CRITICALITY-BASED PROCESS DATA EVALUATION

CARBOGEN AMCIS
A Dishman Group Company

A tool to navigate process data with a focus on process risks and opportunities

The development of robust API processes involves extensive process and analytical data difficult to navigate without powerful tools for data evaluation and visualisation. Criticality-based process data evaluation substantially accelerates exploration of large process datasets, by directing the attention towards the most significant process risks and opportunities.

A criticality dashboard paired with advanced filters supports the communication of the most relevant results by visualising the complete extent of the underlying process data while instantly highlighting the most essential points to an audience that may not be familiar with the intricate details of a particular process.

In the first 3 years after initial implementation, the criticality concept had a significant impact on communication, collaboration, and decision-making between cross-site and cross-functional validation project teams in the interest of process robustness, quality, budget and timelines.

Exploration of extensive multidimensional process and analytical datasets is a significant bottleneck in the absence of proper data management tools that dynamically merge, transform, filter and visualise data from different sources to extract relevant information in a general or specific context.

Process knowledge emerges from process data connected to the matching analytical data, evaluated against the corresponding specifications.

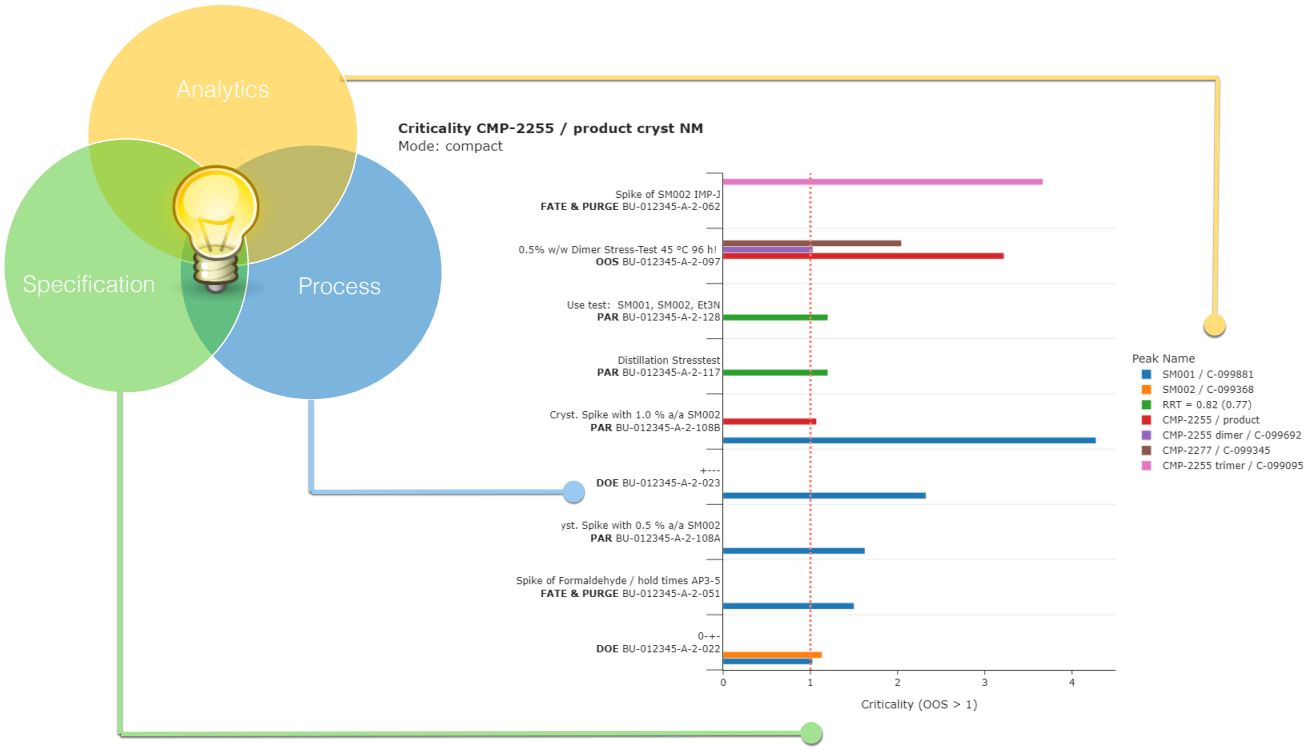


Figure 1: The criticality plot combines process data with quality attributes[1] normalised against their specification limits (OOS > 1, as indicated by the dashed line).

[1]In the current version of the tool, the criticality concept captures HPLC purity data. An extension to cover other types of quality attributes and process parameters is in the pipeline.

