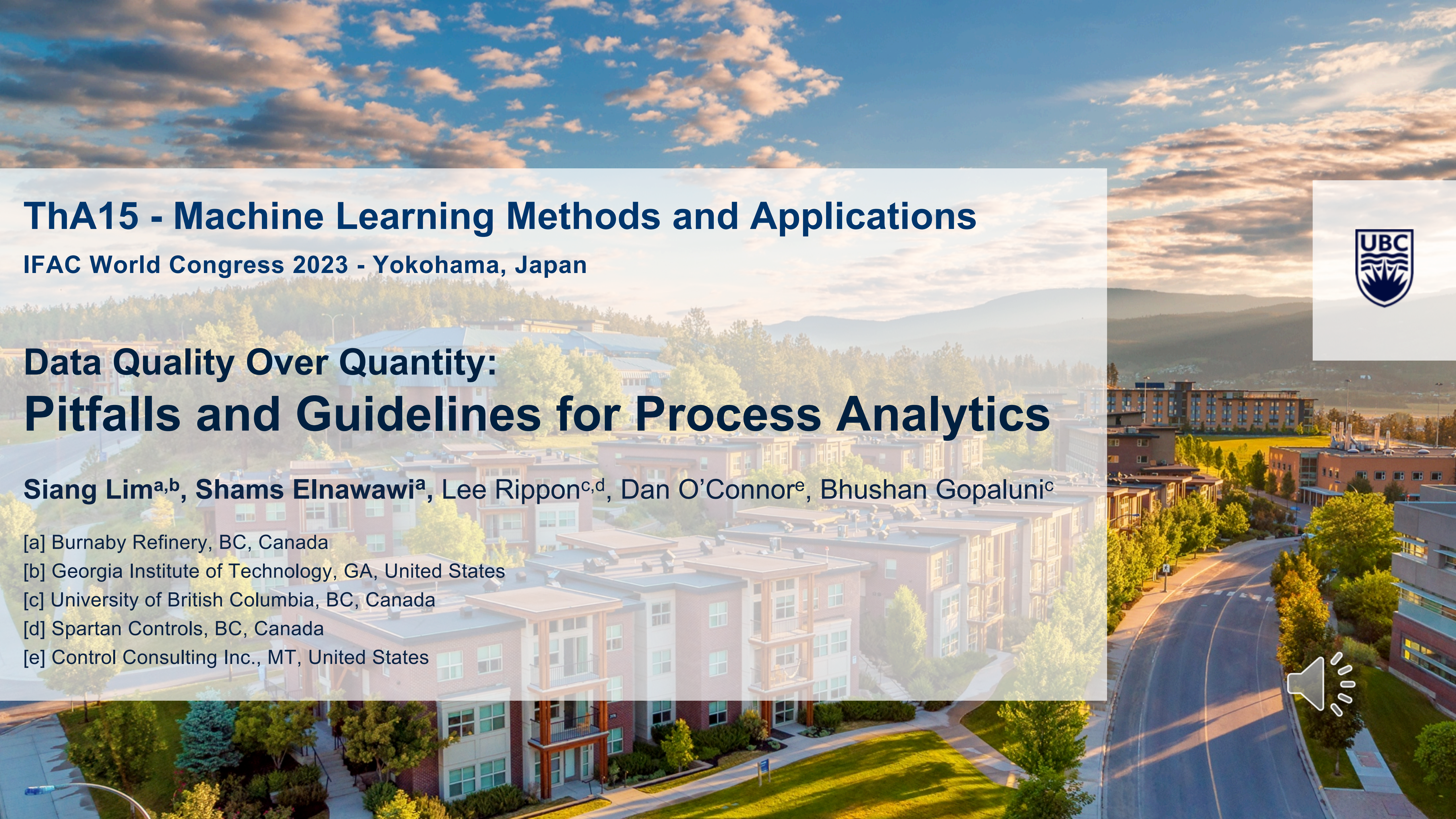


ThA15 - Machine Learning Methods and Applications

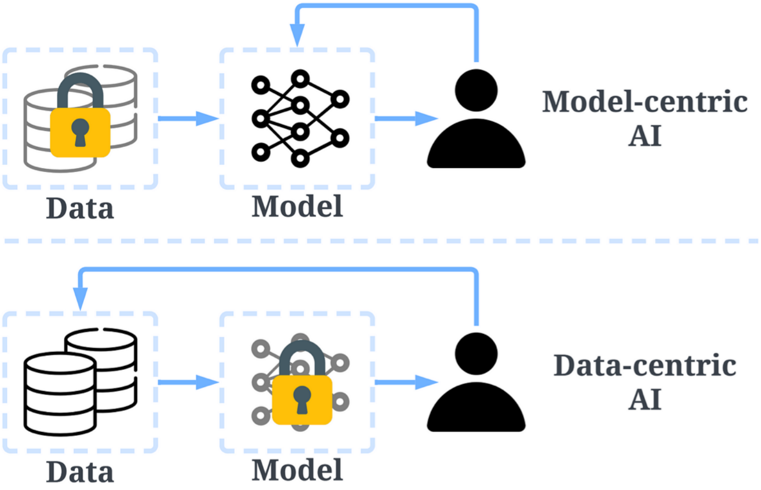
IFAC World Congress 2023 - Yokohama, Japan

Data Quality Over Quantity:
Pitfalls and Guidelines for Process Analytics

Siang Lim[a,b], Shams Elnawawi[a], Lee Rippon[c,d], Dan O'Connor[e], Bhushan Gopaluni[c]

[a] Burnaby Refinery, BC, Canada
[b] Georgia Institute of Technology, GA, United States
[c] University of British Columbia, BC, Canada
[d] Spartan Controls, BC, Canada
[e] Control Consulting Inc., MT, United States

Today’s Topic: **Pitfalls in Industrial Process Analytics**



Zha, Daochen, et al. **"Data-centric Artificial Intelligence: A Survey."** *arXiv preprint arXiv:2303.10158*, 2023.

https://github.com/daochenzha/data-centric-AI

Motivation:

Much of the process analytics literature focuses on modelling and algorithmic techniques, while little attention is paid to the practical aspects of data acquisition and cleaning.

Practitioners unfamiliar with industrial datasets often face difficulties in these areas due to the lack of practical academic resources.

Our Work:

A series of common pitfalls and practical considerations for obtaining and pre-processing data for process analytics applications, intended as a **tutorial** for data practitioners unfamiliar with industrial datasets.

Background

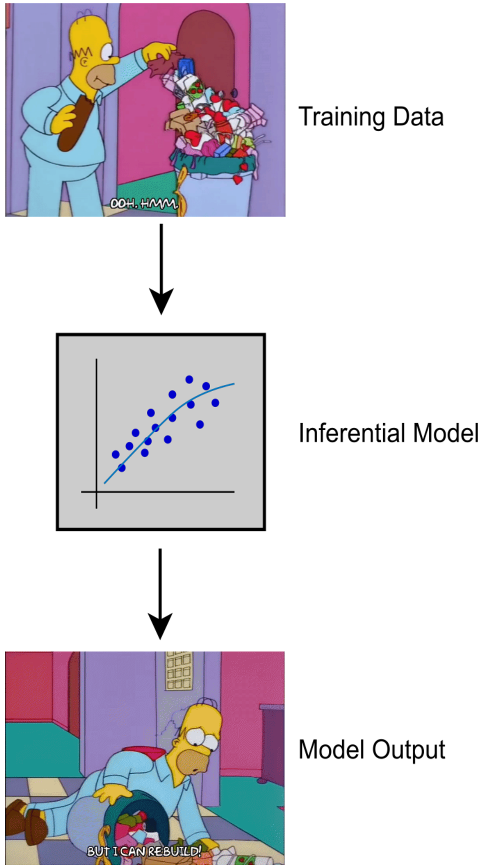


Figure 1: “Garbage in, garbage out” – a colloquialism describing that any advanced model can only be as good as the quality of data used to train it.

- Much of the research in process analytics under-emphasizes *how* finished datasets are obtained.
  - ○ Inferential models work by capturing patterns in the training data;
  - ○ If the data is not descriptive or “clean” enough to allow this, models will not learn well;
- Scope is focused on inferentials/soft sensors.

[1] - Sun, W. and Braatz, R.D. (2020). Opportunities in tensorial data analytics for chemical and biological manufacturing processes. *Comput. Chem. Eng.*, 143, 107099.

Pitfall 1 – Failure in data retrieval and contextualization

- “Process data” describe many aspects of operations in many different formats.
- Exploratory data analysis (EDA) is necessary for more concrete understanding.
  - Self-service analytics tools like Seeq and Spotfire facilitate EDA tasks.
- Practitioners can contextualize data by consulting plant personnel.
  - e.g. data retrieval settings, interpolation methods, tag calculations, etc.

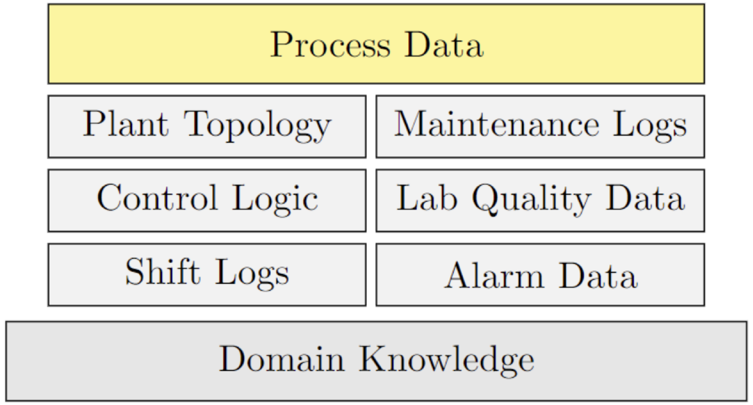
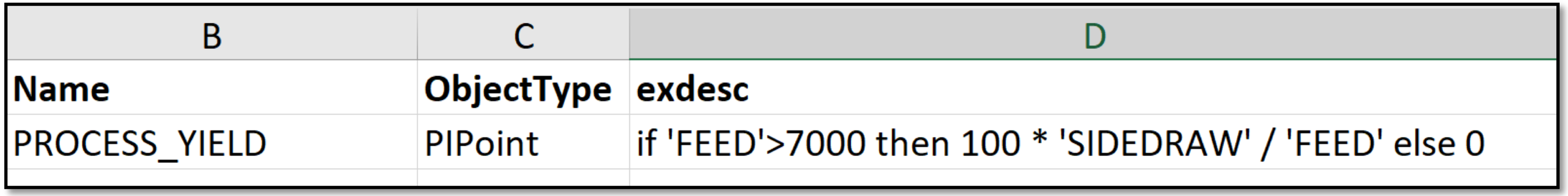


Figure 2: Industrial data come from many different sources in different formats, and these must be contextualized with other datasets to provide actionable information.

Example 1-1: Beware of 'hidden' calculations



B	C	D
Name	**ObjectType**	**exdesc**
PROCESS_YIELD	PIPoint	if 'FEED'>7000 then 100 * 'SIDEDRAW' / 'FEED' else 0

Figure 4: Tag names can be misleading because a process value could undergo transformations through PI calculations before being historized. In this PI tag example, the PROCESS_YIELD is zeroed when the FEED is below threshold, not because the yield is 0.

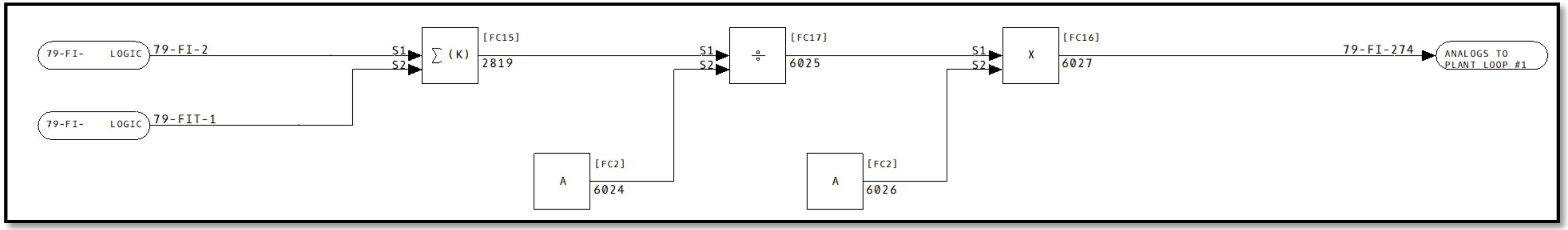


Figure 5: A process value could also undergo transformations in the DCS. In this example, the tag 79FR274 is not a value from a flow meter 79-FIT-274 as we would expect but is calculated based on 79-FIT-1 and 79-FIT-2.

