OCTOBER 2023

Selection and Use of AI, ML, and Chemometrics in Biopharmaceutical Manufacturing: Challenges and Considerations

Johan Hultman

SARTORIOUS

AVEVA

SARTORIUS

Simplifying Progress

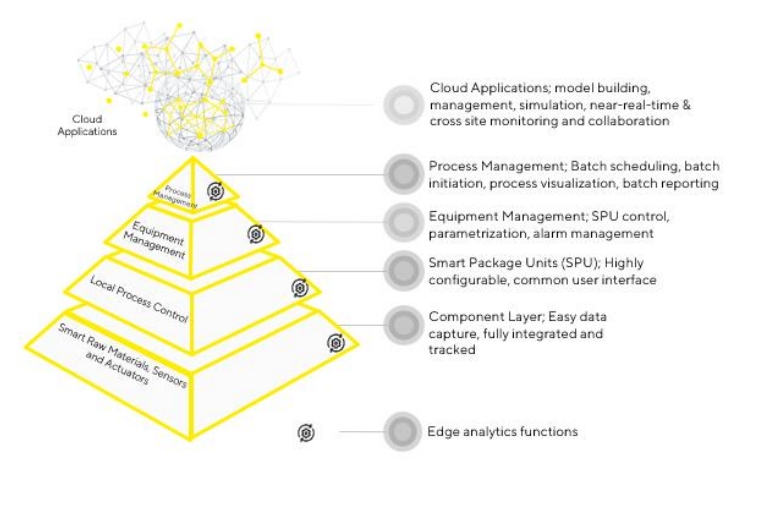
Selection and Use of AI, ML, and Chemometrics in Biopharmaceutical Manufacturing: Challenges and Considerations

Johan Hultman

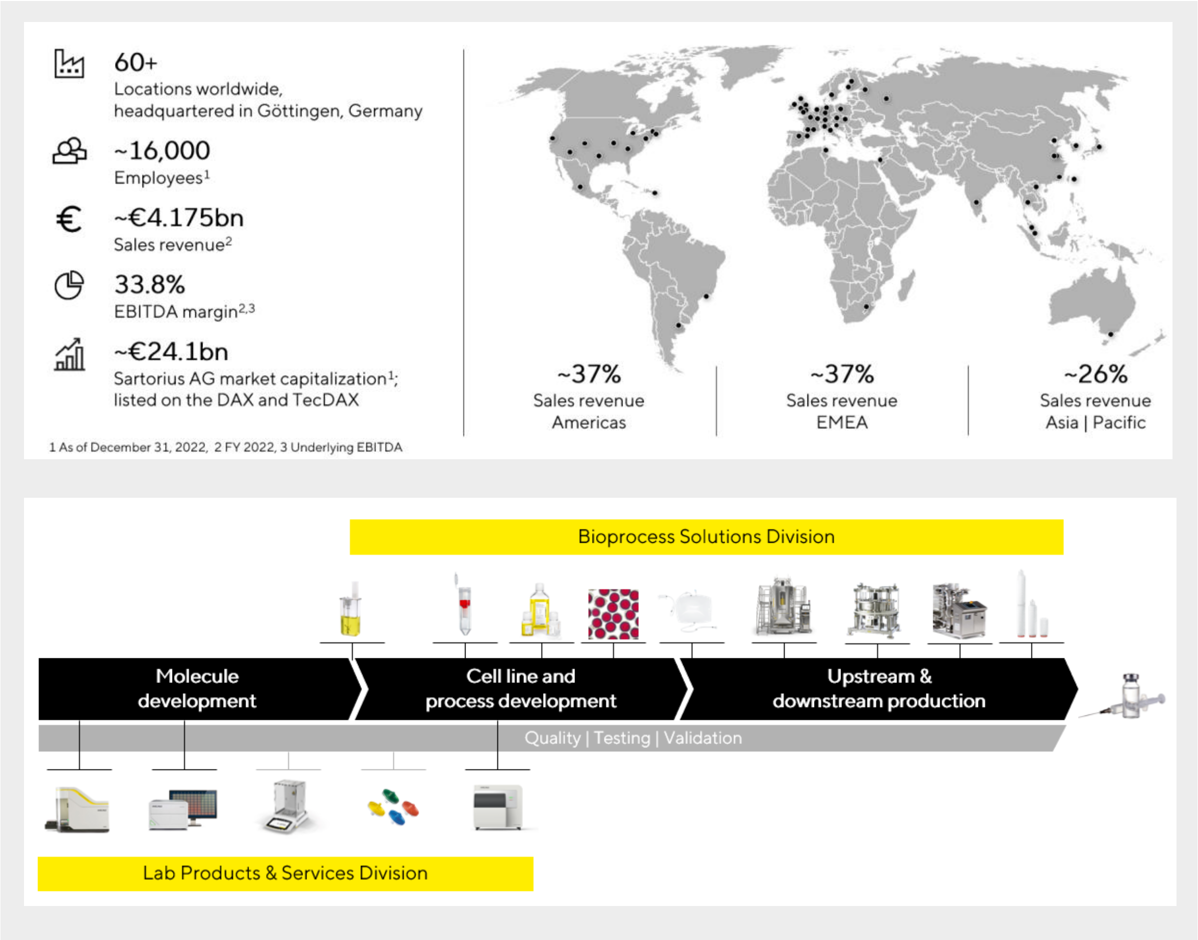
Sartorius Digital Solutions

Agenda:

- Sartorius intro
- Business Challenge:
  - AI, ML, and Chemometrics in Biopharmaceutical Manufacturing
- DPMM
- AI, ML, & Chemometrics:
  - Definitions
  - Methods & Algorithms
  - Differences in Approach Challenges
- Exempels



Sartorius



Digital Solutions Team

- Executing a digital product and services portfolio across **3 continents** and out of **9 global sites**
- **350+** internal & external employees

