



12TH STATISTICS WORKSHOP: FROM DATA TO PATIENTS – CMC STATISTICIANS CONTRIBUTIONS TO QUALITY, SAFETY, AND EFFICACY OF PHARMACEUTICAL PRODUCTS

Beyond Control Charts: Adaptive, Model-Based Condition Monitoring with Generative AI for Biomanufacturing

Mark Wolff, M.S., Ph.D.

mark.wolff@sas.com

*Advisory Industry Consultant
SAS Institute*

Visiting Fellow
Frost Institute for Data Science and Computing
University of Miami



Biomanufacturing Value Chain

Biologics, combination drug-device, advanced therapies



Artificial Intelligence

Systems that **generate, reason, adapt, and act** to achieve goals once requiring human expertise.

Core Capabilities

- Generation – produces novel text, images, code, and structured outputs from natural-language intent
- Reasoning & Inference – chains evidence across context, retrieval, and tools to reach grounded conclusions
- Adaptation – a single foundation model handles many tasks; specialization comes from prompting, retrieval, and fine-tuning rather than retraining
- Agency – plans, calls tools, and executes multi-step tasks across systems with limited human oversight

Key Building Blocks

- Foundation Models – pretrained at scale on broad multimodal data; emergent capability from parameters and data, not hand-coded rules
- Retrieval-Augmented Generation – grounds reasoning in current, institution-specific knowledge bases rather than training data alone
- Tool Use & Orchestration – function calling and agentic frameworks let models act on systems, not just describe them
- Governance & Evaluation – continuous monitoring, citation, evals, and human-in-the-loop review as first-class engineering concerns