CRITICALITY-BASED PROCESS DATA EVALUATION

CARBOGEN AMCIS
A Dishman Group Company

Triggered by a LEAN project in 2019, development of a data pipeline for faster access and dynamic navigation of process and analytical data was officially initiated at CARBOGEN AMCIS. At the core of the initiative is the broadly applicable concept of criticality, which is outlined in further detail below.

An initial version of a criticality-based process data evaluation tool was implemented in 2020, using Microsoft Excel. While Excel macros support some rather basic data cleaning and evaluation tasks, they clearly do not cut it, when it comes to dynamic transformation and merging of large datasets across multiple dimensions. Even in the context of smaller projects, a significantly more powerful and object oriented platform is needed to address the data-heavy tasks in a structured and performant way.

Supported by the very positive feedback from initial customers and users, it was decided to migrate the initial Excel tool to a fully interactive Jupyter Lab data science platform. The new environment, accessed via the web-browser, executes major data transformation tasks on the fly and offers a lot more flexibility when it comes to extending the functionality or adapting the tool to specific needs.

Currently, the following modules are implemented and presented in the following sections:

- Criticality-based Extraction of PAR-Ranges and Severity
- IPC Criticality
- IPC Purge Models
- Root Cause Analysis based on Standardised Data

CRITICALITY: A SIMPLE BUT POWERFUL CONCEPT

Criticality is the normalisation of a quality attribute against its specification limits: OOS results have a criticality > 1, while results with a criticality ≤ 1 are within their specification limits. In contrast to the raw data which typically stretches across multiple dimensions and orders of magnitude, dimensionless criticality is normalised to the same scale, which allows to visually summarize and directly compare any numerical quality attribute within the same plot. When labelled with the corresponding process information, the criticality plot provides an instant insight into the underlying process risks and opportunities. Deviations from typical patterns, or gaps in the data, are easily spotted in the criticality plot, even across a large number of quality attributes.

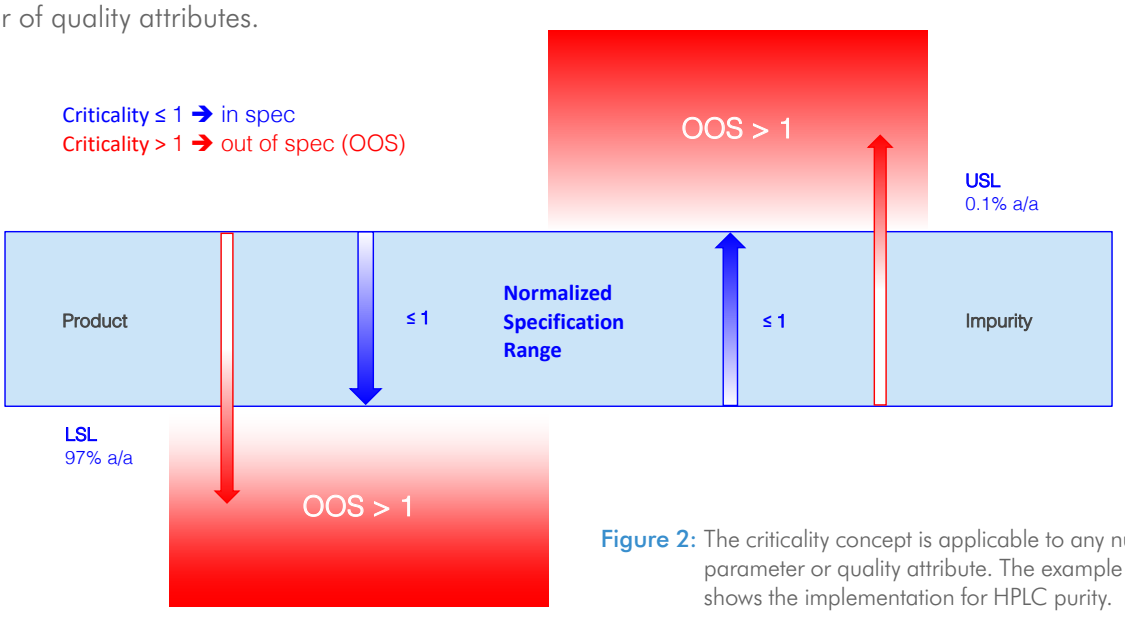


Figure 2: The criticality concept is applicable to any numerical parameter or quality attribute. The example here shows the implementation for HPLC purity.

Using the criticality-based approach, we have successfully established a control strategy and completed a series of very challenging PAR-studies for a process with >100 impurities in a telescoped sequence of metallo-organic reactions in a flow-chemistry setup.

CRITICALITY-BASED EXTRACTION OF PAR-RANGES AND SEVERITY

The fact that experiments with criticality ≤ 1 meet their specifications implies that any underlying process parameter lies within its proven acceptable range (PAR). As a consequence, PAR-ranges can be automatically extracted based on the process parameter ranges covered by the experiments with criticality ≤ 1.