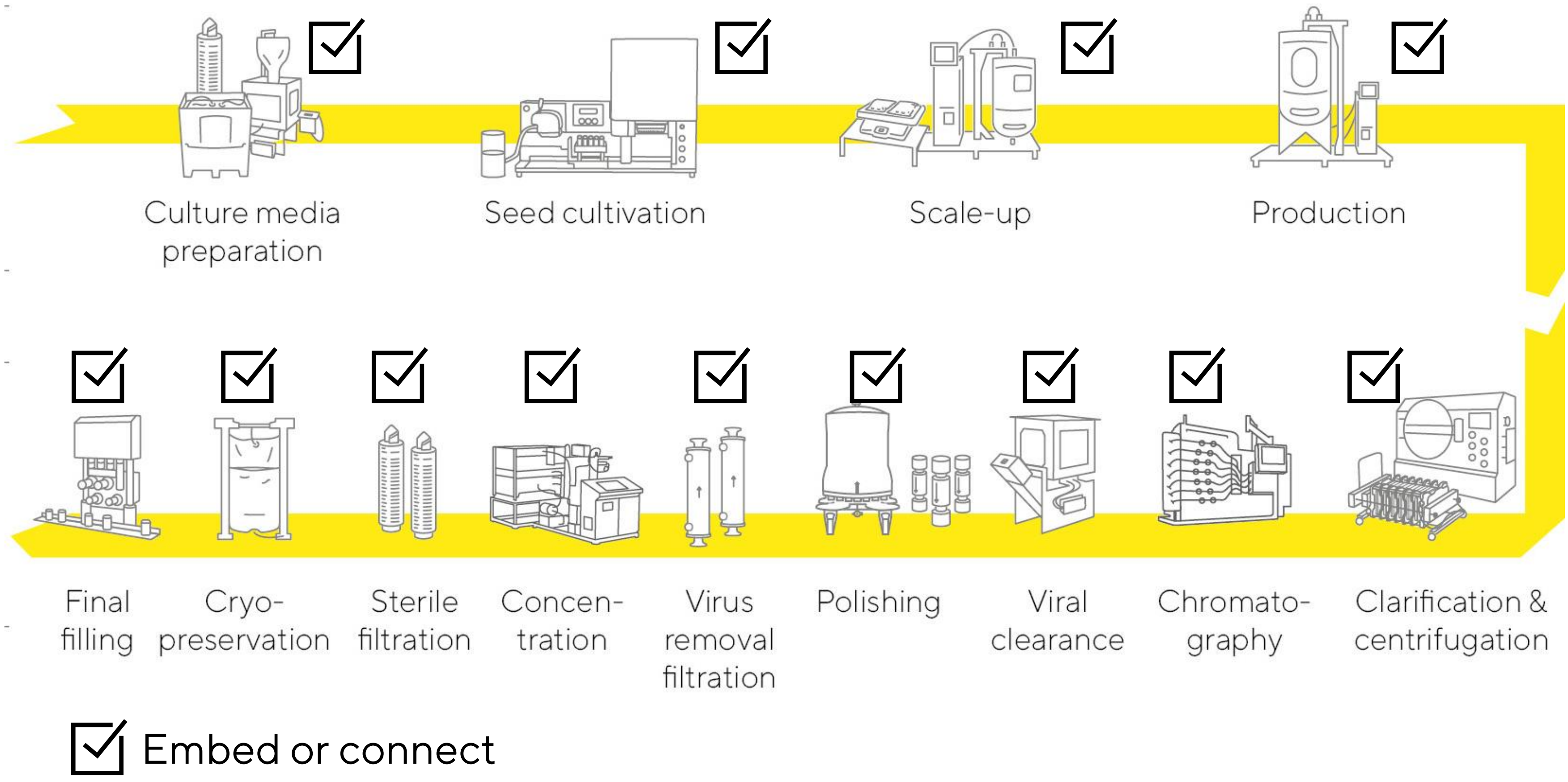
Sharp Focus on Data-Driven Technologies

- That empower end users with **data-driven** decision-making skills to get new medicines to market **quicker, cheaper and with lower risk and complexity**

SARTORIUS

Supporting the Future for the Entire Bioprocess Value Chain

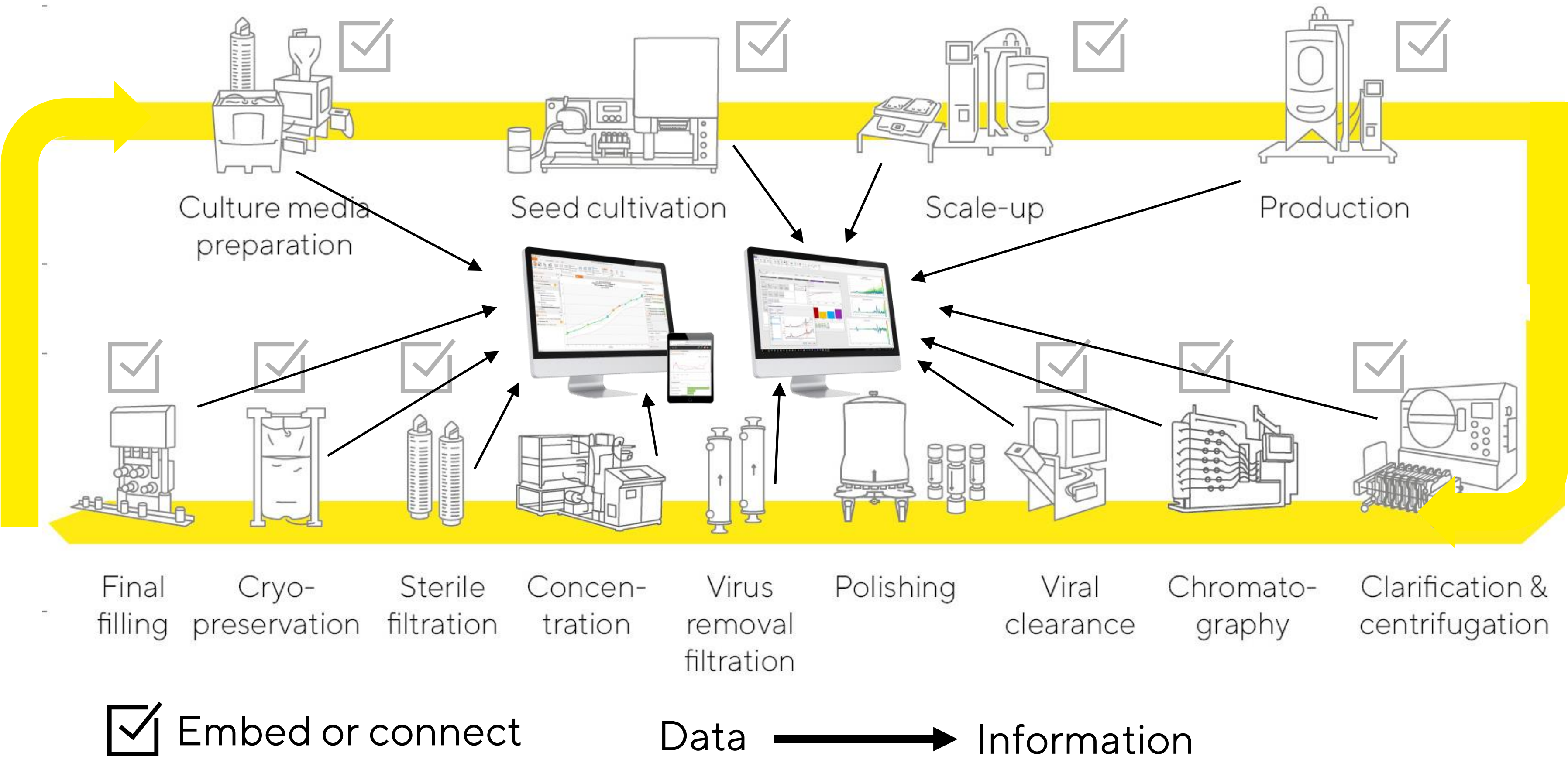
Process Development to Commercial Manufacturing