Tags might not store raw values: watch out for calculations hidden in the data historian/DCS

Pitfall 2 – Ignoring domain knowledge and data cleaning

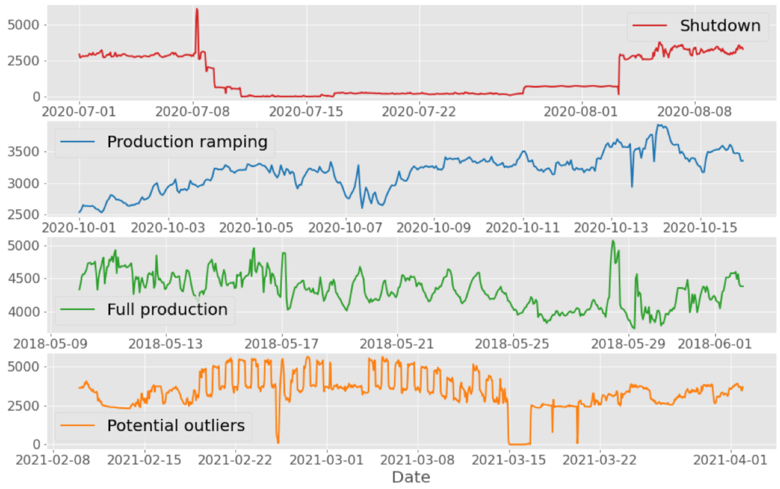


Figure 7a: Categories of operating data in one variable.

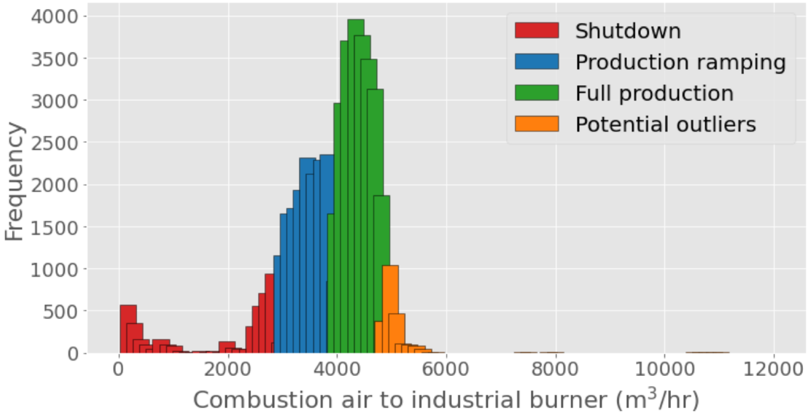


Figure 7b: Histogram of operating regimes from Fig. 8a.

- Continuous chemical processes operate at steady-state – they are “data-rich but information-poor”[2].
- **Accounting for domain knowledge:**
  - Selecting tags that are relevant to the task;
  - Distinguishing different operating regimes;
  - “Curse of dimensionality”: more features does not necessarily lead to better models[3];
- **Data cleaning:** suitable processing of outliers, missing data points, misaligned data.

[2] - Dong, D. and McAvoy, T. (1996). Nonlinear principal component analysis—based on principal curves and neural networks. *Comput. Chem. Eng.*

Example 2-1: Using process values to identify plant operating modes can be misleading during data cleaning

Problem:

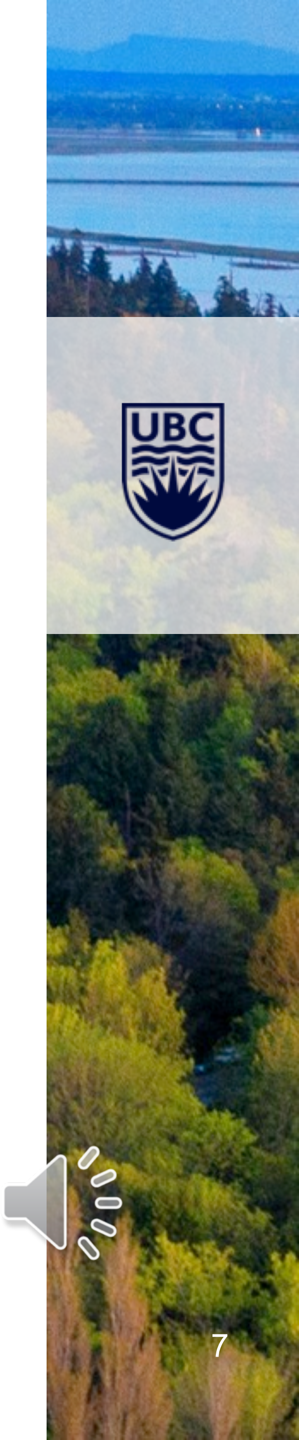
How do we determine if a plant was running during a certain time period using historical process data?

Misconception:

Just apply a threshold filter to the feed rate. If the feed rate was higher than a certain threshold, the plant was running.

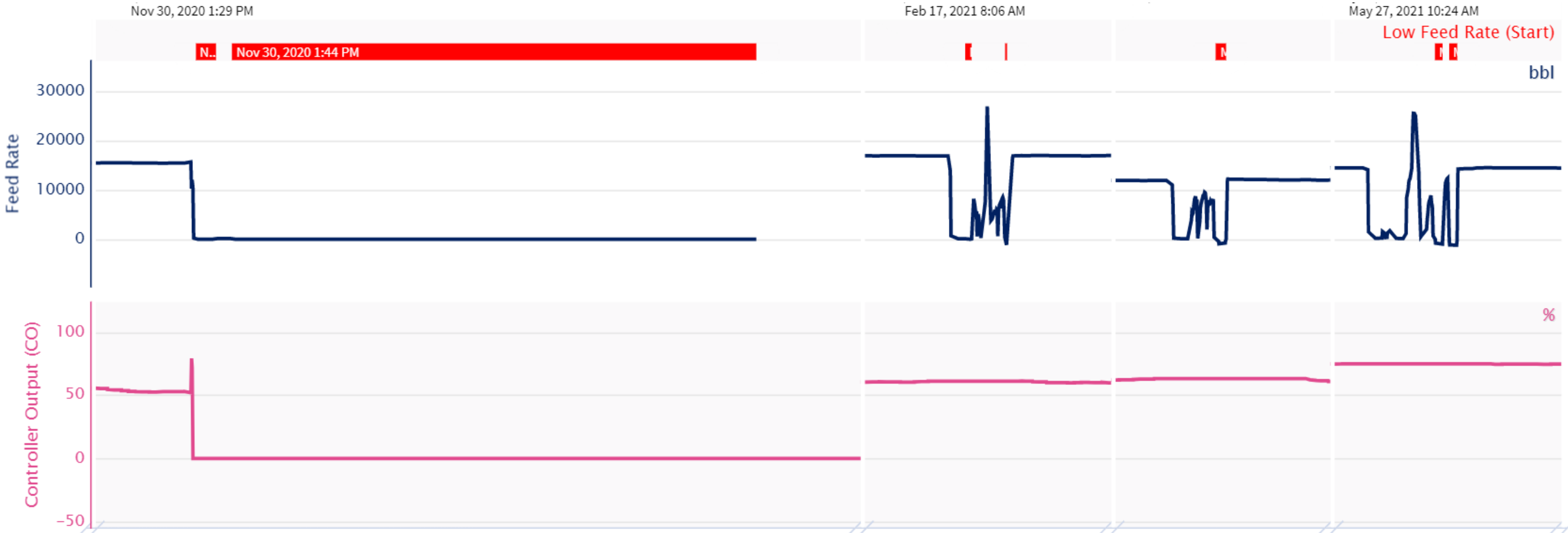
Reason:

Process values can be noisy. Unreliable instruments can fail intermittently to low values. Blindly thresholding the data could incorrectly remove regions when the plant was online.



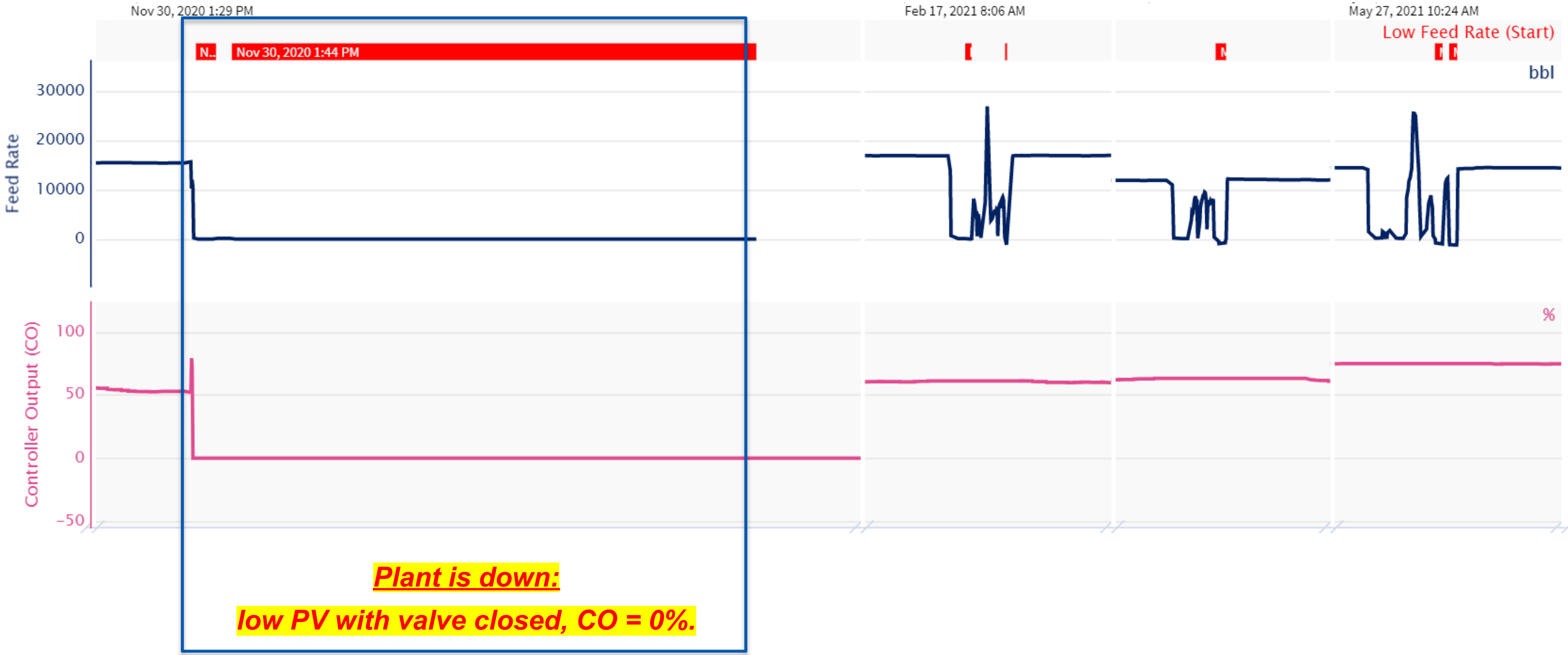
Example 2-1: Using process values to identify plant operating modes can be misleading during data cleaning

Figure 9: Time series of feed rate (PV, blue), feed rate controller valve output (CO, magenta), and regions with feed rate < 100 BPD (red)



Example 2-1: Using process values to identify plant operating modes can be misleading during data cleaning

Figure 9: Time series of feed rate (PV, blue), feed rate controller valve output (CO, magenta), and regions with feed rate < 100 BPD (red)

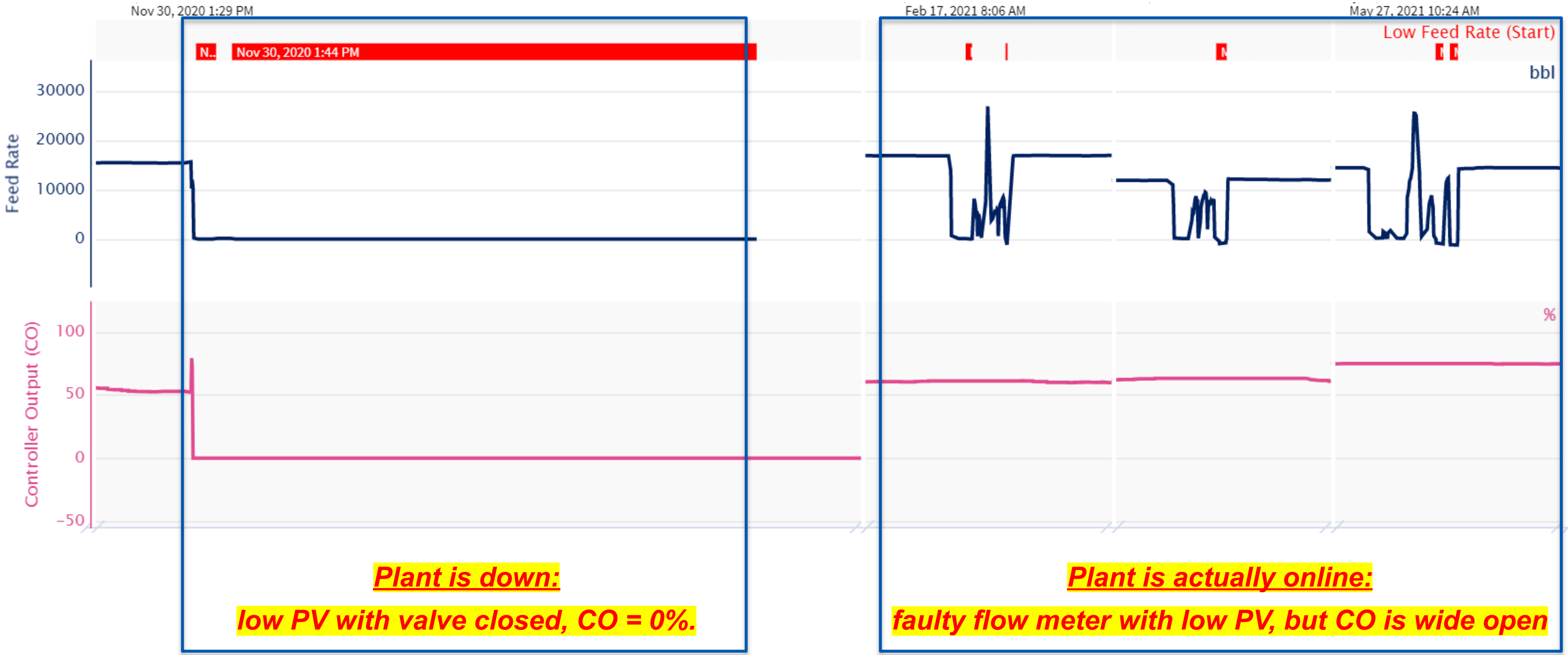


Plant is down:

low PV with valve closed, CO = 0%.