Multiplication of the criticality with the probability of failure for each quality attribute represents a metric for the quality impact or severity inherent to the process parameters tested in an experiment. The tool translates the severity value into a risk priority number (RPN) which may be used as a factor in quality risk assessments.

Tests in the context of a recently executed PAR-study showed that both, the PAR-ranges, and the severity RPN obtained based on criticality was in excellent agreement with the levels assigned by an experienced process chemist.

IPC CRITICALITY

The criticality concept is equally applicable to the evaluation of IPC results. IPC criticality provides valuable preliminary insights into potential process risks based on IPC purity data (Figure 4).

As a first step, an IPC purge model (for details, see section below), is established by correlation of the impurity levels in the isolated product with the corresponding IPC purity data.

Acceptance criteria (IPC limits) for impurities in the IPC are estimated either by plugging impurity specification limits for the isolated product into the IPC purge model or by extrapolation of the regression trendline to the specification limit, the latter being less sensitive to outliers in the data. IPC limits estimated based on IPC purge factors are useful in the evaluation of reaction DoEs (Figure 3) and they are the basis for IPC criticality (Figure 4).

Purge factors are all but invariant. IPC purge models, based on the complete extent of the available purity data, reflect the variability in impurity levels caused by variations in process parameters within the individual IPC samples (Figure 7).

The user has the option to select between the different types of purge models in combination with conservative or extrapolated IPC limits (Figure 5).

Evaluation of IPC data from a reaction DoE against the estimated IPC limits for the impurities identified to be relevant by IPC criticality indicates the potential for improvement of the reaction conditions, which cannot be visualised based on the parameters typically controlled by the IPC alone (e.g. conversion).

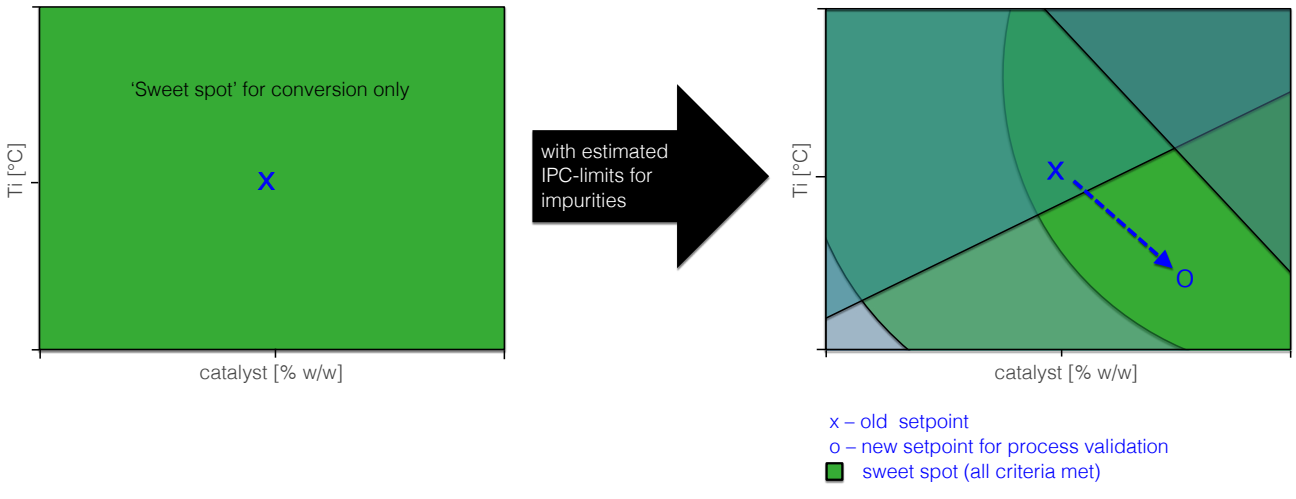


Figure 3: Reaction DoE evaluation against acceptance criteria for impurities, estimated based on IPC purge factors.

CRITICALITY-BASED PROCESS DATA EVALUATION

CARBOGEN AMCIS
A Dishman Group Company

When timelines are tight, IPC criticality helps to prioritise work-up and isolation of experiments based on the IPC purity data. Experiments with a high or untypical IPC criticality usually need to be isolated under standard conditions to assess their potential impact on product quality. Reaction mixtures with a low or typical IPC criticality on the other hand are useful as representative input material to investigate unit operations further downstream.