From selecting optimal cell culture media mixtures to refining chromatography process settings, Sartorius Data Analytics has tools that can solve a wide range of bioprocessing challenges.

Supporting the Future for the Entire Bioprocess Value Chain

Process Development to Commercial Manufacturing



From selecting optimal cell culture media mixtures to refining chromatography process settings, Sartorius Data Analytics has tools that can solve a wide range of bioprocessing challenges.

We have more data than ever, and we have a plan for how to use it.

SARTORIUS

DPMM - Digital Plant Maturity Model

SARTORIUS

Digital Plant Maturity Model (DPMM) 1.0

Level 1 Pre-digital Plant

- Predominately manual processing.
- Low level of automation.

Level 2 Digital Silos

- Site-specific systems; limited integration across functional silos
- Analytics on demand, “why did it happen?” high manual effort
- Plants operate independently with little “real-time” supply chain visibility

Level 3 Connected Plant

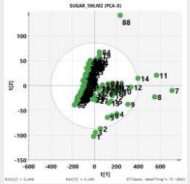
- ERP, LES, MES and Automation layer are fully integrated to support digitized business processes.
- Standard application platform adopted across plant network

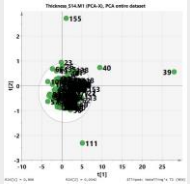
Level 4 Predictive Plant

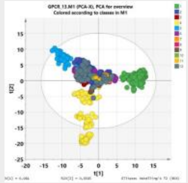
- Online/At-line quality testing with Real Time Release.
- Proactive analytics across plant and internal value chain; “what can happen and when?”
- Integrated Real-time Process analytics

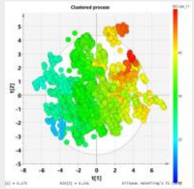
Level 5 Adaptive Plant

- In-line, real-time, continuous, closed loop, process verification and control with automated real-time quality release
- Self-aware, continuously adaptive, “Autonomous” plant; exception conditions handled by remote experts
- Trusted information insights are freely and securely available. Pervasive use of adaptive analytics and Self/Machine learning across value chain.









https://www.biophorum.com/download/digital-plant-maturity-model-3-0/

Digital Plant Maturity Model (DPMM) 3.0

Level 4 Predictive Plant

- Enterprise integration, orchestration of the end-to-end value chain through ecosystems fully integrated across process development, supply chain, maintenance, manufacturing and quality
- Digital twin models exist for certain asset classes and processes to detect anomalies and predict events in near-real-time and recommend courses of action using AI
- Integrated real-time process analytics to enable error-proofing using proactive analytics across plant and internal value chain; ‘what can happen and when?’
- Smart maintenance, utilizing real-time data monitoring

Level 5 Adaptive Plant

- Pervasive use of ‘integrated value chain digital twin’, powered by AI-driven models to autonomously intervene and take action to maintain and enhance efficient, safe, reliable and high-quality operations across the value chain
- Digital continuum allowing AI-driven tech transfer. Process and models are automatically generated using a unified modeling language, powered by AI and based on process, equipment and study data
- In-line, real-time, continuous, closed-loop, process verification and control with automated real-time quality release
- Zero system downtime (including upgrades)— continuous evolution
- Pervasive use of adaptive analytics and self/machine learning across value chain
- AI-based security risk and incident management, automating decision-making to deploy proactive security measures

https://www.biophorum.com/download/digital-plant-maturity-model-3-0/

SARTORIUS

AI, ML, and Chemometrics

AI, ML, and Chemometrics in Biopharmaceutical Manufacturing

The Challenge

The use of different types of instrumentation that generate large amounts of data have exploded in the Life Science, Food & Bev. Industry.

The way of handling the data are many but some of them are better than others and provide additional information and trust.

The Solution

We have enabled the use of such information in our own production and together with partners developed tools to help drive Digital Transformation and Automation by using all the data, *compress or extract* the hidden information and made it easily accessible from partner applications like AVEVA PI or AVEVA Historian

The Value

1. Easily use information from the trusted AI/ML prediction models in OSI PI system or the AVEVA™ Historian (Wonderware)
2. Use AI/ML monitoring and diagnostic tools to improve our process models
3. Use and compress data without losing important information – edge analytics

Artificial Intelligence (AI) in Drug Manufacturing

AI, ML, & Chemometrics/MVDA

Topic 1: Consensus on Global Definitions & Concepts

- Definitions (AI, ML, Chemometrics)
- Methods & Algorithms
- Differences in Approach
- What is MVDA?