Example 2-1: Using process values to identify plant operating modes can be misleading during data cleaning

Figure 9: Time series of feed rate (PV, blue), feed rate controller valve output (CO, magenta), and regions with feed rate < 100 BPD (red)



Plant is down:

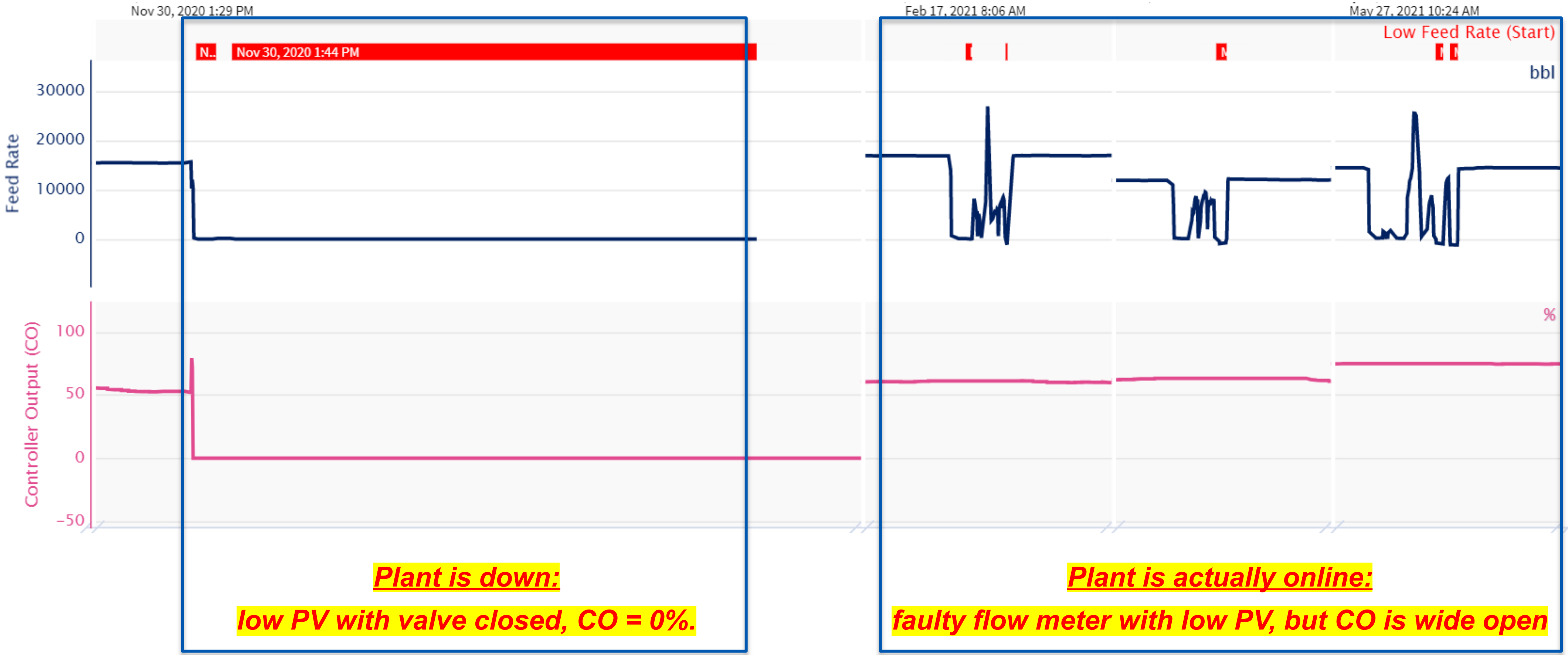
low PV with valve closed, CO = 0%.

Plant is actually online:

faulty flow meter with low PV, but CO is wide open

Example 2-1: Using process values to identify plant operating modes can be misleading during data cleaning

Figure 9: Time series of feed rate (PV, blue), feed rate controller valve output (CO, magenta), and regions with feed rate < 100 BPD (red)



Plant is down:

low PV with valve closed, CO = 0%.

Plant is actually online:

faulty flow meter with low PV, but CO is wide open

Context is important: a single tag (.PV) might not tell the whole story.

Example 2-2: Adding more features might not help you – building soft sensors for tracking ‘green’ molecules

Problem:

To generate carbon credits, refiners must quantify the amount of bio-based ‘green’ content in fuels. The standard method is to use ^{14}C measurements, which is expensive and time-consuming to do online. A soft sensor can help refiners estimate biogenic content in real-time.

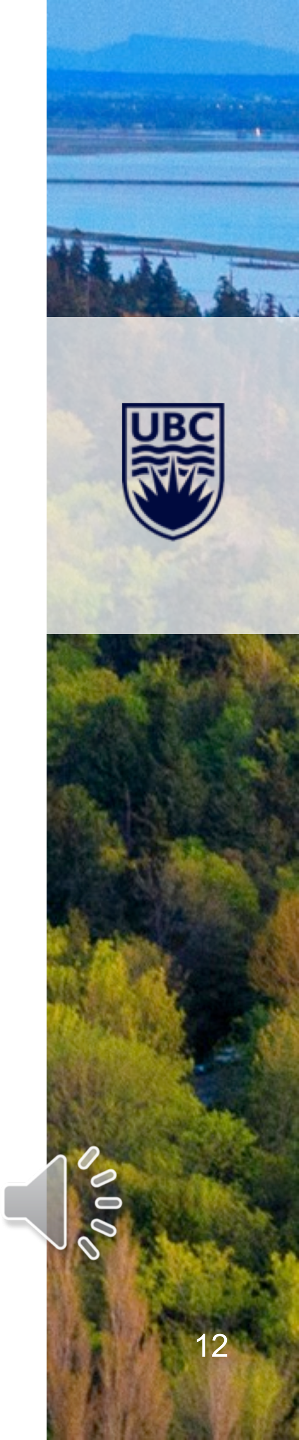
How do we determine which features to use in our soft sensor model?

Misconception:

Just use all tags in the plant, more features = more information. More is always better.

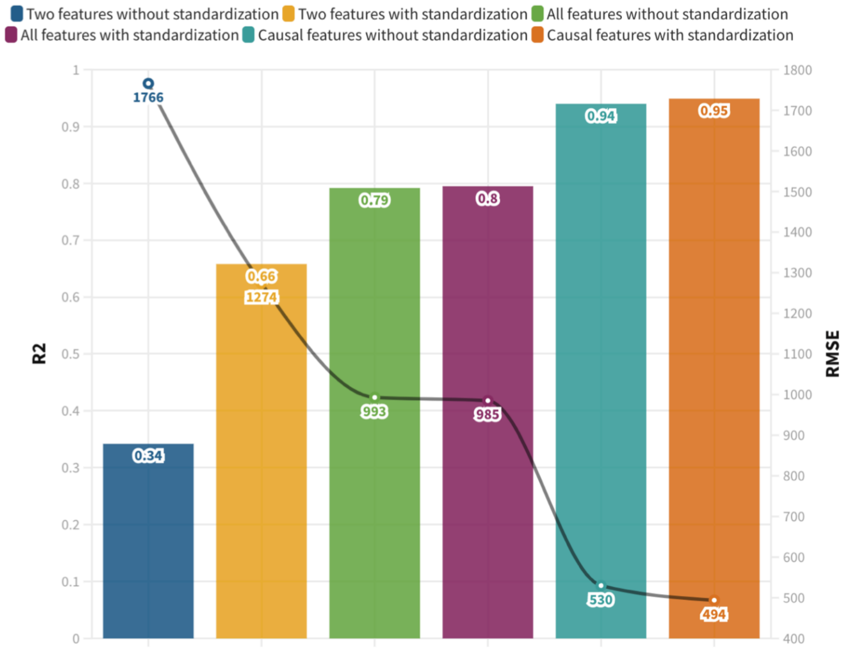
Reason:

Adding ‘non-informative’ or ‘non-useful’ features will confuse your models.



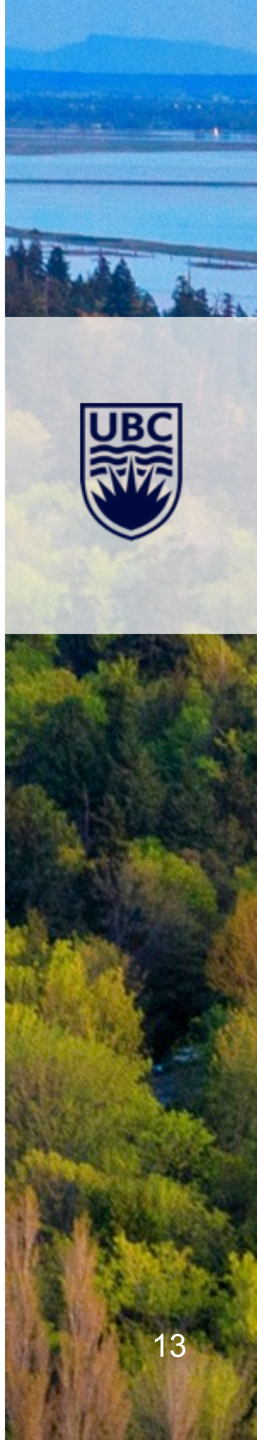
Example 2-2: Adding more features might not help you – building soft sensors for tracking ‘green’ molecules

Figure 10: Bar charts show soft sensor performance (R^2) using all features (green) compared to carefully selecting features using domain knowledge and causal analysis (orange).

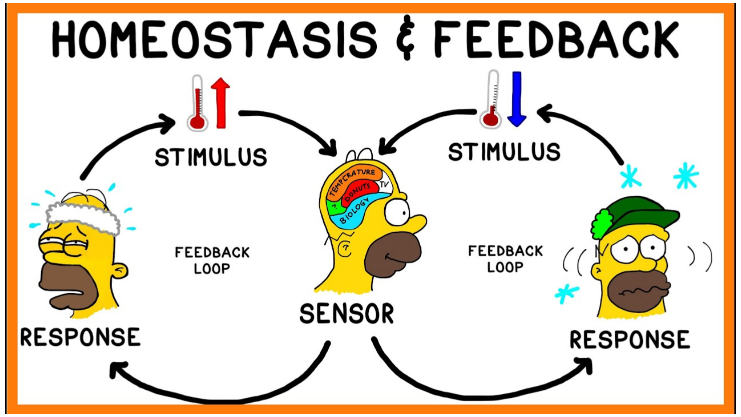


Reference:
Su, Jianping, et al. "Tracking the green coke production when co-processing lipids at a commercial fluid catalytic cracker (FCC): combining isotope ^{14}C and causal discovery analysis." *Sustainable Energy & Fuels* (2022)

- **Goal:** Predict biogenic feed content in fuels using historical process data.
- **Results:** Blindly using all features resulted in a mediocre model. Carefully selecting features using domain knowledge gave much better performance.
- Work led by UBC researchers and PhD students (Jianping Su and Liang Cao) in collaboration with the Burnaby Refinery



Pitfall 3 – Failure to account for closed-loop conditions



Reality:

Closed-loop feedback control is almost always present in industrial process datasets.

Implications:

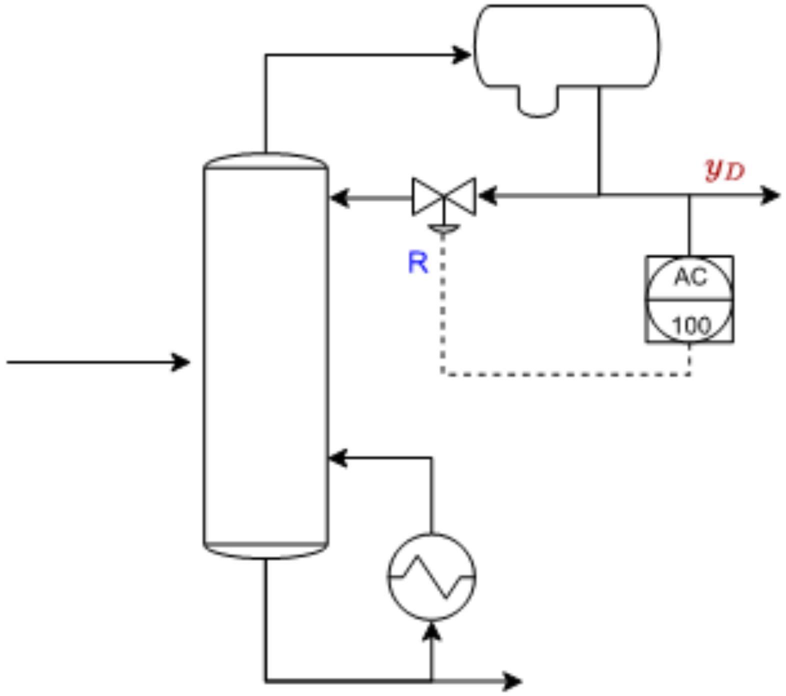
Feedback will influence process data and observed correlations, causing sign changes or spurious correlations.

