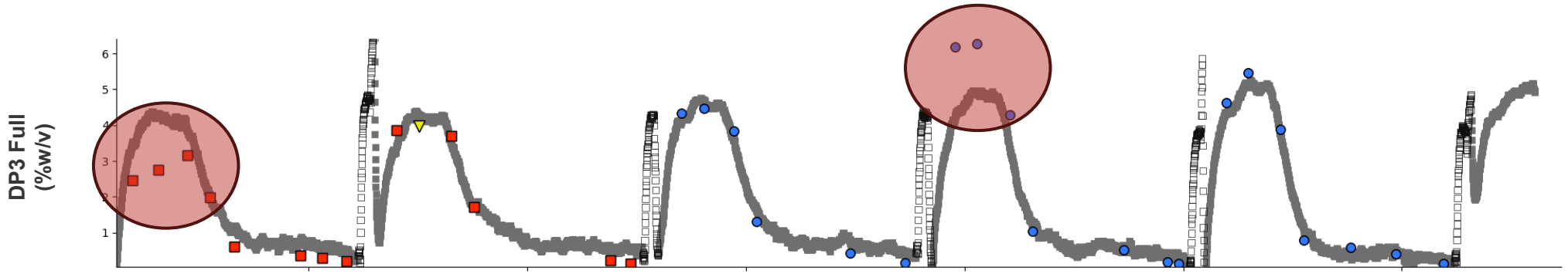
Qualitative notes about the calibration curve

Presenting Results

Report Example – Predictions over time plot



Which is more likely to be real?



The example above shows contradictory readings between the IRmadillo and the HPLC.

While HPLC is the industry standard, it is not perfect, errors can occur.

We have also shown that **the IRmadillo can be trained to be more accurate than the HPLC**.

It is important to evaluate both systems when reviewing results.

Using the models & getting Value

Time	Freeing up the operators by reducing sampling
Insight	Off-line analysis enabled by continuous data
Awareness	Trending in the DCS
Alerts	Alert operators to deviations from the historical batch
Improve	Advisory input for how to improve a current batch
Optimize	Advanced Process Control

CSP – Calibration Service Program

CSP exists because calibration is an ongoing process. When your process changes, calibration models may no longer report appropriate values unless that change is captured.

CSP covers you for this eventuality. If the models need refreshing, we follow the same process for building the first set but encompassing new data to improve the models and better capture your process.

Over time, your models will capture more and more of the variation in your process.

KEIT INDUSTRIAL ANALYTICS

Measure | Trust | Improve

Appendix

Accuracy = ± RMSECV

- Root mean squared errors (RMSE) are an average of the difference between the measured and predicted values.
- RMSECV is the root mean squared error of cross validation. It provides an indication of the how model performs over the range of the model and is defined as the accuracy (±) for our models.
- RMSEP is the root mean squared error of prediction. It is measure of how the model performs on unseen data. During model development, Keit retains some of the data supplied to test the model, both internal and external to the calibration data.