Topic 2: The Need for AI Model Explainability

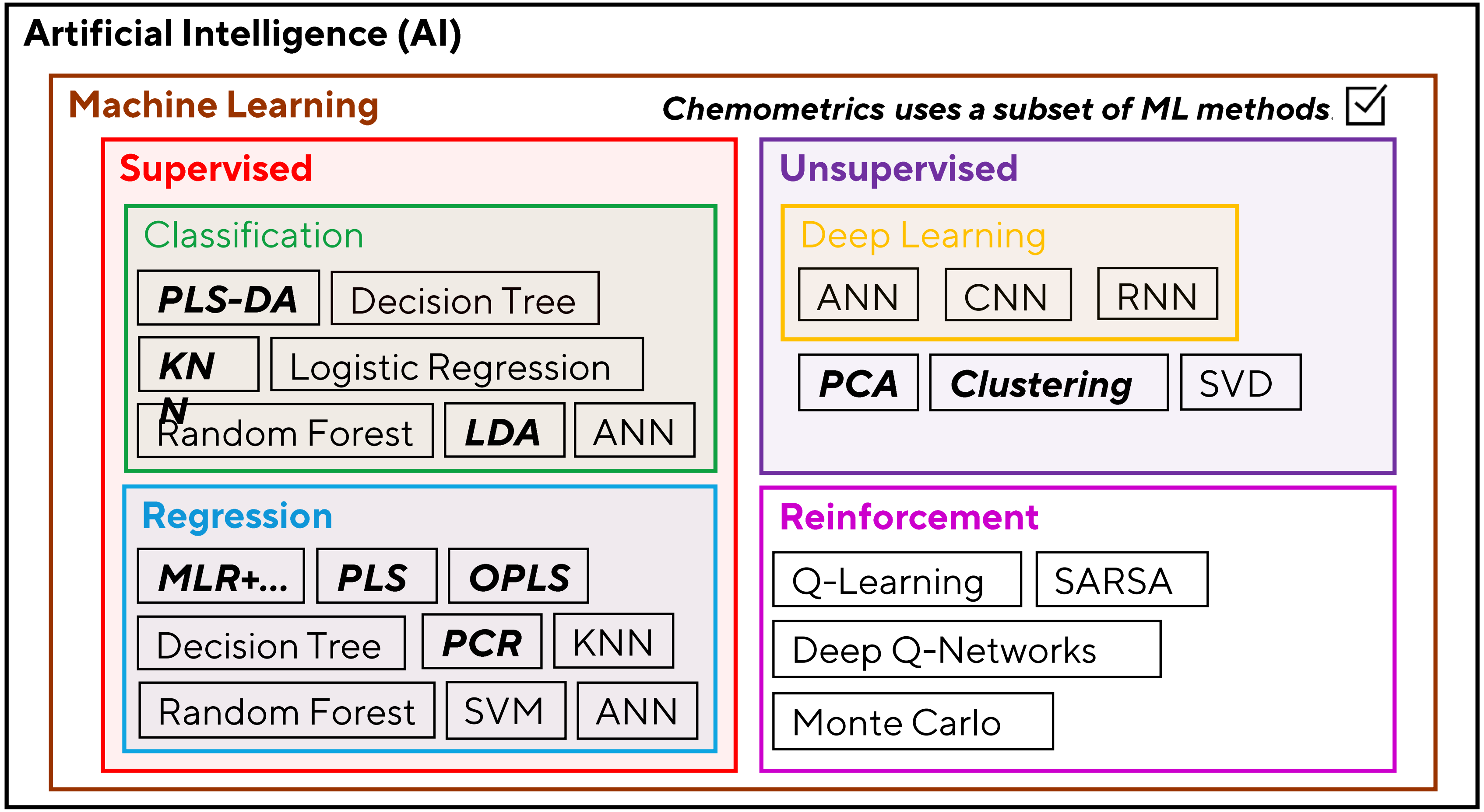
- What is model explainability?
- Explainability & MVDA
- MVDA & explainability in action

SARTORIUS

AI, ML, & Chemometrics: Definitions

	Artificial Intelligence (AI)	Machine Learning (ML)	Chemometrics
Definitions	Development of intelligent systems, including ML and other techniques, to simulate human intelligence	Algorithms and models that enable machines to learn from data and make predictions or decisions.	A discipline for analyzing chemical data and designing experiments to extract meaningful information
Center of Discussion	Sectors: healthcare, finance, gaming, customer service, robotics	Domains: image & speech recognition, anomaly detection, recommendation systems, NLP	Technical operations: pharma development, process monitoring/ optimization/control, QC, environmental

AI, ML, & Chemometrics: Methods & Algorithms



Supervised
: Trained using labeled data

Unsupervised
: No training data set

Reinforcement
: Learn from feedback (robotics)

AI, ML, & Chemometrics: Differences in Approach

(to model creation and selection)

AI/ML Approach

- Starts with expertise & knowledge of many AI/ML methods
- Experiment with many methods to find various solutions - predictability
- Ranking based on best general performance (agreement with relevant scientific principles and algorithm simplicity have lower/no priority)*

Chemometrics Approach

- Starts with contextual **scientific expertise *plus* ML skills**
- Select from specific ML methods that reflect or correlate with relevant scientific principles
- Ranking based on optimal mix of *explainability*, algorithm *simplicity*, and *performance*

The power to explain has always been a key goal of chemometrics!

Visualization, diagnostics and 21 CFR part 11 compliance, Audit trail… ☑ ☑ ☑ ☑

* Based on presenter's experience

What is Multivariate Data Analysis (MVDA)?

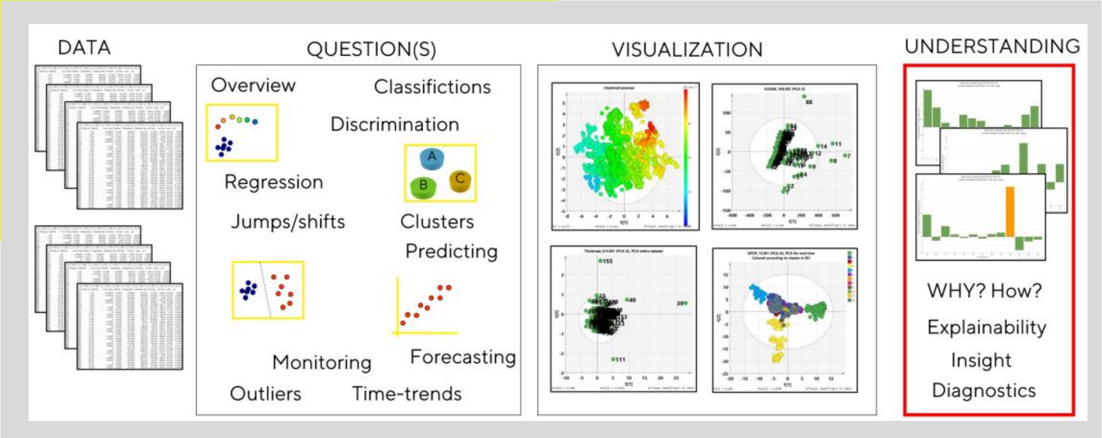
Depends Who You Ask…

MVDA In General

- Use of statistical and ML approaches to analyze data sets containing multiple variables

MVDA In Chemometrics (Historically)

- *Subset of time-tested, supervised and unsupervised ML methods*
- Usually refers to "factor-based" techniques, including PCA, PLS, OPLS, PLS-DA
  - Factorization of a data matrix into lower-dimensional representations called factors or latent variables
- Depending on where you are from (in the world)….
  - MVDA & Chemometrics → terms often used interchangeably
  - MVDA includes Design of Experiments (DOE)



What is Model Explainability?