Significance:

Indiscriminately using industrial process data without accounting for closed-loop conditions could result in models or correlations that do not align with reality or physics.

[BOGObiology]. (2016, December 3. 5 Minute Bio – Homeostasis. [Video File]. Retrieved from https://www.youtube.com/watch?v=kAy-03hIfck

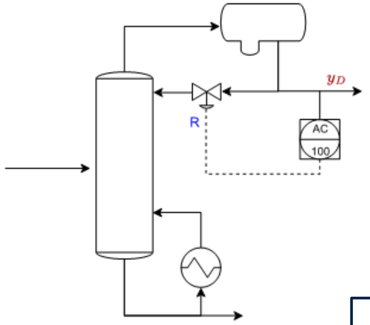
Example 3-0: Simple Distillation Column

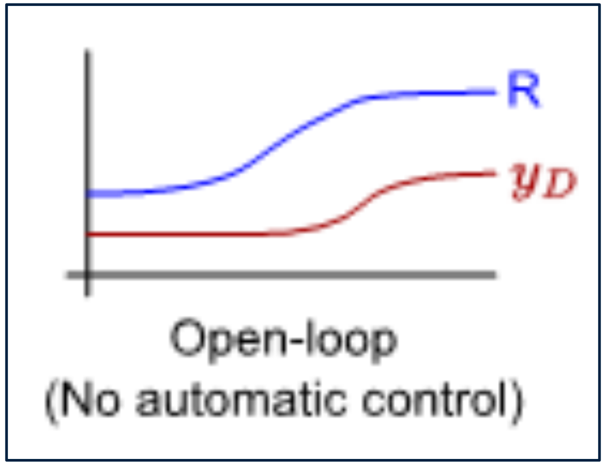


- **[MacGregor, 1991]** Process data was collected at different steady states in a distillation column for reflux ratio, $\boldsymbol{R}$ and overhead purity, $\boldsymbol{y_D}$.

[5] MacGregor, J., Marlin, T., and Kresta, J. (1991). Some comments on neural networks and other empirical modelling methods. In *Proc. of CPC-IV, South Padre Island, TX*, 665–672. CACHE-AIChE.

Example 3-0: Simple Distillation Column



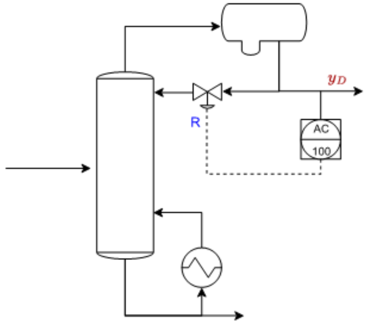


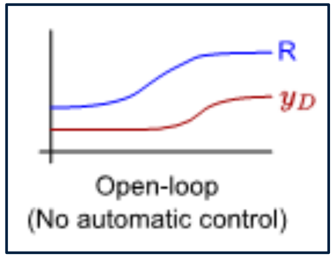
Open-loop
(No automatic control)

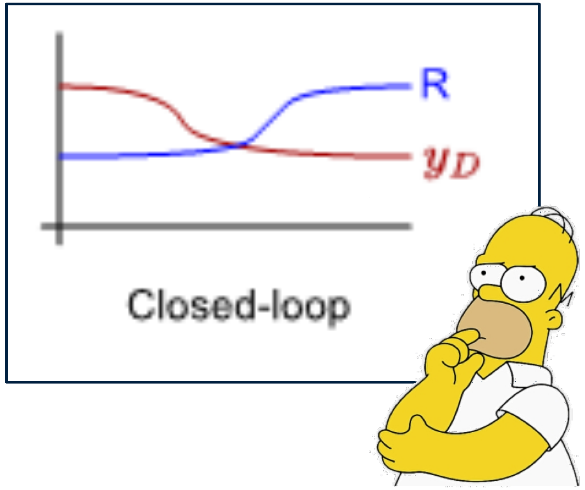
- **[MacGregor, 1991]** Process data was collected at different steady states in a distillation column for reflux ratio, $\boldsymbol{R}$ and overhead purity, $\boldsymbol{y_D}$.
- Basic chemical engineering principles tell us that purity should **increase** with reflux.

[5] MacGregor, J., Marlin, T., and Kresta, J. (1991). Some comments on neural networks and other empirical modelling methods. In *Proc. of CPC-IV, South Padre Island, TX*, 665–672. CACHE-AIChE.

Example 3-0: Simple Distillation Column



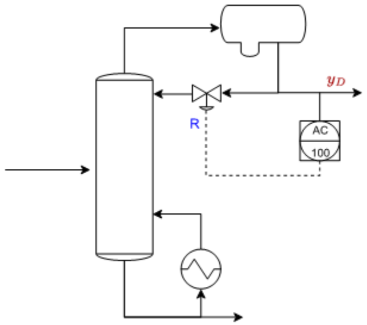


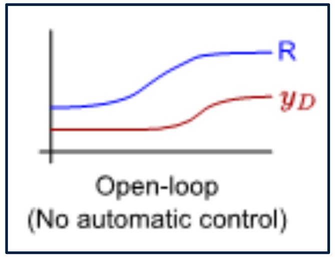


- **[MacGregor, 1991]** Process data was collected at different steady states in a distillation column for reflux ratio, $\boldsymbol{R}$ and overhead purity, $\boldsymbol{y_D}$.
- Basic chemical engineering principles tell us that purity should **increase** with reflux.
- However, an engineer applying regression to the data found a **negative** correlation, which made no physical sense.

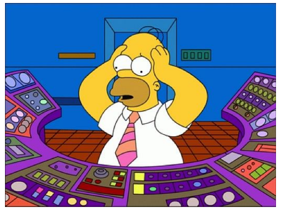
[5] MacGregor, J., Marlin, T., and Kresta, J. (1991). Some comments on neural networks and other empirical modelling methods. In *Proc. of CPC-IV, South Padre Island, TX*, 665–672. CACHE-AIChE.

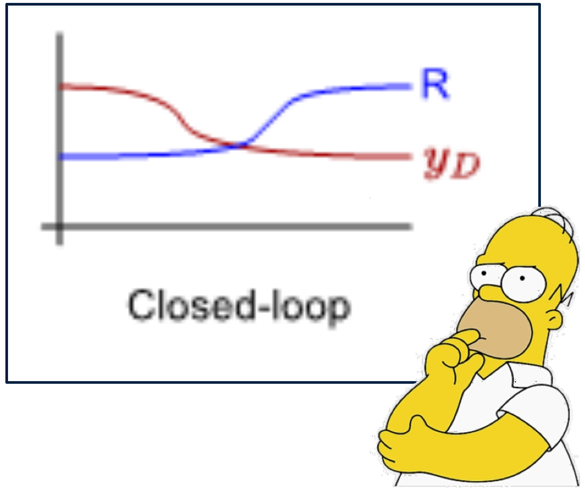
Example 3-0: Simple Distillation Column





Open-loop
(No automatic control)



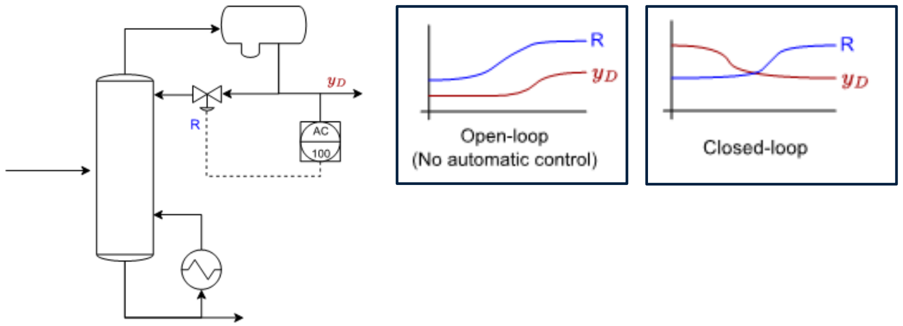


Closed-loop

- **[MacGregor, 1991]** Process data was collected at different steady states in a distillation column for reflux ratio, $\boldsymbol{R}$ and overhead purity, $\boldsymbol{y_D}$.
- Basic chemical engineering principles tell us that purity should **increase** with reflux.
- However, an engineer applying regression to the data found a **negative** correlation, which made no physical sense.
- Eventually, the engineer found that operator was manually increasing reflux ratio when purity was low due to disturbances, and vice-versa.

[5] MacGregor, J., Marlin, T., and Kresta, J. (1991). Some comments on neural networks and other empirical modelling methods. In *Proc. of CPC-IV, South Padre Island, TX*, 665–672. CACHE-AIChE.

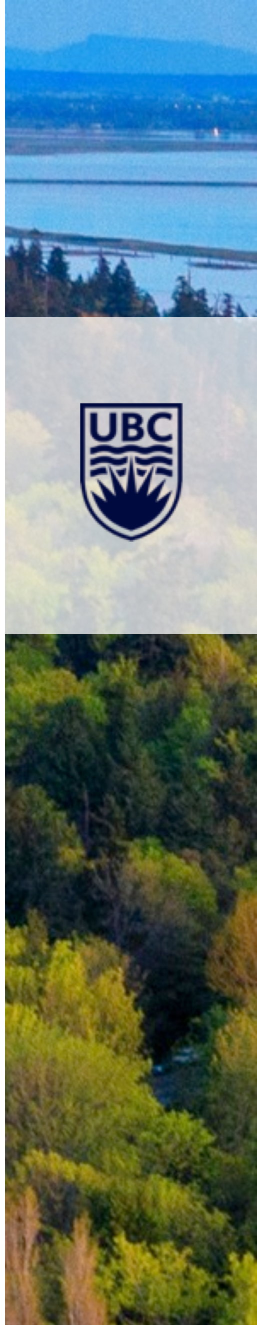
Example 3-0: Simple Distillation Column



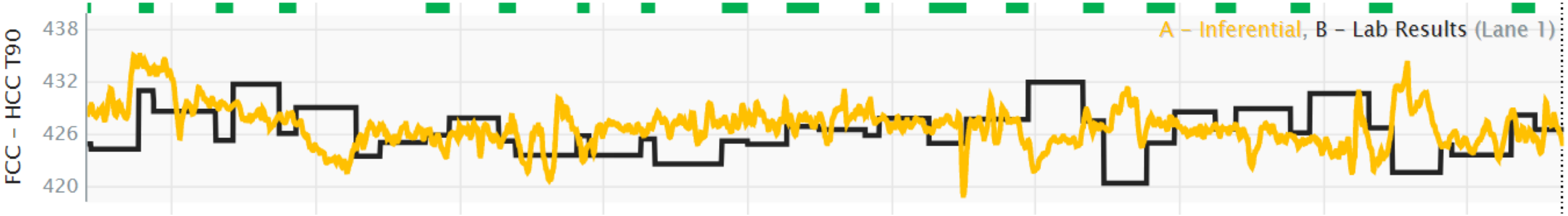
Regression correctly captured the negative correlation between reflux and purity due to operator actions, but failed to provide information on fundamental relationships in the absence of feedback!

- **[MacGregor, 1991]** Process data was collected at different steady states in a distillation column for reflux ratio, $\boldsymbol{R}$ and overhead purity, $\boldsymbol{y_D}$.
- Basic chemical engineering principles tell us that purity should **increase** with reflux.
- However, an engineer applying regression to the data found a **negative** correlation, which made no physical sense.
- Eventually, the engineer found that operator was manually increasing reflux ratio when purity was low due to disturbances, and vice-versa.

[5] MacGregor, J., Marlin, T., and Kresta, J. (1991). Some comments on neural networks and other empirical modelling methods. In *Proc. of CPC-IV, South Padre Island, TX*, 665–672. CACHE-AIChE.