Spectrum of AI Methods



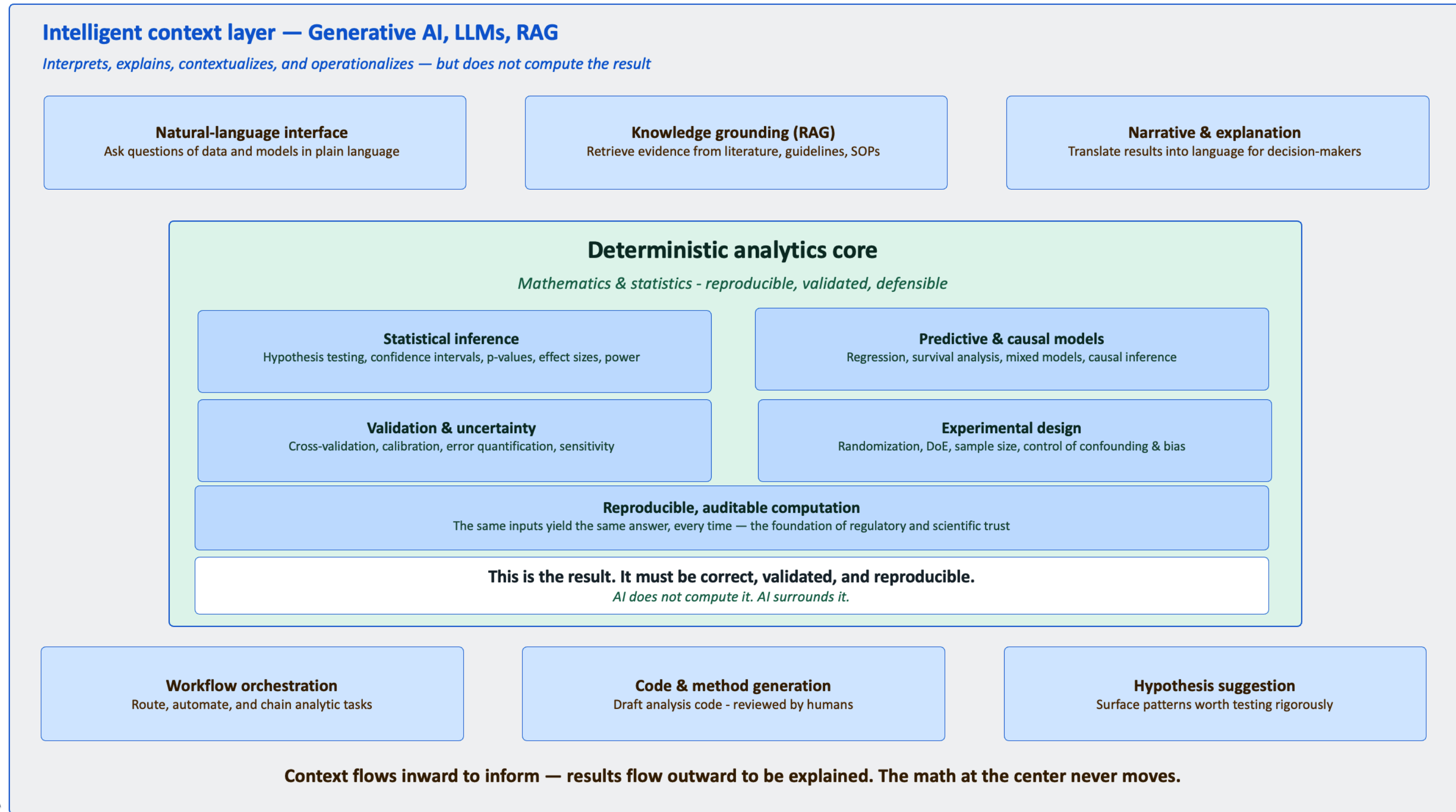
Why It Matters in Biomanufacturing

AI now reads the process, equipment, and batch record that holds most of a plant's knowledge, converting sensor trends and deviations into early warnings, evidence-grounded root causes, and governed actions inside the proven acceptable range.

"Reasoning" here means structured inference over evidence, NOT human understanding.

Intelligent Context

“Hybrid Analytics”



Deterministic vs. Probabilistic

Why it Matters

Deterministic analytics (the core):

- Same inputs always produce the same output.
- Results carry quantified uncertainty, confidence intervals, p-values, error bounds.
- Methods are validated, auditable, and reproducible, the basis of regulatory submission.
- Answers the question: what is true, and how confident can we be?

Probabilistic generative AI (the context):

- Same prompt can produce different outputs, inherently stochastic.
- Fluent and plausible, but can be confidently wrong (hallucination)
- Excellent at language, retrieval, summarization, and orchestration.
- Answers the question: how do we interpret, explain, and act on what is true?

The danger is mistaking fluency for correctness. A well-written wrong answer is still wrong.

What AI Adds

What it Should Not Replace

What intelligent context genuinely adds:

- Natural-language access to data and models for non-statisticians.
- Grounding answers in the literature and guidelines via RAG, with citations.
- Translating rigorous results into narrative for decision-makers.
- Accelerating the analyst, drafting code, surfacing hypotheses, automating workflow.

What it must never replace:

- The statistical estimate itself, with its uncertainty.
 - The validation that the method is correct for the data and question.
 - The human judgment that a result is scientifically and clinically sound.
-
- Every AI-surfaced insight must be confirmed by deterministic, validated analytics before it informs a decision.
 - Trustworthy AI in life sciences means AI that knows the boundary between context and computation, and stays on its side of it.

Manufacturing and Biomanufacturing

Innovation and Transformation

Innovation, what is newly possible

- Multivariate process models and PAT enable real-time release in place of end-of-batch testing.
- Digital twins of bioreactors predict yield and quality from live process data.
- AI-assisted visual inspection detects particles and defects beyond human reliability.

Transformation, how the work changes

- Quality shifts from reactive inspection to predictive assurance, built in, not tested in.
- Yield and titer optimization at scale directly lowers cost of goods for biologics.
- Gene and cell therapy manufacturing, the hardest to scale, becomes more predictable and economical.

The deterministic anchor: process control limits, capability indices, and release criteria remain statistically defined and validated. AI informs. Statistics certifies.

Biomanufacturing generates two kinds of data

The data are high-dimensional, autocorrelated, and unevenly sampled.

High-frequency process data

- pH, dissolved O₂, temperature, pCO₂
- Agitation and gas flows
- PAT probes: glucose, lactate, viable cells
- Seconds to minutes

Intermittent analytical data

- Titer, viability, metabolites, osmolality
- Glycosylation, aggregates, charge variants
- Hours to days
- Some results arrive after harvest

Classical SPC - Fixed limits and stable conditions

Shewhart limits sit at the mean $\pm 3\sigma$.