Why is it important for biomanufacturing?

Definition

- Ability to explain and describe to a human why and how an AI/ML model made a prediction or decision
- (a.k.a.) interpretability or transparency

General Rule

- Use cases with higher risk or potential harm → demand more explainability!

Path to Meeting Regulatory Requirements

- Process Control – CPPs and CQAs
- Process Optimization – high yield, consistent quality
- Quality – disposition/release product
- Root Cause Investigations – "cause-and-effect" relationships
- Validation – traceability, auditability

SARTORIUS

Explainability & MVDA

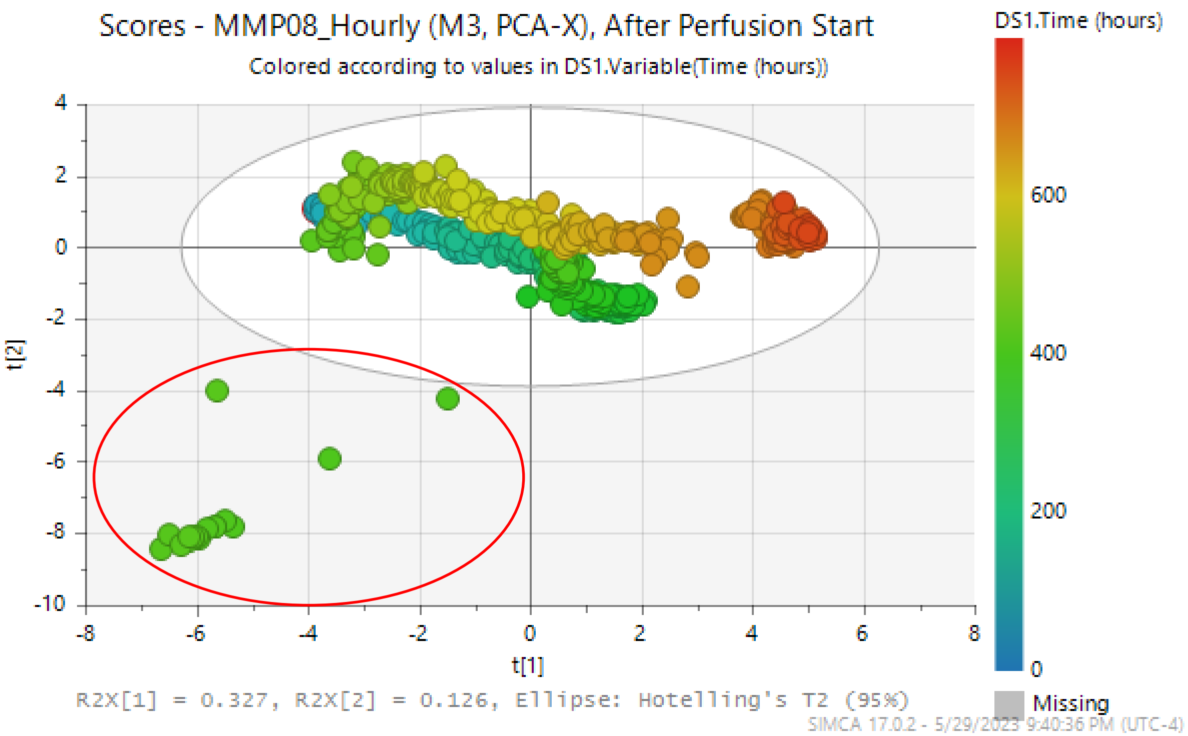
MVDA Methods Offer High Explainabilit☑

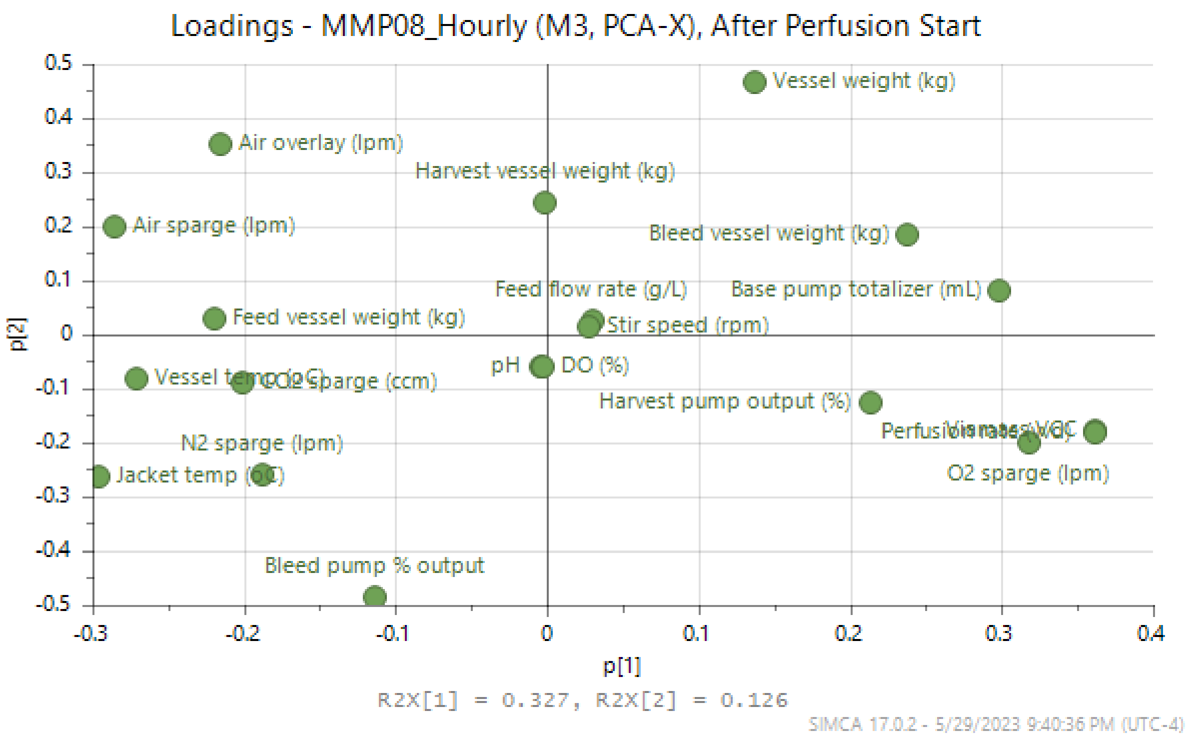
- ✓ Reduce dimensionality – focus on influential sources of variance
- ✓ Provide visualization
- ✓ Simplify interpretation
- ✓ Reveal underlying patterns/relationships
- ✓ Address multicollinearity
- ✓ Accommodate sparse data or missing data