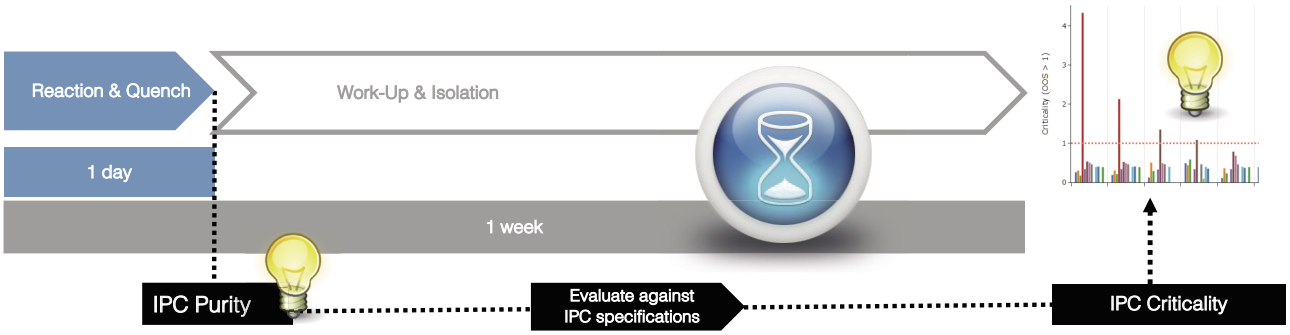


Figure 4: Based on IPC purity, resources for work-up and isolation are allocated to the experiments with the highest IPC criticality.

Since the IPC criticality systematically tracks all impurities from the evaluated IPC samples, the probability of overlooking a potential quality risk hidden in the IPC data is significantly reduced.

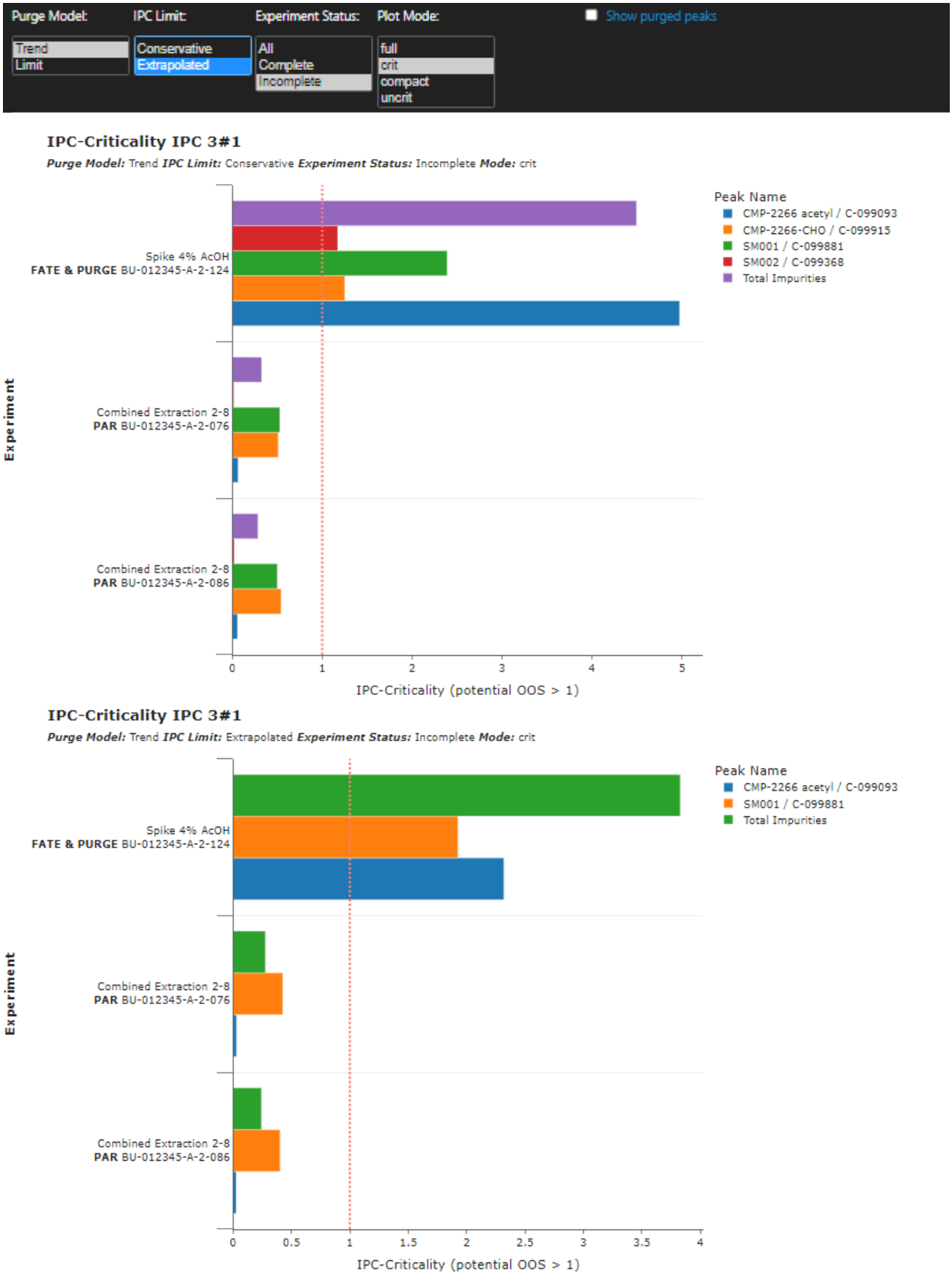


Figure 5: Evaluation of IPC criticality against the 'Conservative' (top) vs. the 'Extrapolated' IPC limits (bottom). Trend-based purge models are typically less outlier-sensitive than Limit-based models.