Assumptions

- Control limits are fixed from a stable reference period.
- Observations are assumed independent.
- The process is assumed stationary and predictable.
- Univariate charts watch one variable at a time.

1 in 370

Under normality and independence, a point falls outside $\pm 3\sigma$ with probability 0.0027. A false alarm occurs about once in 370 points.

Montgomery DC. *Introduction to Statistical Quality Control*. 8th ed. Wiley; 2019, Ch. 5.

Biological production systems violate these assumptions

Fixed limits flag normal trajectories as faults and may not detect real departures.

SPC assumption → reality in cell culture

Fixed limits

Variables trend with culture age and phase

Independence

Successive measurements are autocorrelated

Stationarity

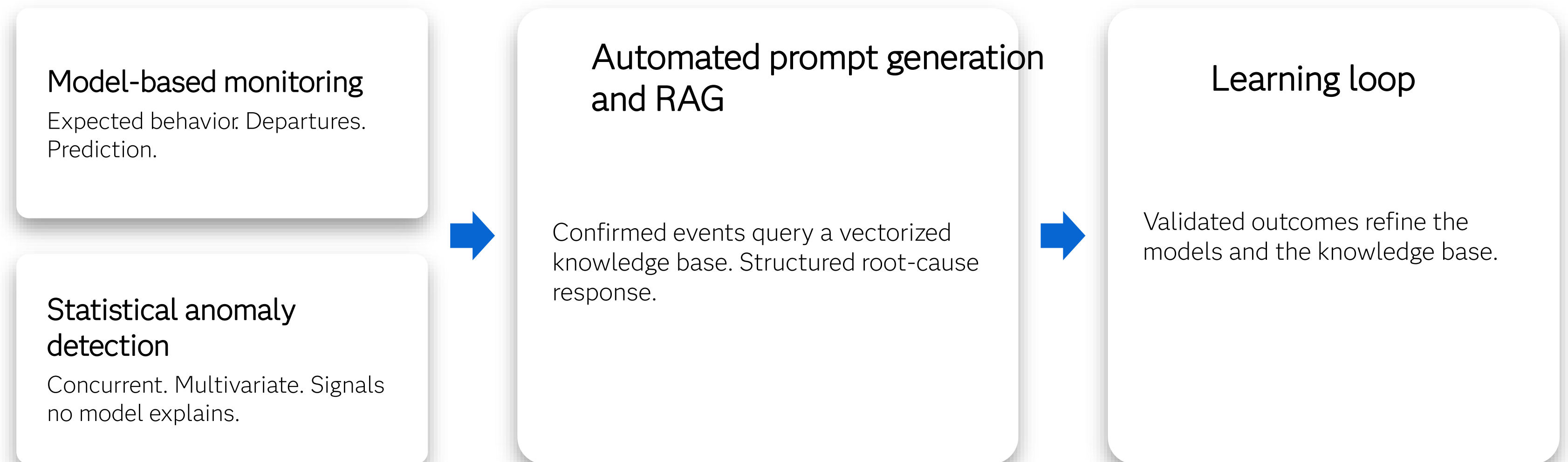
Cells, raw materials, and operating conditions vary

One variable at a time

Variables couple through metabolism and gas transfer

An intelligent condition monitoring framework

The framework extends classical SPC. It does not replace it.



Model-based monitoring represents expected behavior

Monitoring targets the departure from expectation, not the raw value.

Empirical

Form Batch-evolution PLS on time-aligned trajectories

Strength Learns from reference batches

Mechanistic

Form Growth, uptake, and production kinetics

Strength Extrapolates with physical meaning

Hybrid

Form Mechanistic structure with empirical correction

Strength Combines both

Model-based limits adapt where fixed limits cannot

Limits follow the expected trajectory at each culture time.

$$r(t) = y(t) - \hat{y}(t \mid \text{age, setpoints})$$

$$z(t) = r(t) / \sigma(t)$$

$$\text{Limits: } \hat{y}(t) \pm 3\sigma(t)$$

- Residuals are closer to stationary than raw values.
- Departures are evaluated in real time.
- Normal growth and scheduled shifts no longer raise alarms.

Prediction detects drift early and guides adjustment

Drift is flagged before it crosses a limit.