Pitfall 4 – Mishandling multi-rate data



- **Background:** Industrial datasets have **multiple, irregular sampling rates.**
    - *Fast Process Data:* real-time measurements like temperatures, flows etc. **(yellow)**
    - *Slow Quality Data:* offline measurements like lab samples **(black)**
- **Problem:** the time interval between sample collection and lab results could take hours/days
- **Implications:** must consider how to align fast data and slow data during data cleaning

Pitfall 4 – Mishandling multi-rate data

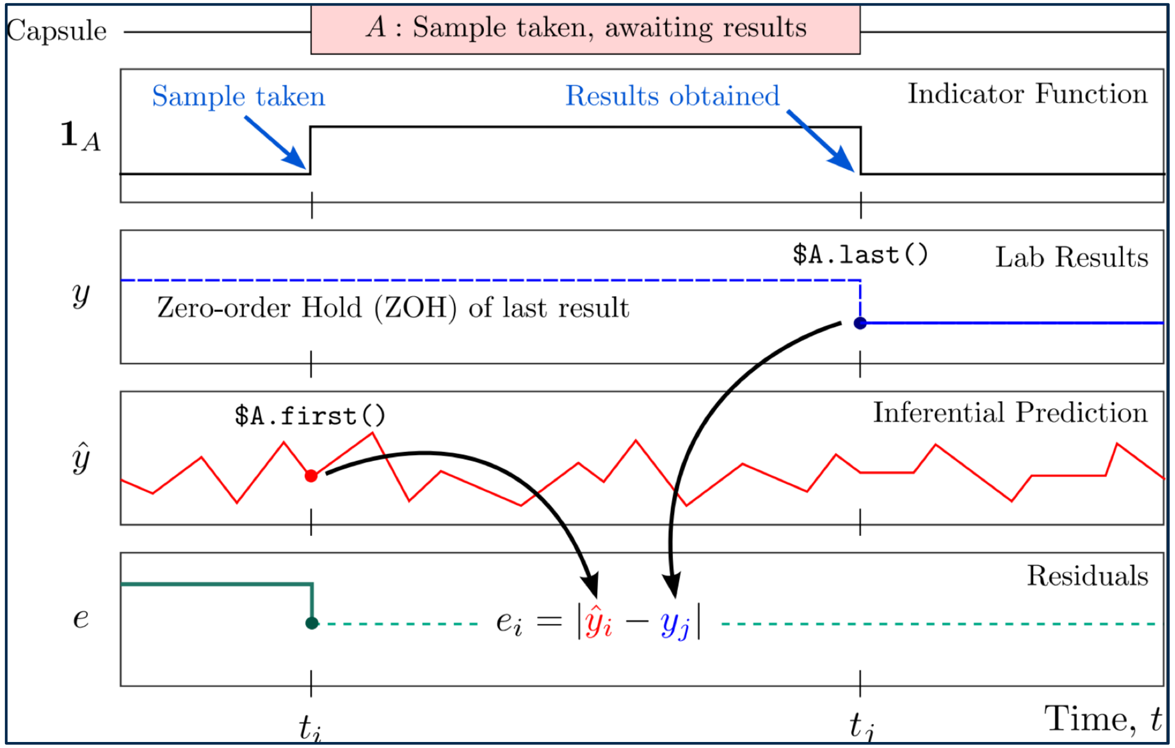


Figure 13: Visual representation of time shift adjustment for updating lab sample inferentials.

Siang, L. C., Elnawawi, S., & Steele, D. (2022). Self-Service Analytics and the Processing of Hydrocarbons. Digital Chemical Engineering, 100021.

Example - soft sensor maintenance:

- Monitor model performance by calculating residuals between predictions and lab results.
- Lab results may only be available hours or days after sample was taken
- Re-aligning lab results back to sample collection time is a critical data cleansing step that is often overlooked.

Pitfall 5 – Chasing irrelevant model metrics

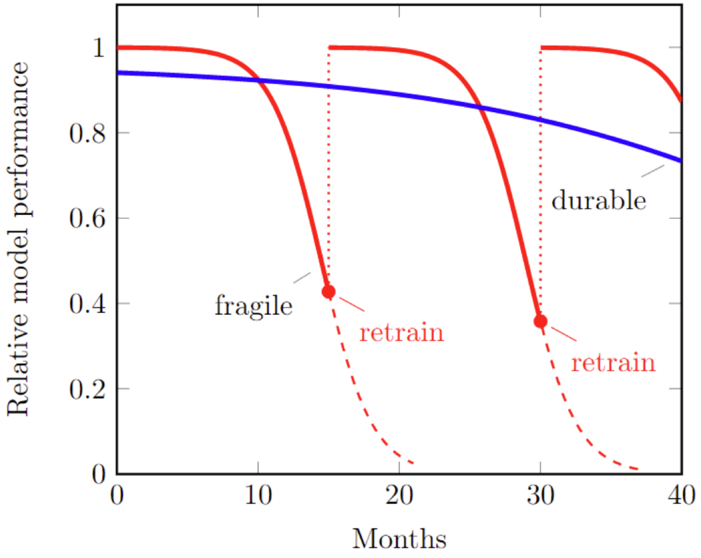


Figure 15: Durable inferential sensors reduce the burden of model maintenance. Model accuracy is important but should also be considered in the broader context of model maintenance and other performance trade-offs.

- **Be mindful of the 80/20 rule:** Squeezing out marginal improvements in model accuracy can lead to diminishing (business) returns.
  - *Example:* If you can build a simple model that works reasonably well and satisfies business stakeholders, it probably won't make sense to spend months trying to build the 'perfect' model with very low prediction error.

Pitfall 5 – Chasing irrelevant model metrics

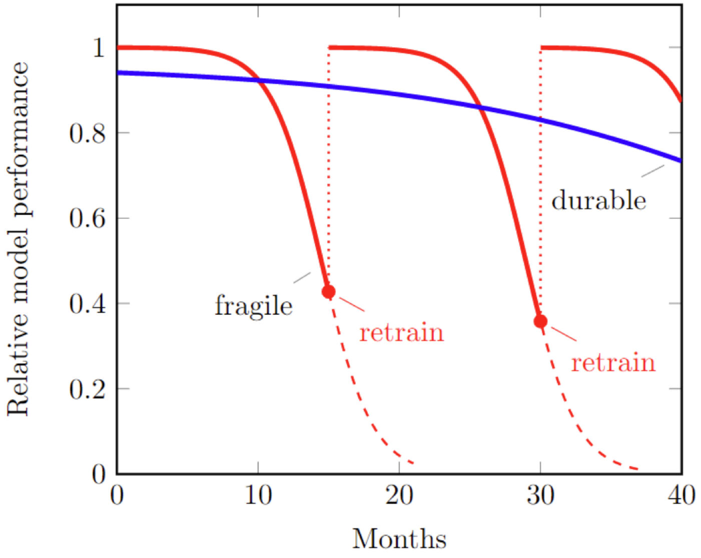
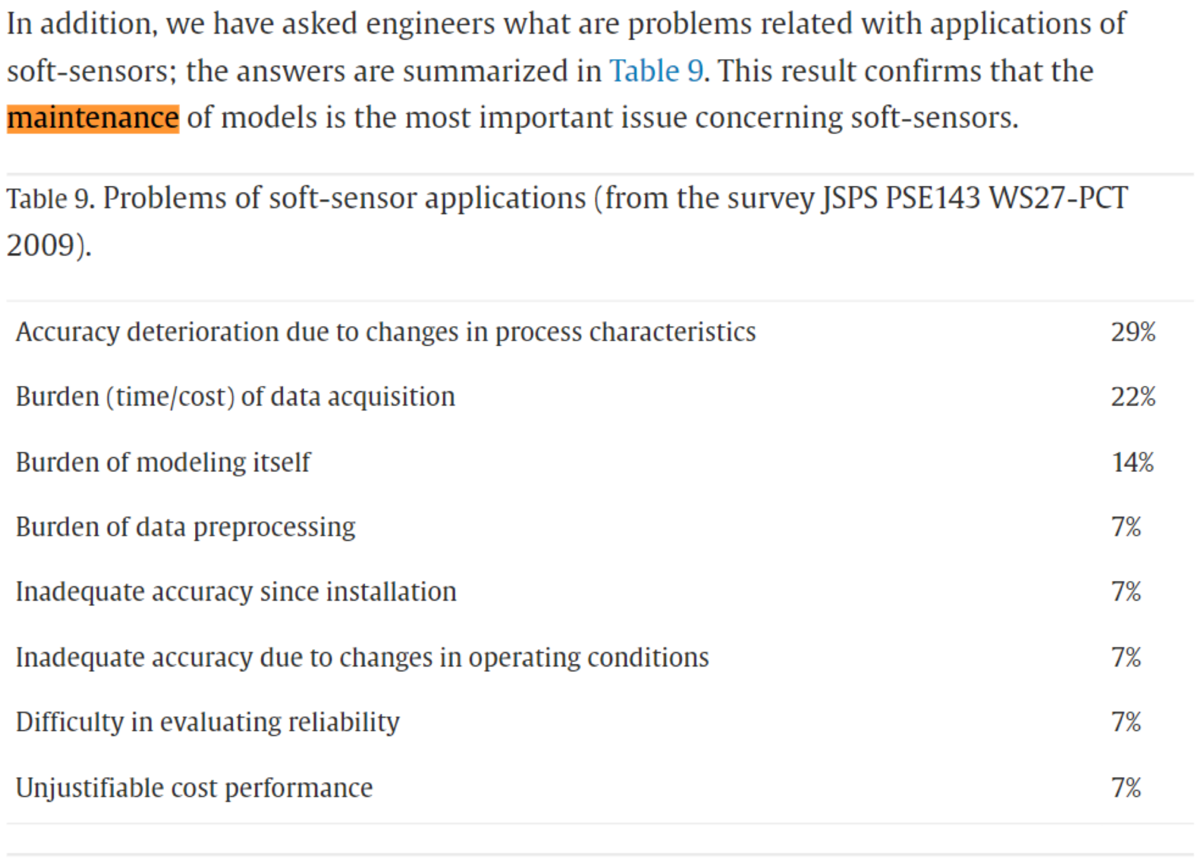


Figure 15: Durable inferential sensors reduce the burden of model maintenance. Model accuracy is important but should also be considered in the broader context of model maintenance and other performance trade-offs.

- **Consider not just accuracy, but also model maintenance and other performance trade-offs:** Soft sensors degrade over time due to process drifts, instrumentation issues etc.
  - Accuracy metrics R^2, *RMSE, MAE* etc. are just one part of the story for industrial implementation
  - *Example:* A durable model may be more valuable than a slightly more accurate one that degrades faster and needs frequent maintenance.

Pitfall 5 – Chasing irrelevant model metrics



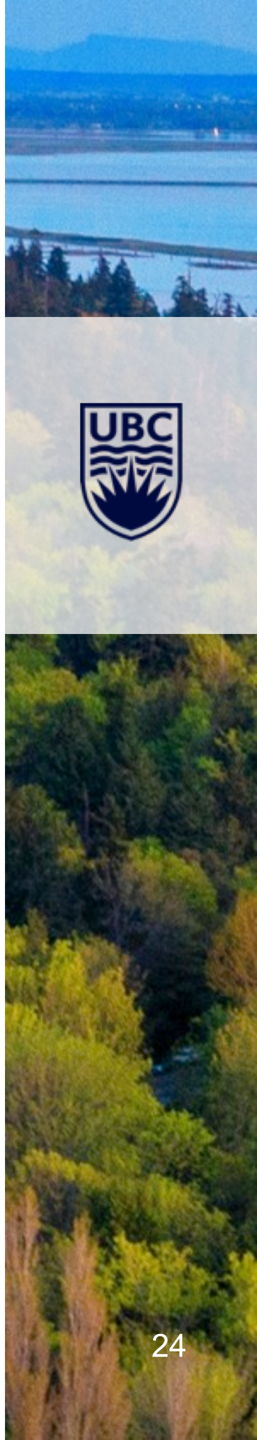
In addition, we have asked engineers what are problems related with applications of soft-sensors; the answers are summarized in Table 9. This result confirms that the maintenance of models is the most important issue concerning soft-sensors.

Table 9. Problems of soft-sensor applications (from the survey JSPS PSE143 WS27-PCT 2009).

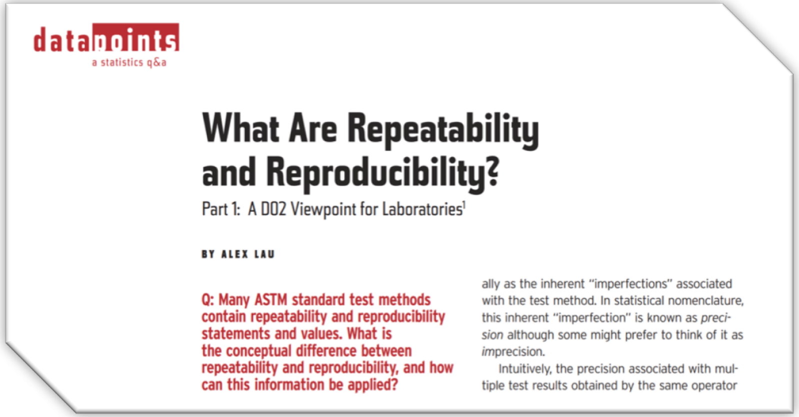
Accuracy deterioration due to changes in process characteristics	29%
Burden (time/cost) of data acquisition	22%
Burden of modeling itself	14%
Burden of data preprocessing	7%
Inadequate accuracy since installation	7%
Inadequate accuracy due to changes in operating conditions	7%
Difficulty in evaluating reliability	7%
Unjustifiable cost performance	7%

- Survey of the chemical process industry in Japan conducted in 2010
- For soft sensors, model maintenance is the most important issue faced by practitioners, more so than accuracy, modeling etc.

Kano, M., & Ogawa, M. (2010). The state of the art in chemical process control in Japan: Good practice and questionnaire survey. *Journal of Process Control*



Example 5-1: Consider uncertainties in the ground truth



datapoints
a statistics q&a

What Are Repeatability and Reproducibility?

Part 1: A D02 Viewpoint for Laboratories[1]

BY ALEX LAU

Q: Many ASTM standard test methods contain repeatability and reproducibility statements and values. What is the conceptual difference between repeatability and reproducibility, and how can this information be applied?