IPC PURGE MODELS

IPC purge factors correspond to the slope of the linear regression trendline obtained by correlation of the impurity levels in the IPC purity data. IPC Purge Factors are summarised in a bar plot with levels > 1 indicating enrichment (Figure 6).

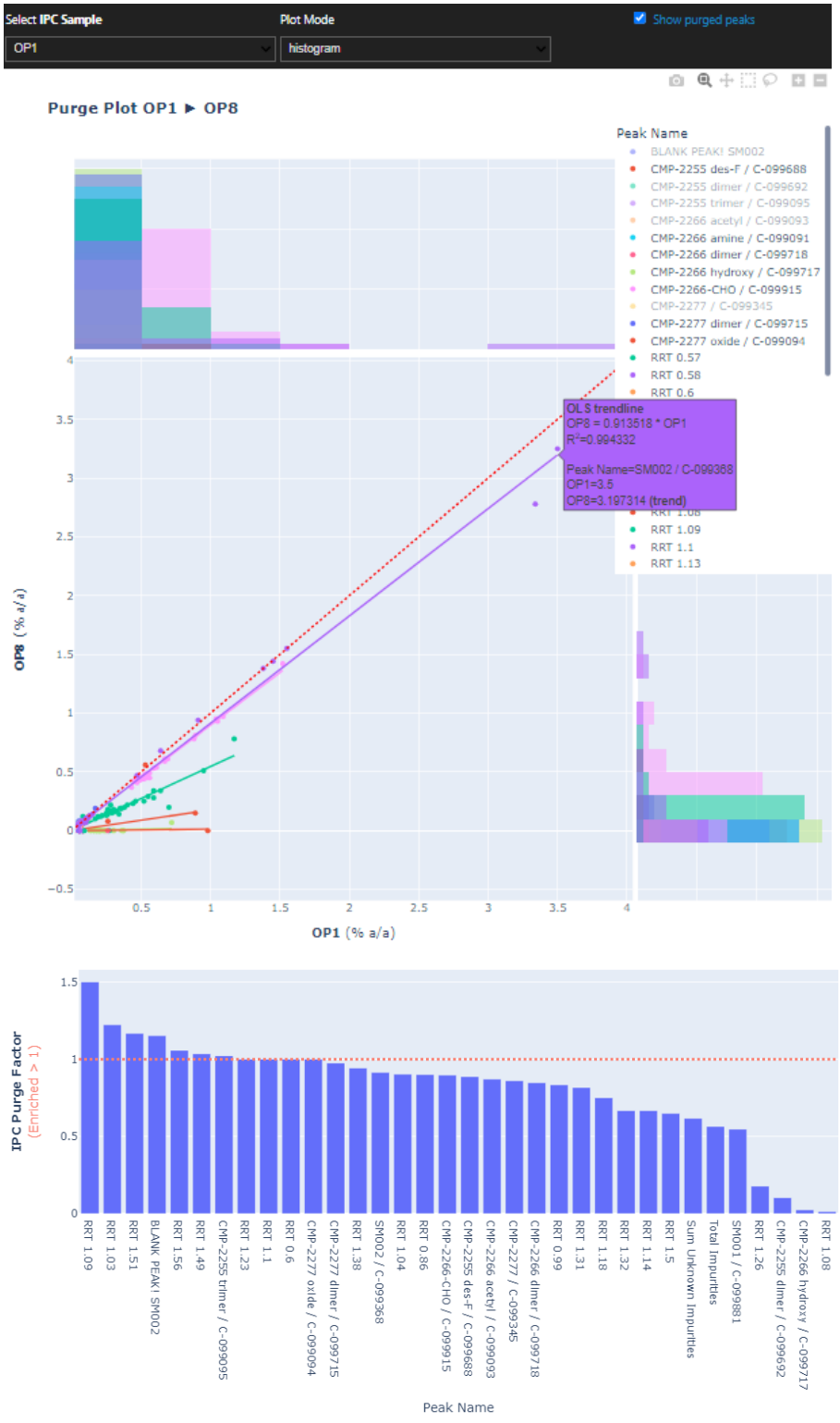


Figure 6: IPC Purge Factors between any two points in the process are obtained based on linear regression (top) and summarised in a bar plot (bottom). The example here shows purge factors across 8 extraction steps.