Example

PCA Model of Perfusion Process Data

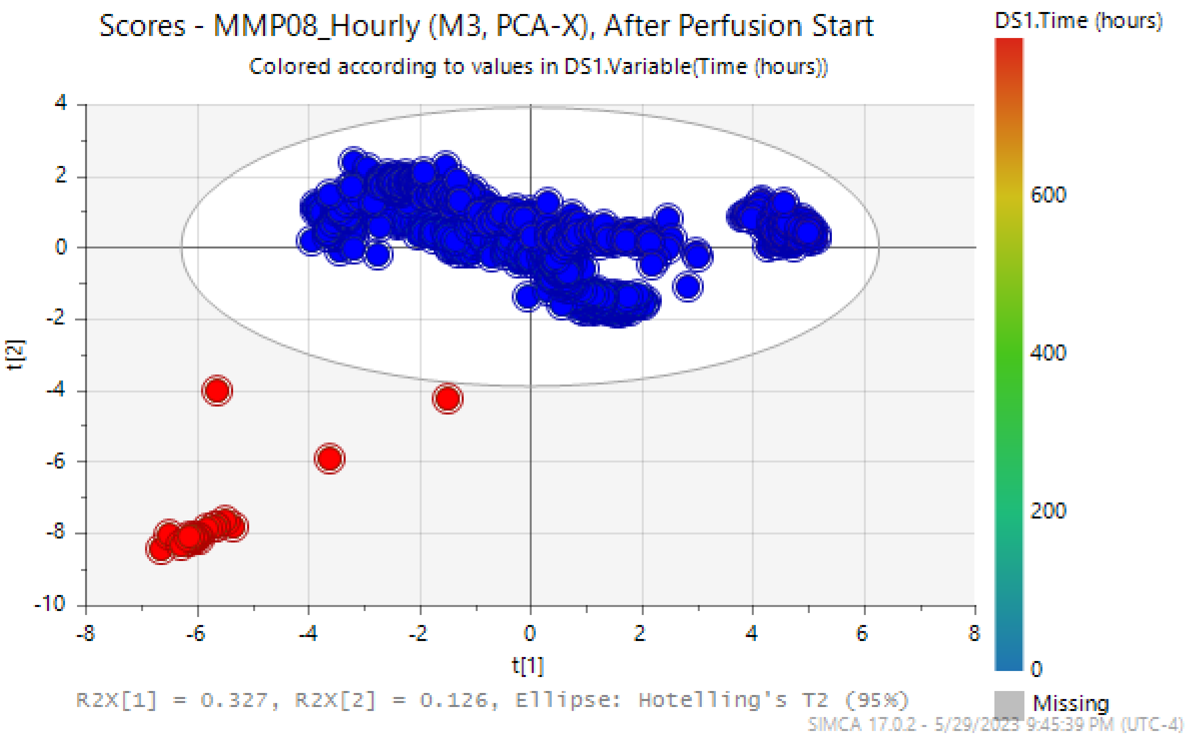
After Start of Perfusion

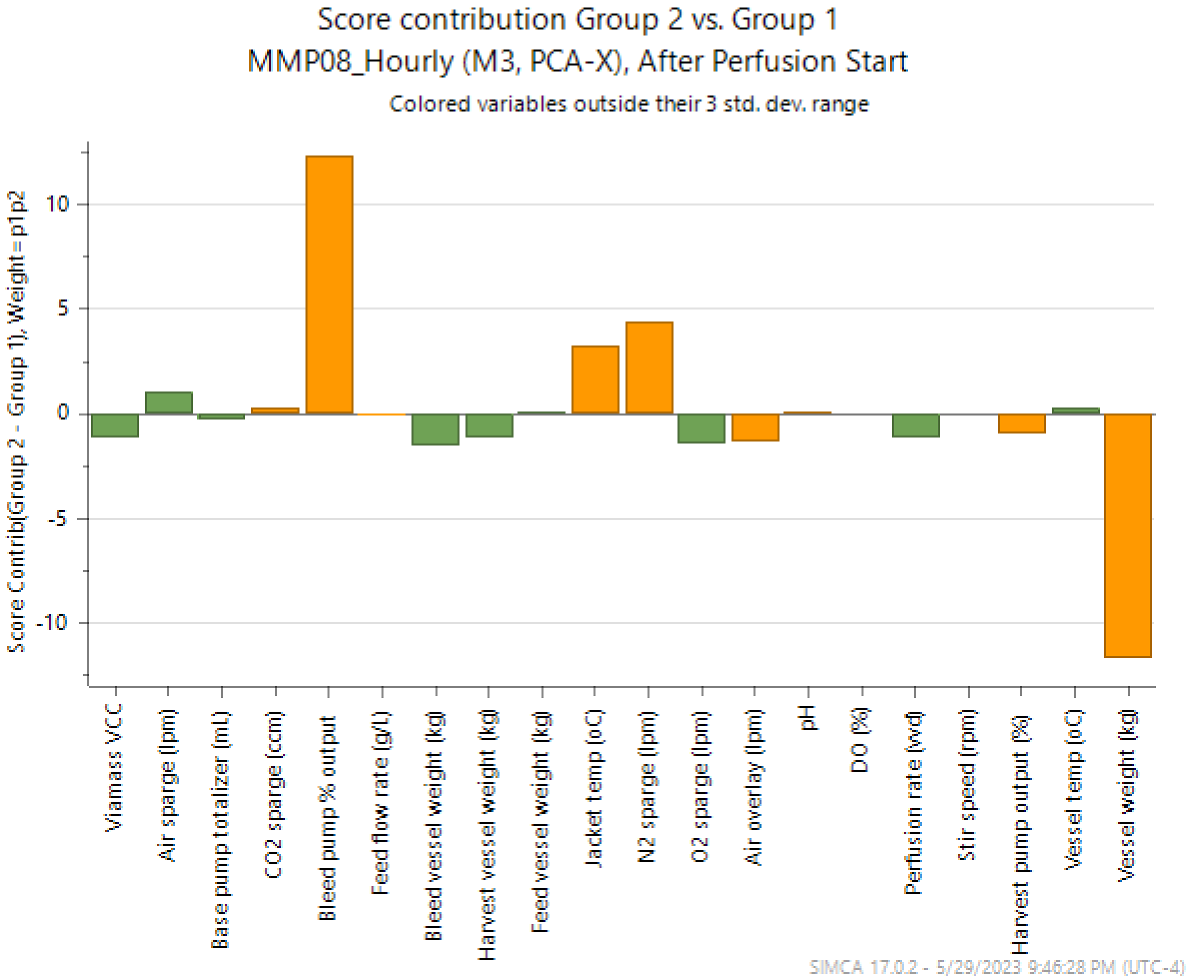




Scores plot identifies a number of observations outside of 95% Hotelling's T^2

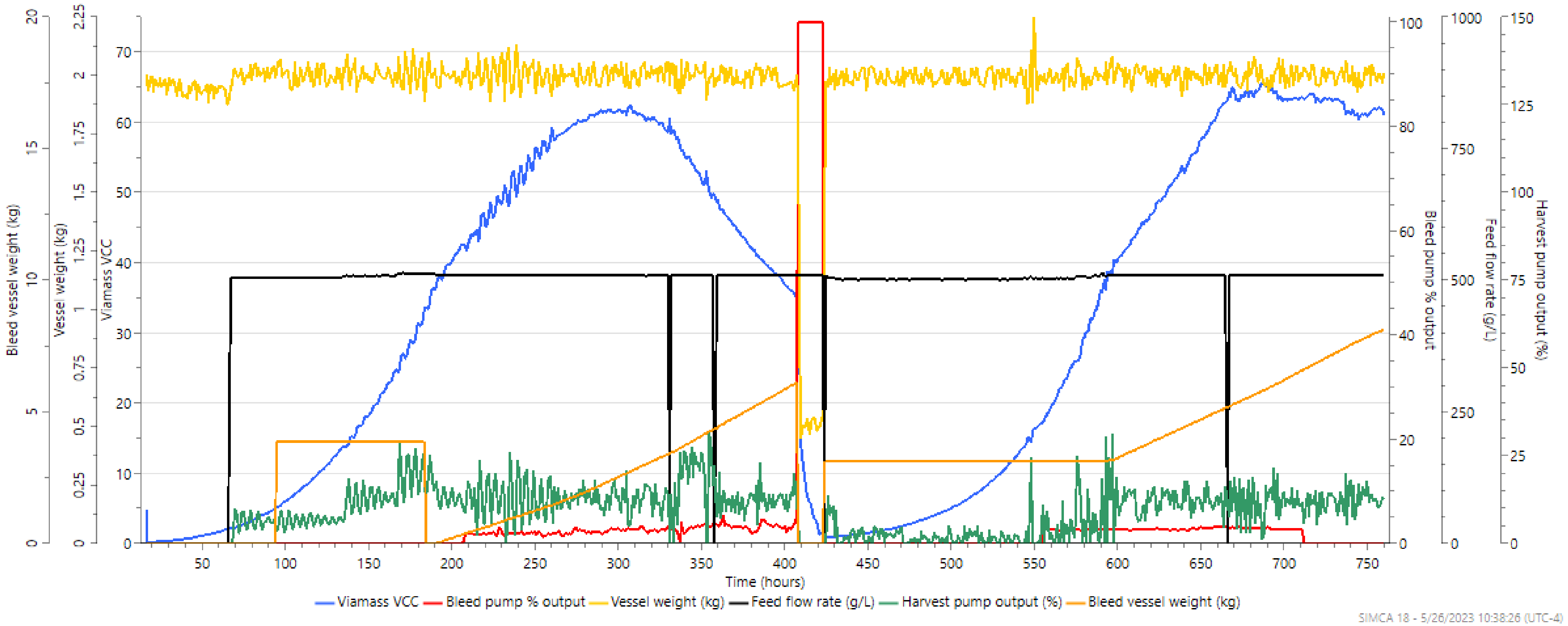
Group-to-Group Score Contribution Comparison





Group-to-group comparison: Group 1 (blue) to Group 2 (red)

The Big Picture of What Happened…

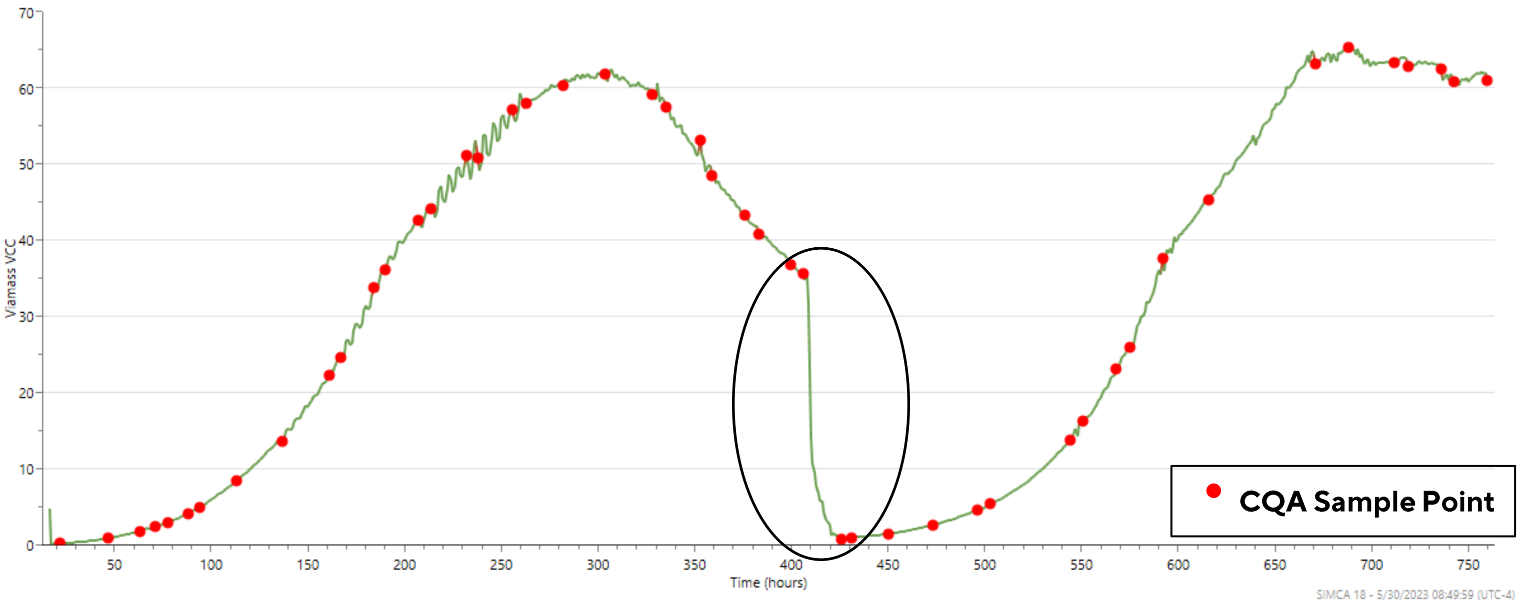


SARTORIUS

PCA Provides Explainability and Opportunity

Key Points

- *CQA sampling completely missed the event!*
  - CQA samples at 406 and 426 hours
- Event first detected and confirmed by PCA at 408 hours (18 hours earlier)
- PCA performed in ≤5 minutes!
- Monitoring process data could have saved time and material loss
  - Took 7.3 days to re-achieve optimal perfusion



Future Opportunity

→ Run PCA model in real-time on incoming data!