Predictive detection

- Soft sensors estimate titer and CQAs between samples
- Forecasts project trajectories to harvest
- Prediction intervals quantify uncertainty

Informed adjustment

- Candidate setpoint changes are evaluated by the model
- Constrained optimization selects the change
- Every change stays inside the design space
- Pre-approved changes follow the control strategy

Statistical anomaly detection runs concurrently

Purpose: identify signals outside the explanatory scope of existing process models.

Multiway PCA / PLS

Reference model from unfolded batch trajectories

Hotelling T^2

Unusual distance within the reference model

SPE (Q statistic)

A broken correlation structure, away from the model

MEWMA

Multivariate EWMA. Small, sustained shifts.

SVDD

Support vector data description. A nonparametric boundary.

Contribution plots

The variables that drive the signal

Examples: novel raw material lots, sensor faults, unseen metabolic states, gas transfer limits.

The two components classify every signal

Model status and anomaly status place each signal in one of four classes.

	No anomaly	Anomaly
Model deviation	Known deviation Pre-approved response inside the design space	Compound event Model and anomaly agree. Escalate.
No model deviation	Normal operation No action	Novel event No model explains the signal. Automated root-cause analysis.

Confirmation precedes escalation

Only confirmed events trigger the generative AI layer.

Confirmation

- Persistence: the signal holds for consecutive samples.
- Fusion: model residuals and multivariate statistics are read together.

Calibration

- Performance is stated as in-control and out-of-control average run length (ARL).
- Autocorrelation raises false alarm rates for run rules.
- Rules are calibrated on reference batches.

Automated prompt generation turns a signal into a query

No person writes the prompt.

- A confirmed event selects a template by event class.
- Slots fill from the monitored state.
- The prompt states role, context, signal, and task.

ROLE	Process deviation analyst
CONTEXT	Product, vessel, batch, day, phase
SIGNAL	Detector, statistic, persistence
STATE	Values, residuals, contributions
HISTORY	Prior events this batch
TASK	Explain, rank causes, recommend actions within the proven acceptable range

(12) United States Patent Wolff	(10) Patent No.: US 12,443,635 B1 (45) Date of Patent: Oct. 14, 2025
(54) SYSTEMS AND METHODS FOR AUTOMATED PROMPT GENERATION FOR INTELLIGENT ALERTING IN CONDITION MONITORING USING CONTEXTUAL LANGUAGE MODELS AND RETRIEVAL-AUGMENTED GENERATION	(56) References Cited U.S. PATENT DOCUMENTS 2021/0241884 A1* 8/2021 Swisher G06V 10/82 2024/0402662 A1* 12/2024 Ramanasankaran G05B 13/028 2025/0045256 A1* 2/2025 Gottlob G06F 16/211 2025/0111095 A1* 4/2025 Lecroart G06N 20/00 2025/0118066 A1* 4/2025 Hattori G06F 40/40 2025/0147992 A1* 5/2025 Ding G06F 16/345 2025/0194978 A1* 6/2025 Doyen G16H 50/20 2025/0202762 A1* 6/2025 Izquierdo Franco ... H04L 41/16 * cited by examiner <i>Primary Examiner</i> — Eric Yen (74) <i>Attorney, Agent, or Firm</i> — Padowitz Alce; Quintin Scheitlin; Alce PLLC (57) ABSTRACT A system, method, and computer-program product includes receiving a machine-generated alert indicating a data outlier for a physical asset; obtaining, from a computer database, a dataset including observed data for the physical asset within a predefined temporal window of the machine-generated alert; generating, via a description generator, a contextual description of the data outlier based at least on the observed data for the physical asset; searching an embedding-based knowledgebase for a subset of embedding representations of the plurality of embedding representations within a similar- ity threshold of an embedding representation of the contex- tual description; obtaining, from the computer database, a subset of the plurality of reference artifacts that correspond to the subset of embedding representations; and generating, via a large language model, a resolution suggestion for resolving the data outlier based at least on the subset of reference artifacts obtained from the computer database. 30 Claims, 23 Drawing Sheets
(71) Applicant: SAS Institute Inc. , Cary, NC (US)	
(72) Inventor: Mark A. Wolff , Cary, NC (US)	
(73) Assignee: SAS INSTITUTE INC. , Cary, NC (US)	
(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.	
(21) Appl. No.: 19/229,708	
(22) Filed: Jun. 5, 2025	
Related U.S. Application Data	
(60) Provisional application No. 63/704,226, filed on Oct. 7, 2024.	
(51) Int. Cl. G06F 16/3329 (2025.01) G06F 16/334 (2025.01)	
(52) U.S. Cl. CPC G06F 16/3329 (2019.01); G06F 16/334 (2019.01)	
(58) Field of Classification Search CPC G06F 16/3329; G06F 16/334 See application file for complete search history.	