For further inspection, IPC purge data for each individual impurity is visualised in a scatter plot including the USL from the product specification and IPC limits (Figure 7).

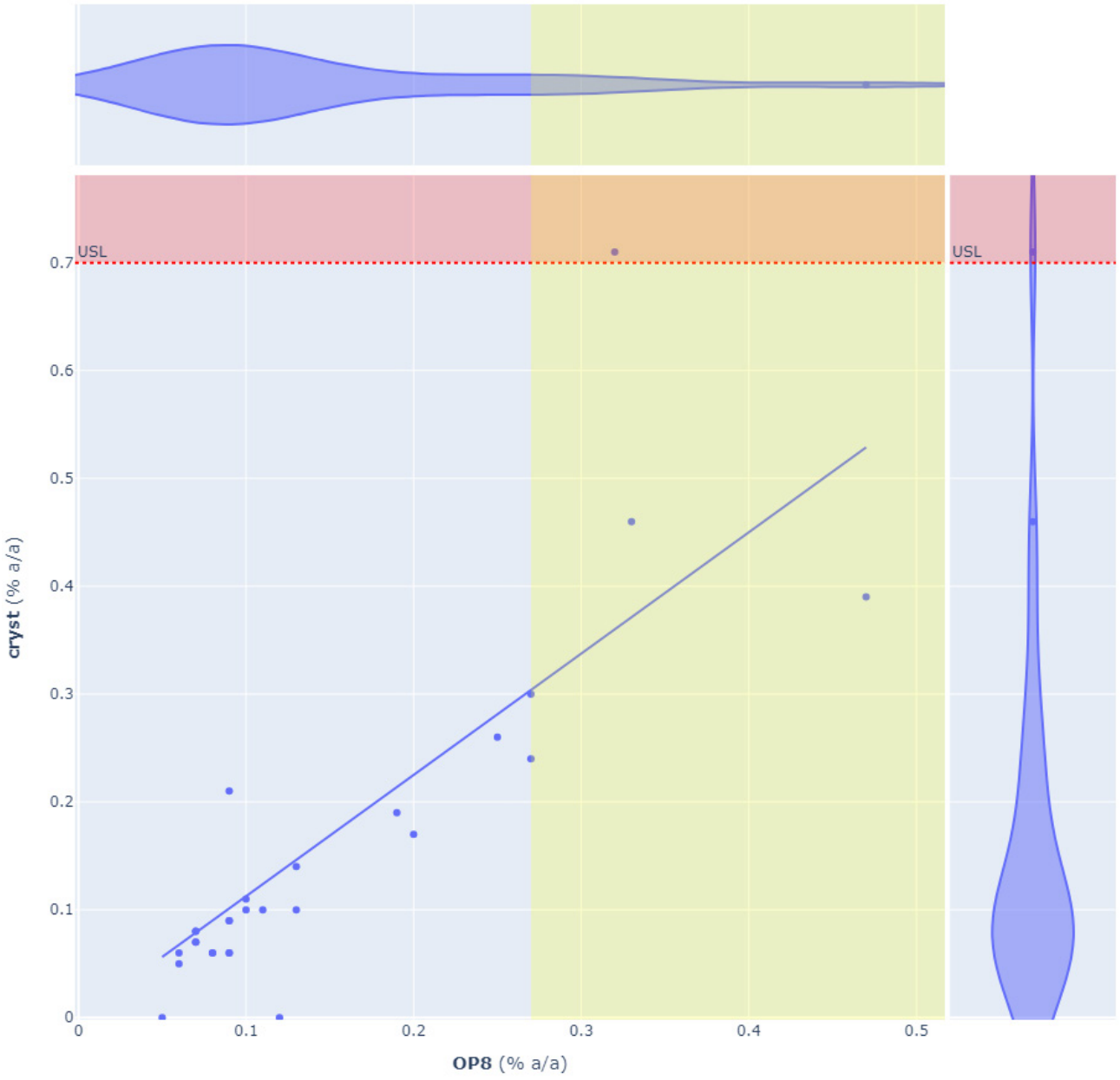


Figure 7: Impurity-specific purge data with highest tolerated IPC level (left edge of the yellow area) and the upper specification limit (USL) for the isolated product (red). The plot shows that (in this example), the IPC limit obtained by extrapolation of the trendline to the USL is significantly higher than the highest tolerated IPC level at the left edge of the yellow area. Data points within the yellow area but below USL belong to experiments OOS in a different quality attribute. The user has the option to select between the conservative or the extrapolated limits, where applicable (see also Figure 5).