The adverse event could have been detected 18 hours sooner, allowing an immediate response!

"**At Takeda** a six month drug product particle issue investigation could have been avoided if they used data analytics from the start."

Challenge

Use SIMCA® to Analyze Final Drug Product Quality to Help Ensure Product Safety and Stability

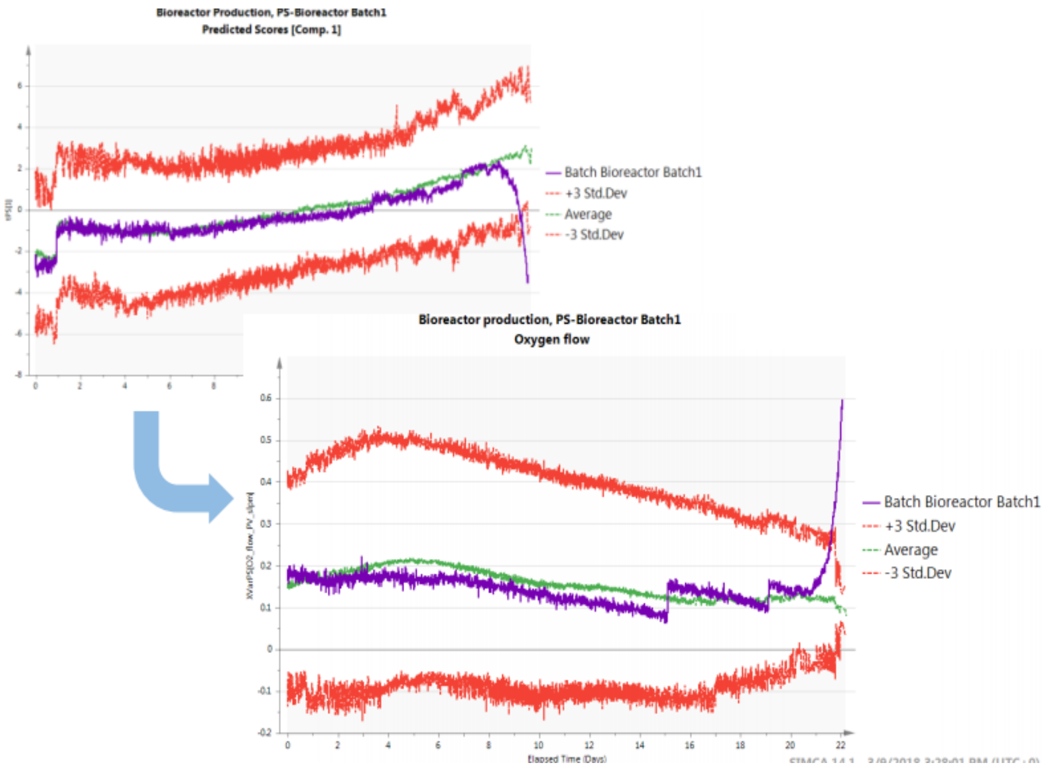
- Variation in volume and quality of final product directly influences commercial success
- Drug product investigations are time consuming and require QA | QC teams to review all USP and DSP data

SARTORIUS

"**At Takeda** a six month drug product particle issue investigation could have been avoided if they used data analytics from the start."

Solutions

Use SIMCA® MVDA to accelerate batch genealogy (tracking and tracing) studies



SARTORIUS

"**At Takeda** a six month drug product particle issue investigation could have been avoided if they used data analytics from the start."

Benefits

Accelerate quality investigations – *use AI/ML*

Enhance process traceability – *select the right AI/ML algorithms*

Ensure product safety – focus on *explainability and traceability not only predictions*

Takeaway Messages

- If you are doing ML, you are also doing AI
- If you are practicing chemometrics, you are also doing ML and AI

- It is important to choose models that correlate with relevant scientific principles and make common sense
- In general, use cases with higher risk or potential harm demand a higher degree of model explainability
- If your best algorithm has acceptable explainability, simplicity, and performance, your regulatory experience will be happier ☺

- Explainability of chemometrics/MVDA methods provides a commonsense path for meeting regulatory requirements in biopharma manufacturing

Thank you!

With contributions from Scott Carpenter, Timo Schmidberger, Andrew Hines

SARTORIUS

This presentation may include predictions, estimates, intentions, beliefs and other statements that are or may be construed as being forward-looking. While these forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could result in actual outcomes differing materially from those projected in these statements. No statement contained herein constitutes a commitment by AVEVA to perform any particular action or to deliver any particular product or product features. Readers are cautioned not to place undue reliance on these forward-looking statements, which reflect our opinions only as of the date of this presentation.

The Company shall not be obliged to disclose any revision to these forward-looking statements to reflect events or circumstances occurring after the date on which they are made or to reflect the occurrence of future events.

linkedin.com/company/aveva

@avevagroup

ABOUT AVEVA

AVEVA is a world leader in industrial software, providing engineering and operational solutions across multiple industries, including oil and gas, chemical, pharmaceutical, power and utilities, marine, renewables, and food and beverage. Our agnostic and open architecture helps organizations design, build, operate, maintain and optimize the complete lifecycle of complex industrial assets, from production plants and offshore platforms to manufactured consumer goods.

Over 20,000 enterprises in over 100 countries rely on AVEVA to help them deliver life's essentials: safe and reliable energy, food, medicines, infrastructure and more. By connecting people with trusted information and AI-enriched insights, AVEVA enables teams to engineer efficiently and optimize operations, driving growth and sustainability.

Named as one of the world's most innovative companies, AVEVA supports customers with open solutions and the expertise of more than 6,400 employees, 5,000 partners and 5,700 certified developers. The company is headquartered in Cambridge, UK.

Learn more at www.aveva.com

AVEVA