A vectorized knowledge base supplies the evidence

Retrieval ranks passages by similarity to the prompt.

Process development records

Design space, ranges, characterization studies

Batch histories

Prior deviations and their responses

Regulatory guidance

Risk and compliance expectations

Curated scientific literature

Mechanisms and published evidence

Documents are chunked, embedded, versioned, and access-controlled.

RAG returns a structured root-cause response

The result is automated root-cause analysis.

Contextual explanation

What the deviation means in process terms

Ranked contributing factors

Weights, evidence for and against, verification steps

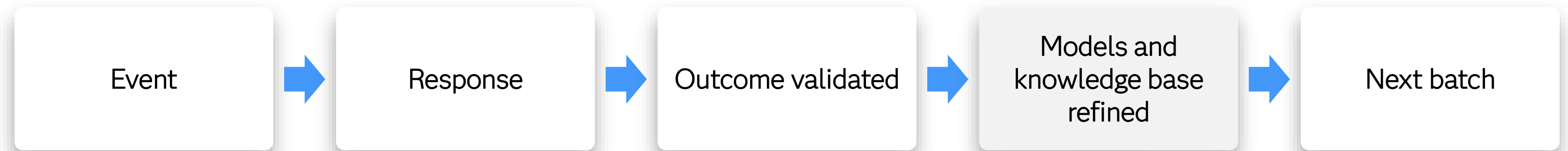
Recommended corrective actions

Bounded by the design space, with an approval tier

- The model answers only from retrieved sources and cites every claim.
- A fixed output schema and fixed generation settings constrain variation. Every response is logged.
- Unsupported claims and out-of-range actions are removed before display.

Validated outcomes create a learning system

Responses stay consistent across batches and campaigns.



- Validated cases enter the knowledge base.
- Confirmed findings update the ranking of causes.
- Recurring patterns become known deviation classes.
- Monitoring models update after revalidation under change control.

The framework is designed for regulatory expectations

Each expectation maps to a design feature and a reference.

Expectation	How the framework meets it	Reference
Process control	Actions bounded by the design space and control strategy	ICH Q8(R2), Q10
Risk management	Deviation severity assessed by risk	ICH Q9(R1)
Lifecycle change	Models and responses updated under change control	ICH Q12
Ongoing verification	Monitoring feeds continued process verification	FDA PV 2011, PAT 2004
Traceability	Prompt, sources, response, and approval logged	21 CFR Part 11
Explainability	Every claim cited. Statistics kept separate from text.	EU GMP Annex 22 (draft)
AI governance	Defined context of use and model risk	FDA AI draft guidance 2025

The role of the CMC statistician

The framework depends on statistical rigor.

Monitoring models need validation and a revalidation schedule.

The models themselves drift and need monitoring.

False alarm rates must account for autocorrelation.

Every prediction carries an interval, and actions respect it.

Statistical evidence stays separate from generated text.

Pre-approved automated responses need documented evidence.

Expected outcomes

Benefits for operations, quality, and investigations.