Whenever an experiment with a potentially critical impurity is isolated within specifications (criticality $\leq$ 1), the highest tolerated IPC level for the impurity is updated, pushing the yellow area further to the right.

When an impurity is not impacted by any of the process variations, there may be a very high correlation coefficient ($R^2 > 0.99$) of the purity data in the purge plot. A low correlation coefficient does not mean poor data quality, but a higher impact of process variations on the impurity levels.

ROOT CAUSE ANALYSIS BASED ON STANDARDISED DATA

If an experiment turns out to be OOS, it is essential to find a root cause. Based on data standardisation (distance of the value from the mean normalised against the standard deviation), a module for impact assessment/root cause analysis links outliers in the quality attributes (responses) to the corresponding outliers in the process parameters (factors). The interactions of outliers between factors and responses, above a user-defined threshold, are visualised to rule out or to identify a potential root cause within the tracked process parameters.

PROCESS DATA NAVIGATION

Criticality provides the high-level compass for in-depth navigation of the underlying process data.

The dashboard is fully functional with early-stage experiment level data, also in the absence of detailed process parameter data.

The evaluation tool offers a workflow starting with a dashboard summarizing the most relevant high-level project and process information based on criticality and additional filters applied to the merged data (Figure 8).

Figure 8: High-level information, such as experiment status and yields are displayed in the dashboard.

The dashboard offers an overview on the general status of the project regarding yields, project progress, criticality, PAR-ranges and the expected quality impact (severity) extracted based on criticality and probability of failure of the complete underlying data.

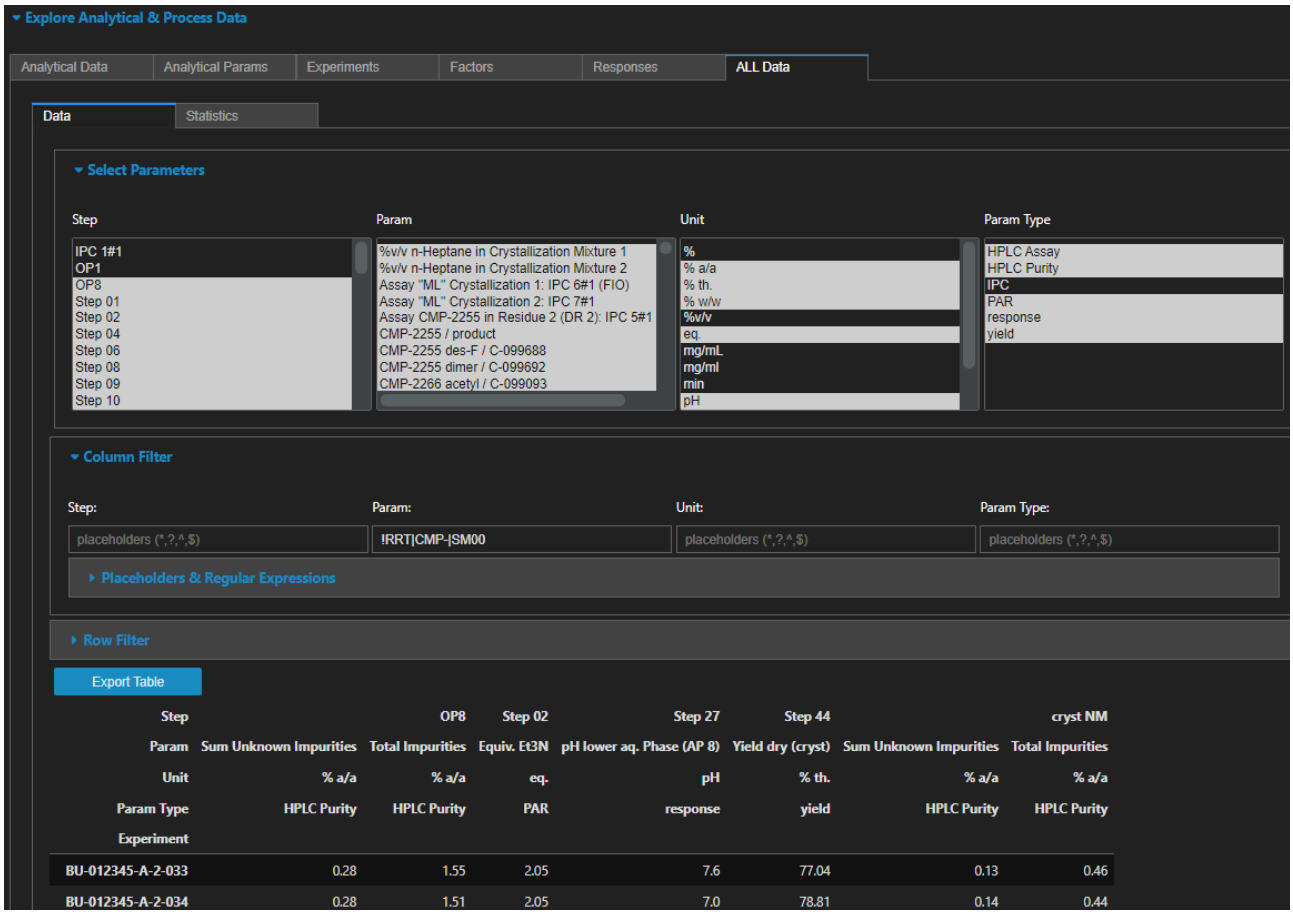
The interface also offers a summary of incomplete or misaligned data, as a hint to better align the data or to render the process more robust.