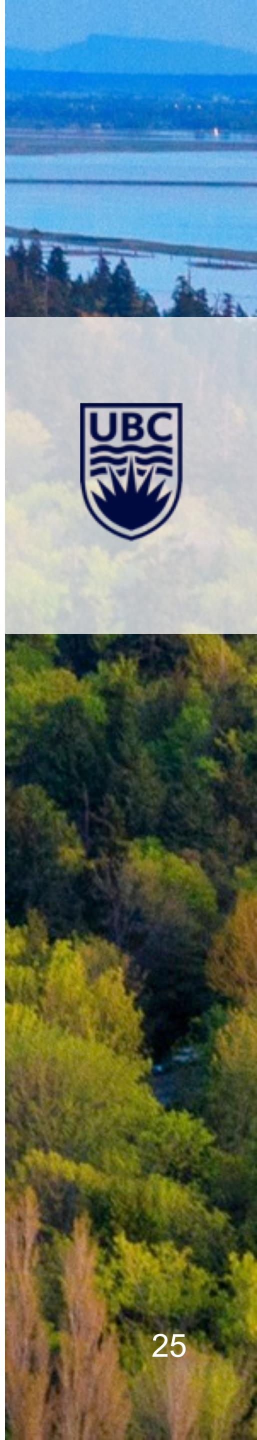
ally as the inherent "imperfections" associated with the test method. In statistical nomenclature, this inherent "imperfection" is known as *precision* although some might prefer to think of it as *imprecision*.

Intuitively, the precision associated with multiple test results obtained by the same operator

Reference: https://sn.astm.org/data-points/what-are-repeatability-and-reproducibility-ma09.html

- ASTM definition for measurement precisions under specific test conditions: **Reproducibility (R)** – uncertainty with test conducted in different labs, operators and apparatus
- **Example:** ASTM D2700 for measuring fuel motor octane number (MON) has a reproducibility of $R = 0.9$.
  - This means that a sample tested for MON in 2 different labs with differences below 0.9 can be explained only by the test method precision, not equipment/operator error.
- **Implications:** A soft sensor for predicting MON with a typical value of 90 can't possibly have an error of less than 1% using lab samples tested in different facilities as the ground truth.

Understand how the ground truth data is obtained, as well as its limits and uncertainties



Takeaways: Pitfalls and Guidelines for Process Analytics

1. Contextualize your process data

2. Apply domain knowledge

3. Account for closed-loop conditions in industrial processes

4. Align multi-rate data correctly

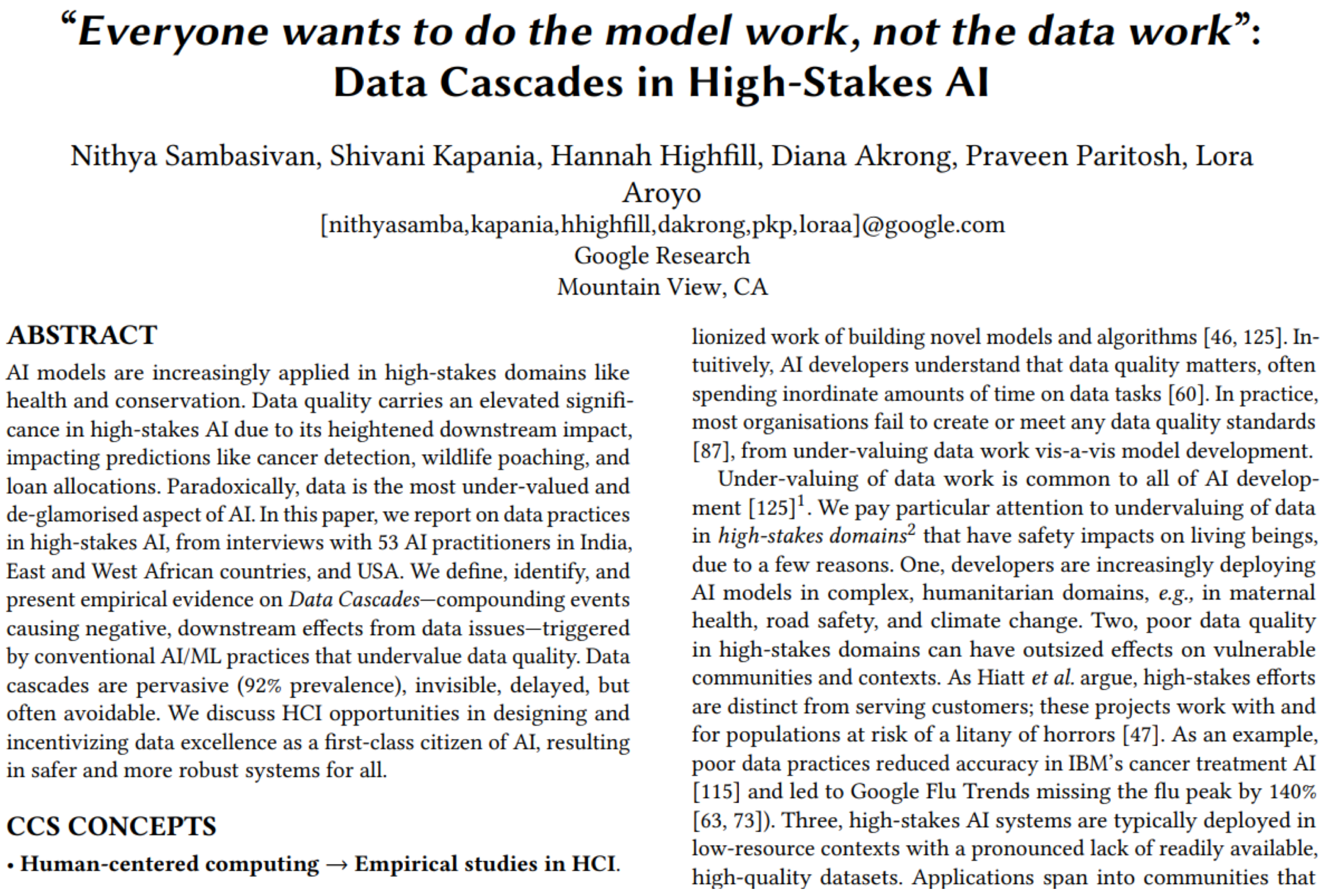
5. Consider model durability and maintenance, not just accuracy

Burnaby Refinery, BC, Canada

https://APCPapers.github.io/

THE UNIVERSITY OF BRITISH COLUMBIA

Appendix & Extra Slides



"*Everyone wants to do the model work, not the data work*": Data Cascades in High-Stakes AI

Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, Lora Aroyo
[nithyasamba,kapania,hhighfill,dakrong,pkp,loraa]@google.com
Google Research
Mountain View, CA

ABSTRACT

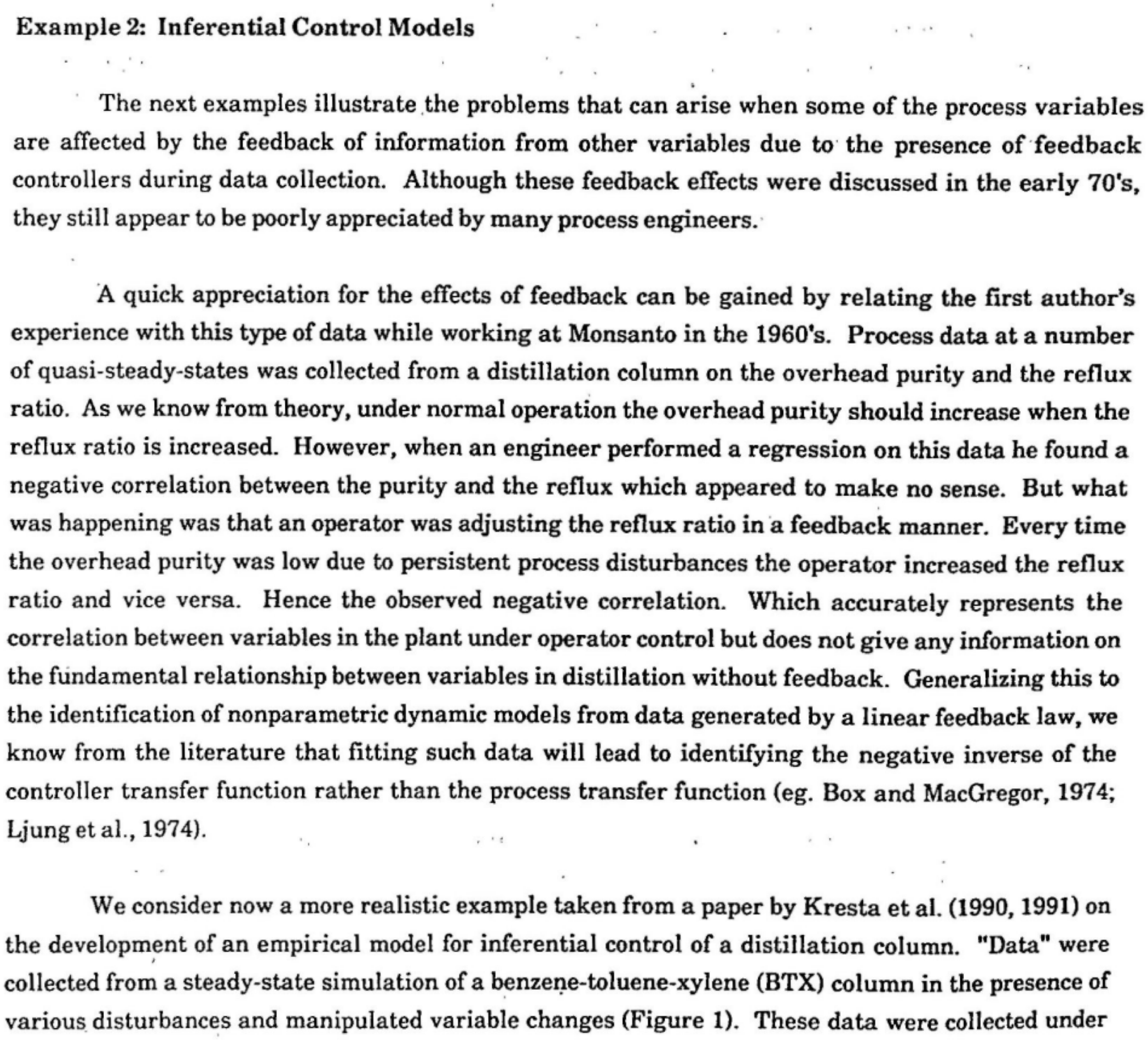
AI models are increasingly applied in high-stakes domains like health and conservation. Data quality carries an elevated significance in high-stakes AI due to its heightened downstream impact, impacting predictions like cancer detection, wildlife poaching, and loan allocations. Paradoxically, data is the most under-valued and de-glamorised aspect of AI. In this paper, we report on data practices in high-stakes AI, from interviews with 53 AI practitioners in India, East and West African countries, and USA. We define, identify, and present empirical evidence on *Data Cascades*—compounding events causing negative, downstream effects from data issues—triggered by conventional AI/ML practices that undervalue data quality. Data cascades are pervasive (92% prevalence), invisible, delayed, but often avoidable. We discuss HCI opportunities in designing and incentivizing data excellence as a first-class citizen of AI, resulting in safer and more robust systems for all.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI.**

lionized work of building novel models and algorithms [46, 125]. Intuitively, AI developers understand that data quality matters, often spending inordinate amounts of time on data tasks [60]. In practice, most organisations fail to create or meet any data quality standards [87], from under-valuing data work vis-a-vis model development.

Under-valuing of data work is common to all of AI development [125][1]. We pay particular attention to undervaluing of data in *high-stakes domains*[2] that have safety impacts on living beings, due to a few reasons. One, developers are increasingly deploying AI models in complex, humanitarian domains, *e.g.,* in maternal health, road safety, and climate change. Two, poor data quality in high-stakes domains can have outsized effects on vulnerable communities and contexts. As Hiatt *et al.* argue, high-stakes efforts are distinct from serving customers; these projects work with and for populations at risk of a litany of horrors [47]. As an example, poor data practices reduced accuracy in IBM's cancer treatment AI [115] and led to Google Flu Trends missing the flu peak by 140% [63, 73]). Three, high-stakes AI systems are typically deployed in low-resource contexts with a pronounced lack of readily available, high-quality datasets. Applications span into communities that

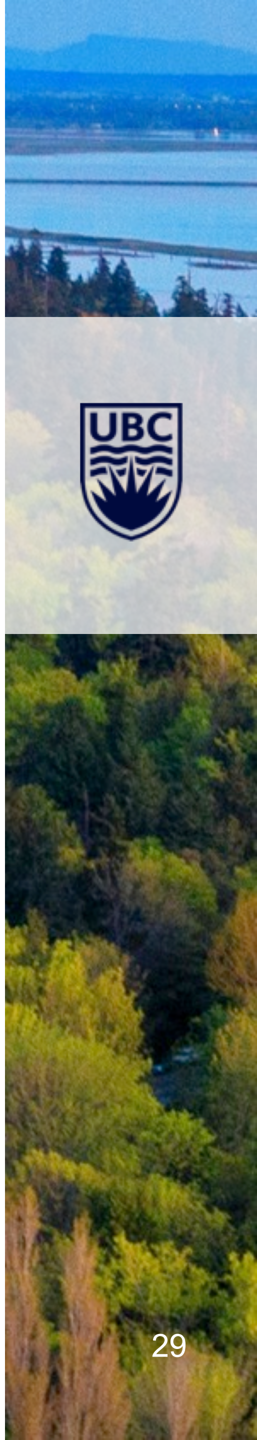


Example 2: Inferential Control Models

The next examples illustrate the problems that can arise when some of the process variables are affected by the feedback of information from other variables due to the presence of feedback controllers during data collection. Although these feedback effects were discussed in the early 70's, they still appear to be poorly appreciated by many process engineers.