Earlier detection

Deviations and anomalies found before they cross a limit

Lower investigational burden

Root-cause analysis drafted with cited evidence

Less waste

Fewer lost batches and fewer false alarms

Higher yield and quality

Timely, consistent adjustment inside the design space

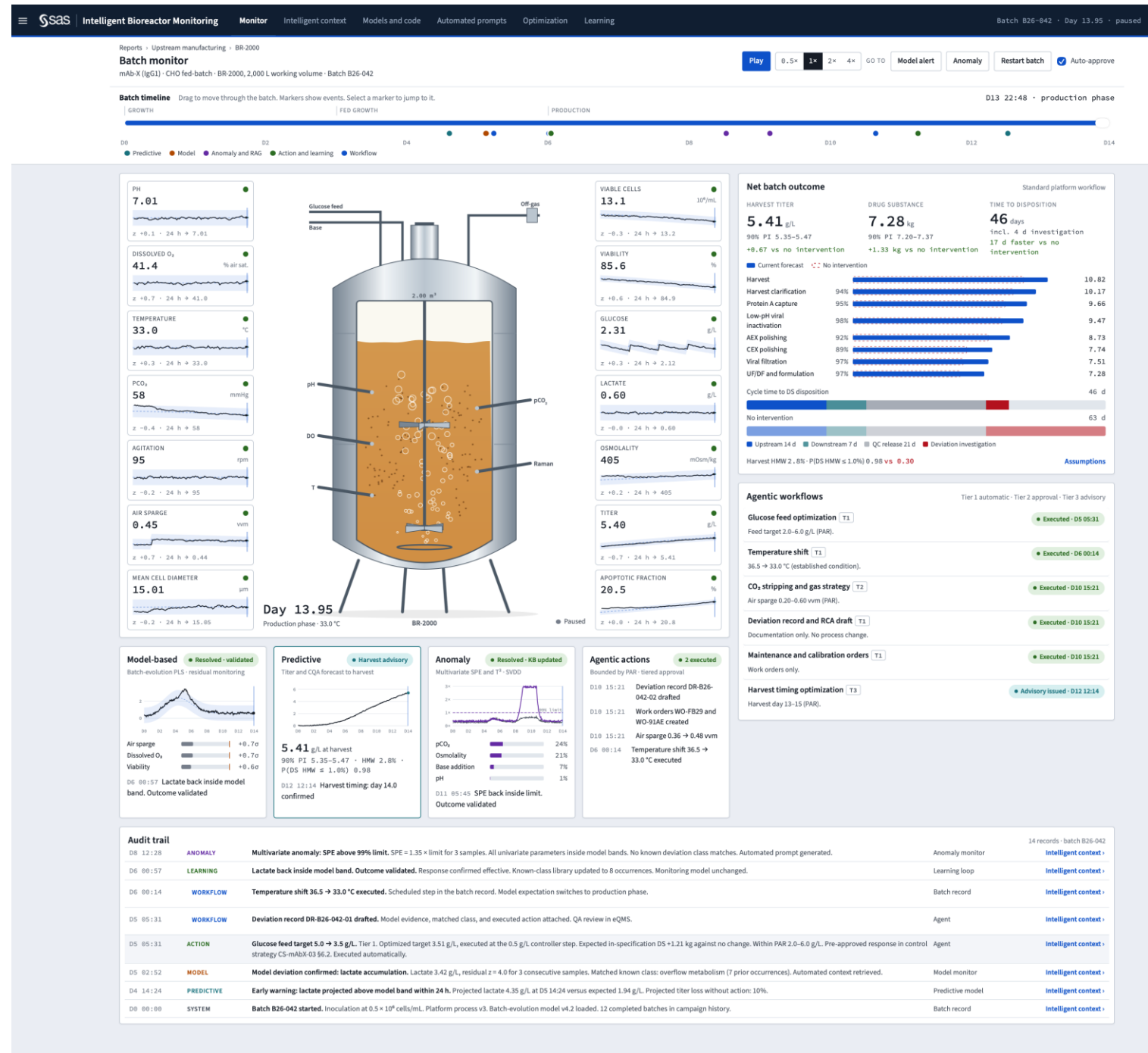
Summary

Fixed limits do not fit living production systems.