DATA EVALUATION AT THE PROCESS PARAMETER OR QUALITY ATTRIBUTE LEVEL

For deeper investigations, the tool offers interactive filters for in-depth evaluation of the process data in a specific context down to the process parameter or quality attribute level of detail.

Every filter widget is equipped with a statistics tab that shows the typical parameter ranges including the mean and standard deviations, as well as a histogram indicating the number of the data points supporting the statistics. Resulting data tables are exportable for sharing, for further evaluation, or for reporting (Figure 9).

Users familiar with the Python programming language can use the power of the Jupyter Lab platform to extend the functionality, or to further explore the data according to their specific needs



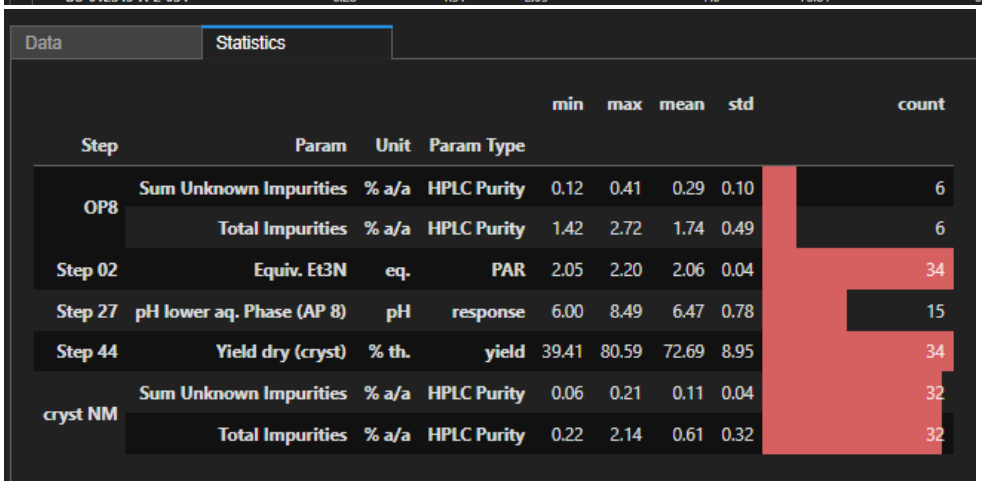


Figure 9: For detailed evaluation, analytical and process datasets are filtered based on user input, then merged. Numerical parameters are summarised in a statistics tab.

CRITICALITY-BASED PROCESS DATA EVALUATION

CARBOGEN AMCIS
A Dishman Group Company

FLEXIBILITY THROUGH A MODULAR APPROACH

The functionality of the different modules reflects the extent of data typically available at the different stages of a project. For early phase projects, the extent of manually collected process parameter data is usually rather limited, while the body of the electronically tracked analytical data may already be substantial. Therefore, the dashboard is fully functional with experiment level data, in the absence of detailed process parameter data.

OUTLOOK

This document covers the functionality of the new Python-based process data evaluation tool as of August 2023. Additional features, such as a module for the exploration of specification scenarios are in the pipeline. The functionality of the tool will be improved and further extended based on user feedback and input.

The modular architecture of the tool offers the flexibility to be adapted based on task-specific, data-specific or user-specific requirements at any time.

ABOUT CARBOGEN AMCIS

CARBOGEN AMCIS has more than 40 years of experience in drug development and commercialisation services to the pharmaceutical and biopharmaceutical industries, at all stages of drug development. Our capabilities span from contract chemical process research and development to the supply of Active Pharmaceutical Ingredients (APIs) - as well as drug products for preclinical studies, clinical trials and commercial use.

CARBOGEN AMCIS covers the complete development pipeline starting from small-scale early phase API process development, strategic sourcing of raw materials, scale-up from gram to kilogram quantities, process characterization, process optimization, and process validation, including regulatory support and project management services along the full lifecycle of a product.

CONTACT INFORMATION:

Lukas Brändli, Manager PR&D
lukas.braendli@carbogen-amcis.com
www.carbogen-amcis.com