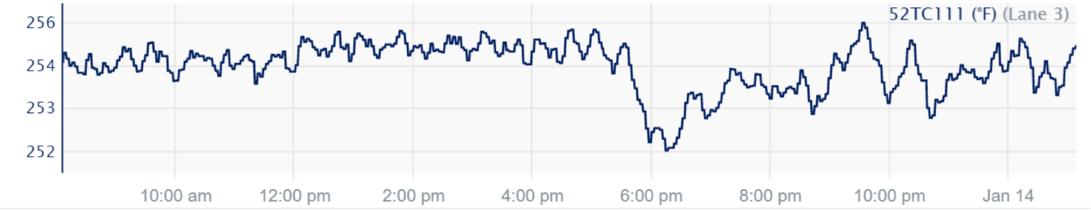
A quick appreciation for the effects of feedback can be gained by relating the first author's experience with this type of data while working at Monsanto in the 1960's. Process data at a number of quasi-steady-states was collected from a distillation column on the overhead purity and the reflux ratio. As we know from theory, under normal operation the overhead purity should increase when the reflux ratio is increased. However, when an engineer performed a regression on this data he found a negative correlation between the purity and the reflux which appeared to make no sense. But what was happening was that an operator was adjusting the reflux ratio in a feedback manner. Every time the overhead purity was low due to persistent process disturbances the operator increased the reflux ratio and vice versa. Hence the observed negative correlation. Which accurately represents the correlation between variables in the plant under operator control but does not give any information on the fundamental relationship between variables in distillation without feedback. Generalizing this to the identification of nonparametric dynamic models from data generated by a linear feedback law, we know from the literature that fitting such data will lead to identifying the negative inverse of the controller transfer function rather than the process transfer function (eg. Box and MacGregor, 1974; Ljung et al., 1974).

We consider now a more realistic example taken from a paper by Kresta et al. (1990, 1991) on the development of an empirical model for inferential control of a distillation column. "Data" were collected from a steady-state simulation of a benzene-toluene-xylene (BTX) column in the presence of various disturbances and manipulated variable changes (Figure 1). These data were collected under



Example 3-1: The complexity of a simple control loop

Figure 12: Plot of PV for a temperature controller 52TC111. This data passed the initial check, there were no compression issues or hidden calculations.

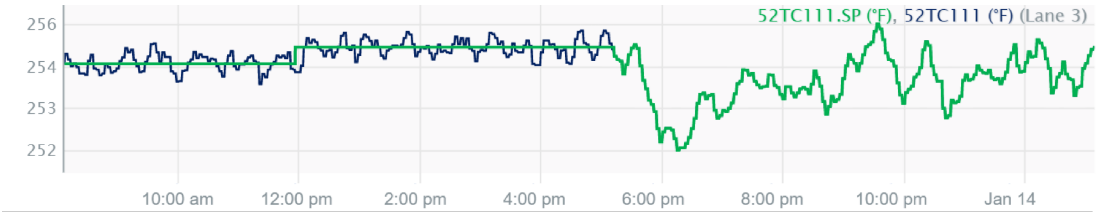


Observations

- PV fluctuates more after 5PM. **Why?**

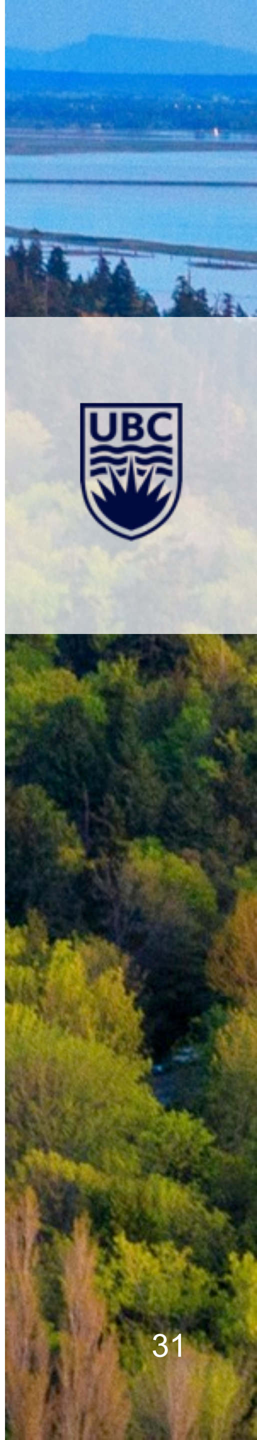
Example 3-1: The complexity of a simple control loop

Figure 12: Plot of SP and PV for a temperature controller 52TC111. The SP started tracking the PV after 5PM.



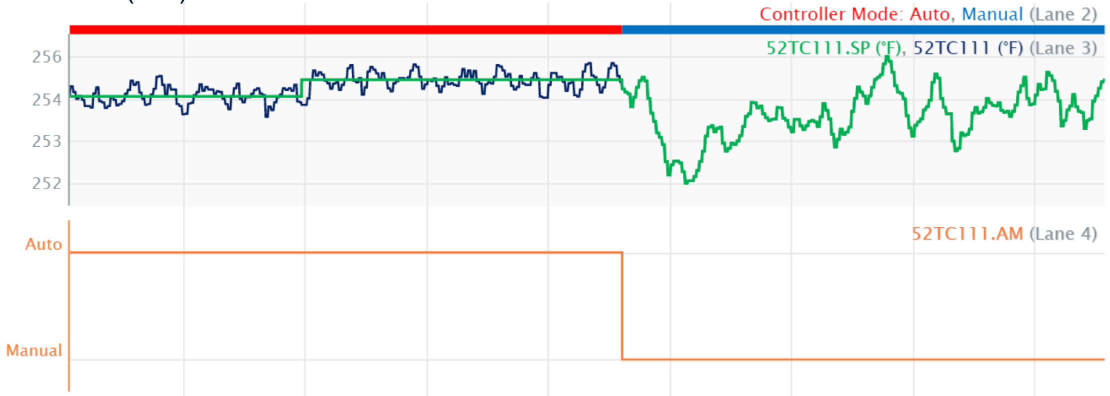
Observations

- PV fluctuates more after 5PM.
- Plot SP: initially the PV was tracking the SP, but after 5PM the SP tracked the PV. **Why?**



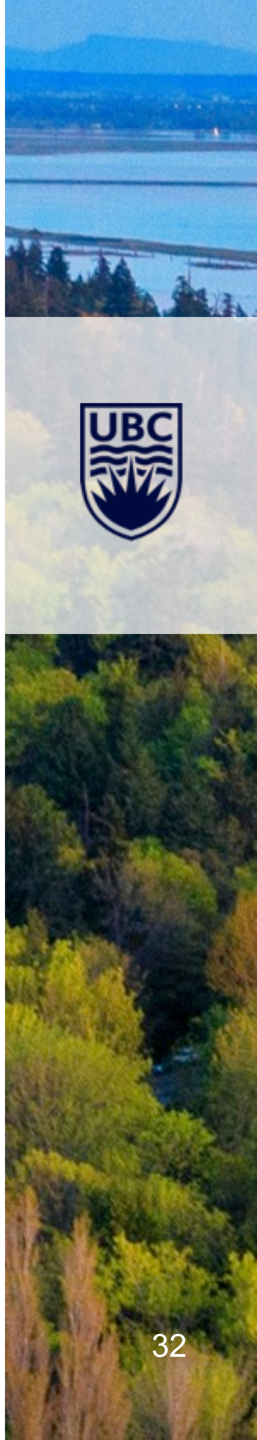
Example 3-1: The complexity of a simple control loop

Figure 12: PV and SP plot for a temperature controller 52TC111 showing loop mode (AM) switch from Auto to Manual.



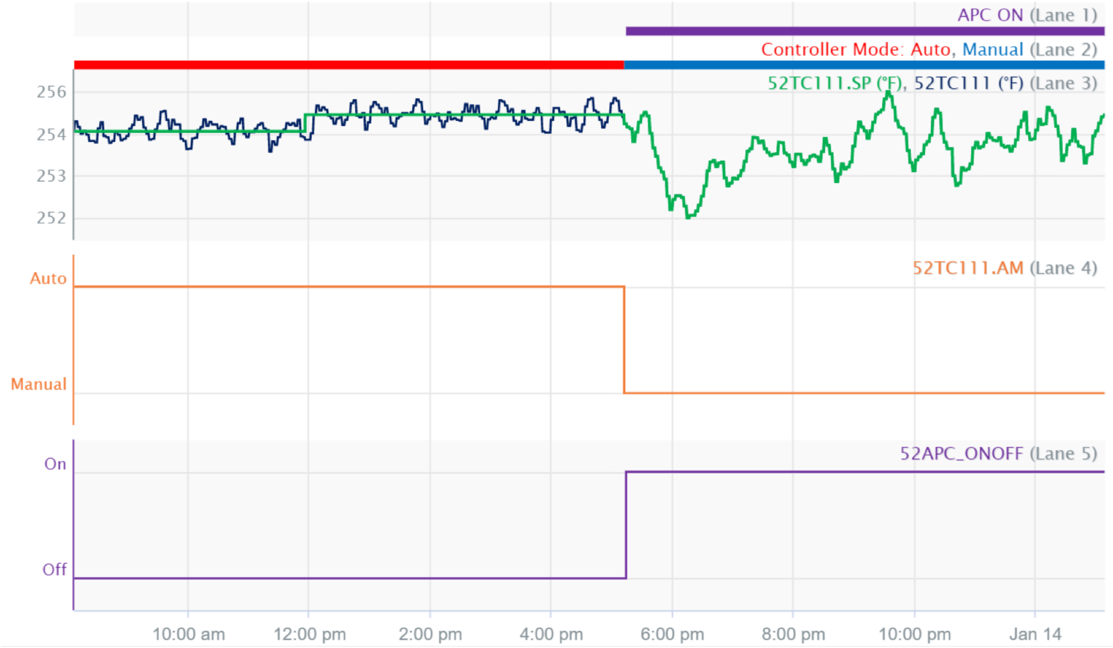
Observations

- PV fluctuates more after 5PM.
- Plot SP: initially the PV was tracking the SP, but after 5PM the SP tracked the PV.
- Plot .MODE or .AM tag: in manual mode, SP tracks PV for bumpless transfer. **Why?**



Example 3-1: The complexity of a simple control loop

Figure 12: Plot for a temperature controller 52TC111 showing loop mode (AM) switch from Auto to Manual as well as APC On switch.

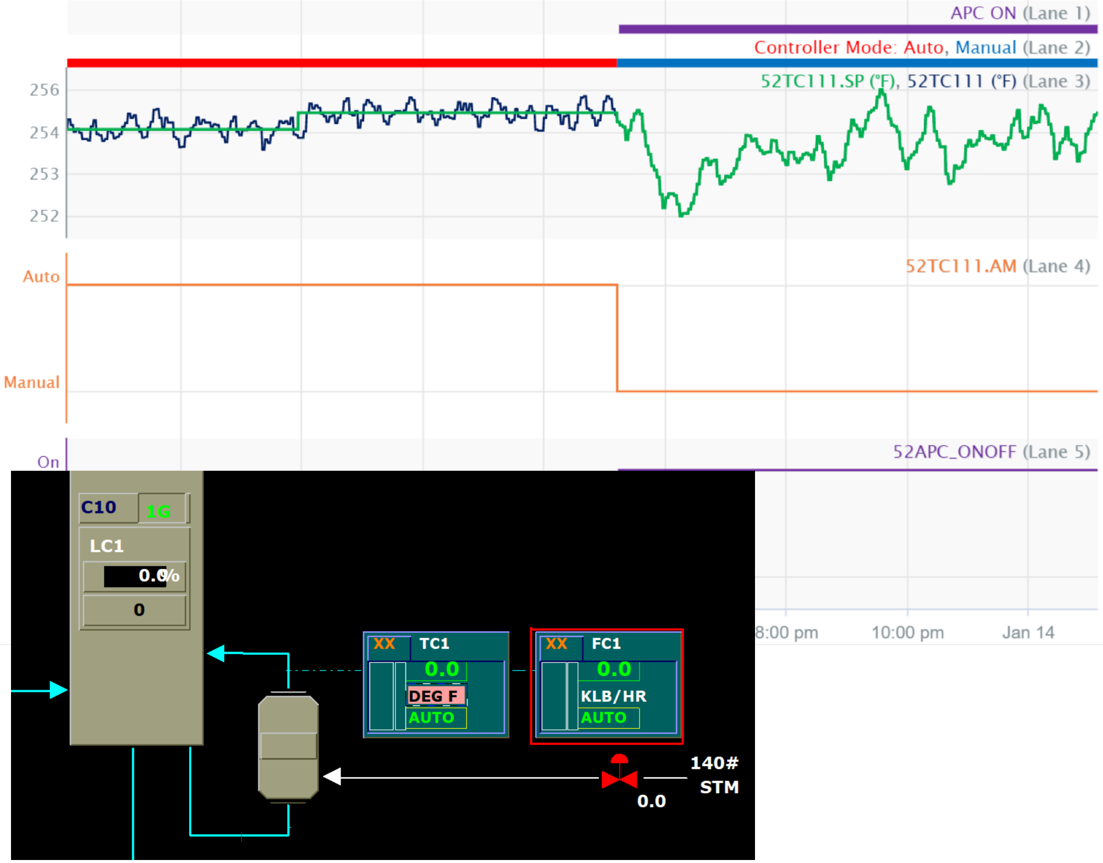


Observations

- PV fluctuates more after 5PM.
- Plot SP: initially the PV was tracking the SP, but after 5PM the SP tracked the PV.
- Plot .MODE or .AM tag: in manual mode, SP tracks PV for bumpless transfer.
- APC was turned on, which caused the loop to go from AUTO to MANUAL. **Why?**

Example 3-1: The complexity of a simple control loop

Figure 12: Plot for a temperature controller 52TC111 showing PID and APC loop mode changes. APC/cascade configuration is shown at the bottom.