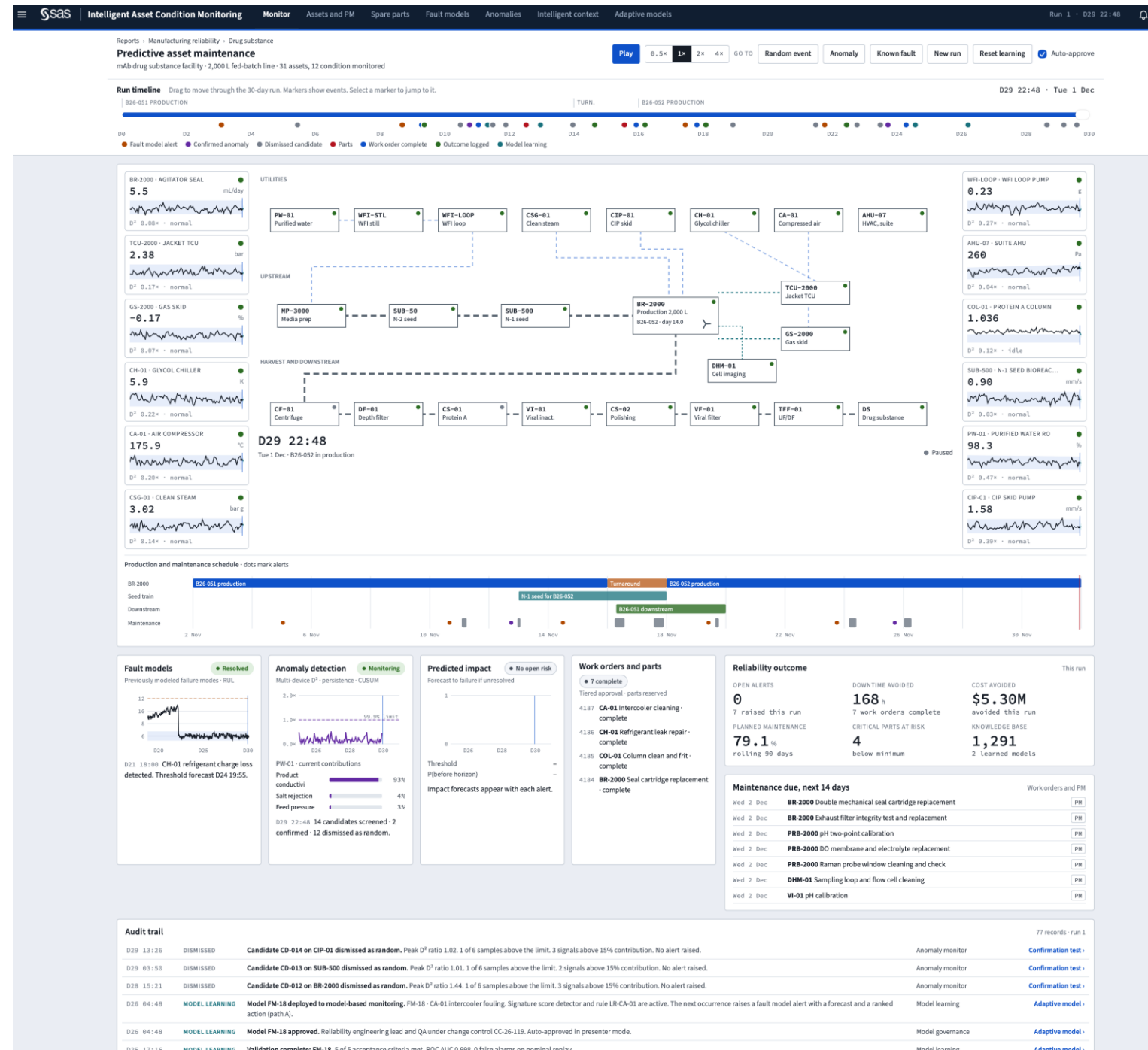
- Model-based monitoring follows the expected trajectory and predicts drift.
- Anomaly detection covers deviations for which no model exists.
- Automated prompts and RAG turn confirmed events into cited root-cause analysis.
- Validated outcomes make the system learn across batches and campaigns.
- Governance keeps the system controlled, traceable, and explainable.

DEMO

Intelligent Bioreactor Monitoring



Intelligent Asset Condition Monitoring





12TH STATISTICS WORKSHOP: FROM DATA TO PATIENTS – CMC STATISTICIANS CONTRIBUTIONS TO QUALITY, SAFETY, AND EFFICACY OF PHARMACEUTICAL PRODUCTS

Beyond Control Charts: Adaptive, Model-Based Condition Monitoring with Generative AI for Biomanufacturing

Mark Wolff, M.S., Ph.D.

mark.wolff@sas.com

*Advisory Industry Consultant
SAS Institute*

Visiting Fellow
Frost Institute for Data Science and Computing
University of Miami