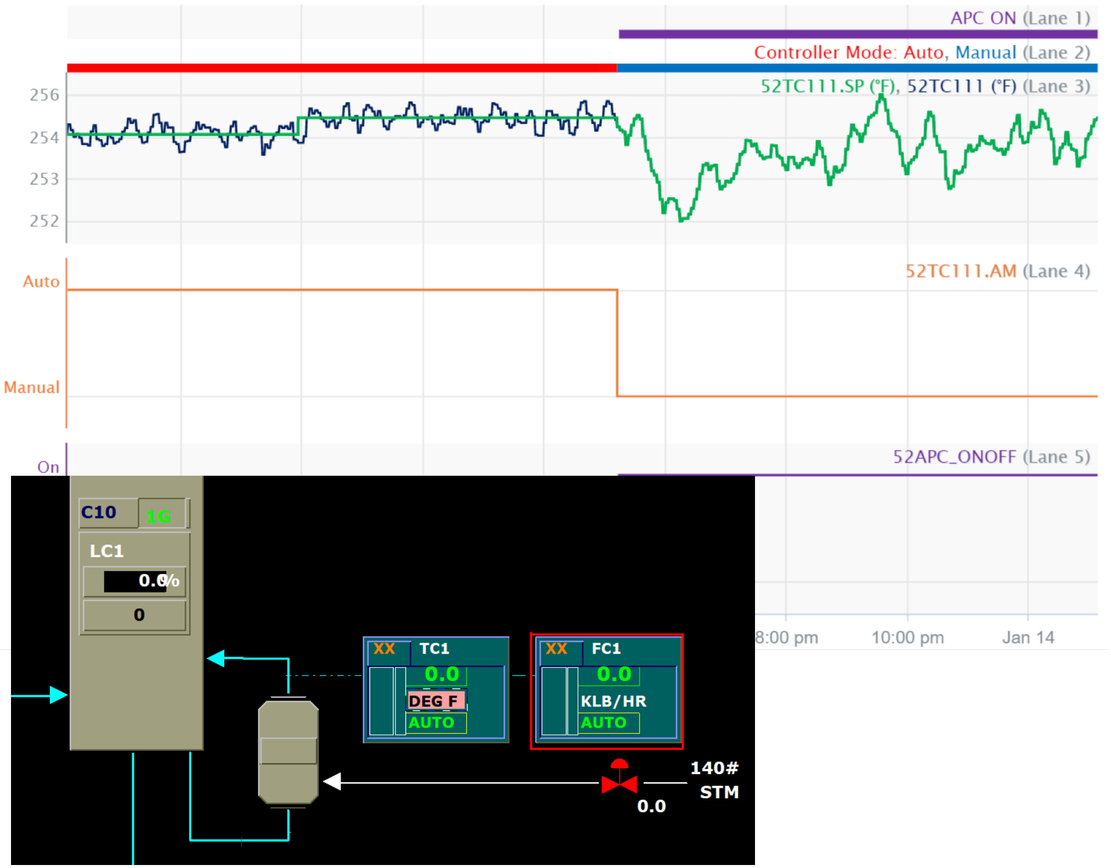


Observations

- PV fluctuates more after 5PM.
- Plot SP: initially the PV was tracking the SP, but after 5PM the SP tracked the PV.
- Plot .MODE or .AM tag: in manual mode, SP tracks PV for bumpless transfer.
- APC was turned on, which caused the loop to go from AUTO to MANUAL.
- 52TC111 is part of a cascade loop. When APC is ON, the cascade loop is broken and APC controls 52FC111 directly. 52TC111 no longer involved in control.

Example 3-1: The complexity of a simple control loop

Figure 12: Plot for a temperature controller 52TC111 showing PID and APC loop mode changes. APC/cascade configuration is shown at the bottom.

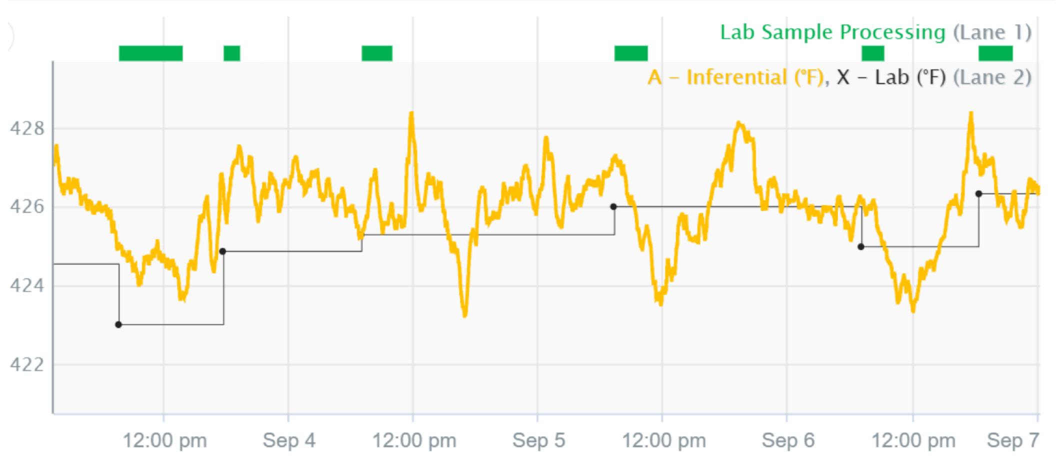


- PV fluctuates more after 5PM.
- Plot SP: initially the PV was tracking the SP, but after 5PM the SP tracked the PV.
- Plot .MODE or .AM tag: in manual mode, SP tracks PV for bumpless transfer.
- APC was turned on, which caused the loop to go from AUTO to MANUAL.
- 52TC111 is part of a cascade loop. When APC is ON, the cascade loop is broken and APC controls 52FC111 directly. 52TC111 no longer involved in control.

Loop configurations can be complex: check your operating modes before using the data

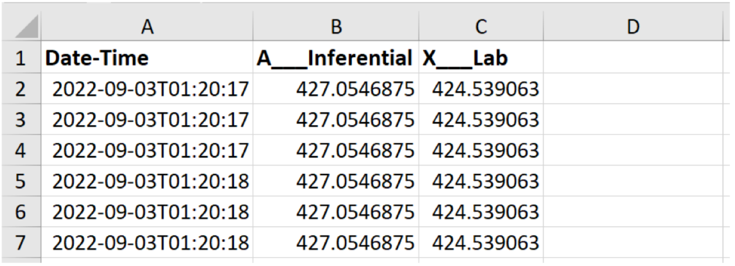
Example 4-1: Multi-rate data retrieval and alignment

Figure 14: Time-series of inferential prediction (high frequency), lab results (twice daily) and time required for lab sample processing (green bar).



Problem: Process data can be acquired at high frequency, but lab measurements are only available twice daily.

Export

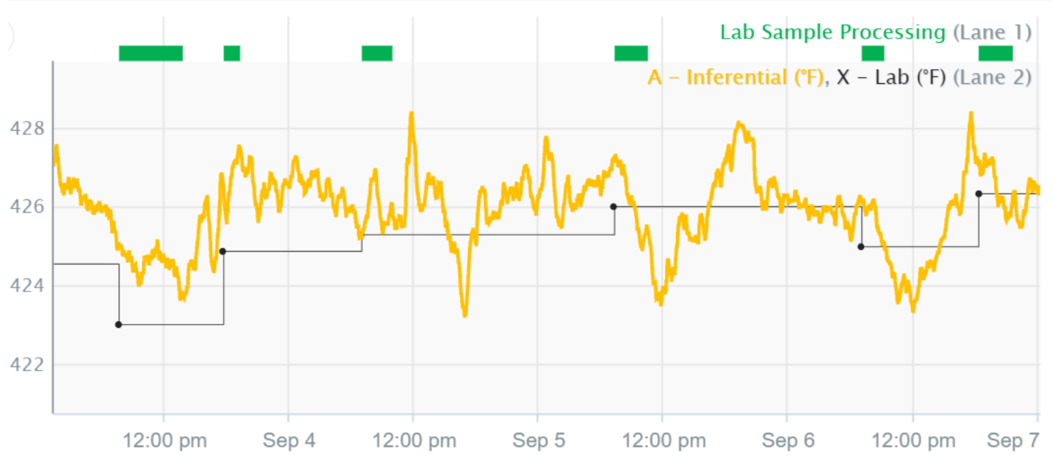


	A	B	C	D
1	Date-Time	A___Inferential	X___Lab	
2	2022-09-03T01:20:17	427.0546875	424.539063	
3	2022-09-03T01:20:17	427.0546875	424.539063	
4	2022-09-03T01:20:17	427.0546875	424.539063	
5	2022-09-03T01:20:18	427.0546875	424.539063	
6	2022-09-03T01:20:18	427.0546875	424.539063	
7	2022-09-03T01:20:18	427.0546875	424.539063	

Single grid: apply zero-order hold on lab samples and align to the same timestamp grid *(default export settings)*

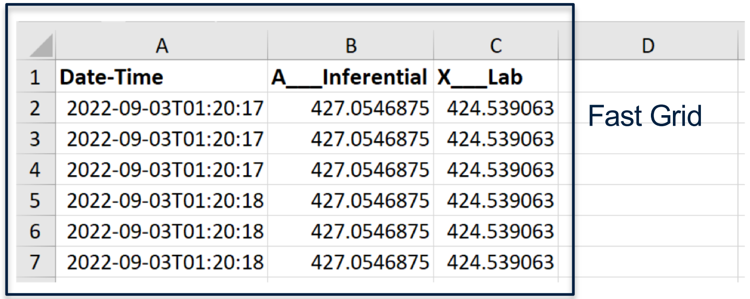
Example 4-1: Multi-rate data retrieval and alignment

Figure 14: Time-series of inferential prediction (high frequency), lab results (twice daily) and time required for lab sample processing (green bar).



Problem: Process data can be acquired at high frequency, but lab measurements are only available twice daily.

Export

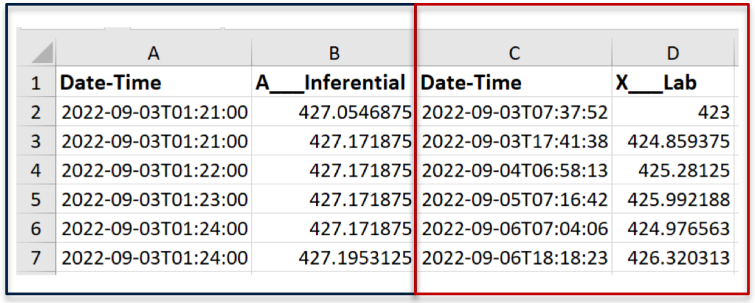


	A	B	C	D
1	Date-Time	A___Inferential	X___Lab	
2	2022-09-03T01:20:17	427.0546875	424.539063	Fast Grid
3	2022-09-03T01:20:17	427.0546875	424.539063	
4	2022-09-03T01:20:17	427.0546875	424.539063	
5	2022-09-03T01:20:18	427.0546875	424.539063	
6	2022-09-03T01:20:18	427.0546875	424.539063	
7	2022-09-03T01:20:18	427.0546875	424.539063	

Single grid: apply zero-order hold on lab samples and align to the same timestamp grid *(default export settings)*

Fast Grid

Slow Grid



	A	B	C	D
1	Date-Time	A___Inferential	Date-Time	X___Lab
2	2022-09-03T01:21:00	427.0546875	2022-09-03T07:37:52	423
3	2022-09-03T01:21:00	427.171875	2022-09-03T17:41:38	424.859375
4	2022-09-03T01:22:00	427.171875	2022-09-04T06:58:13	425.28125
5	2022-09-03T01:23:00	427.171875	2022-09-05T07:16:42	425.992188
6	2022-09-03T01:24:00	427.171875	2022-09-06T07:04:06	424.976563
7	2022-09-03T01:24:00	427.1953125	2022-09-06T18:18:23	426.320313

Multiple grids: use different grids and decide how to treat 'missing' lab data

Export settings matter: consider separate timestamp grids for multi-rate data