Generalized Likelihood Ratio (GLR) control charts with composite hypotheses

An application to high-purity processes in chemical industry

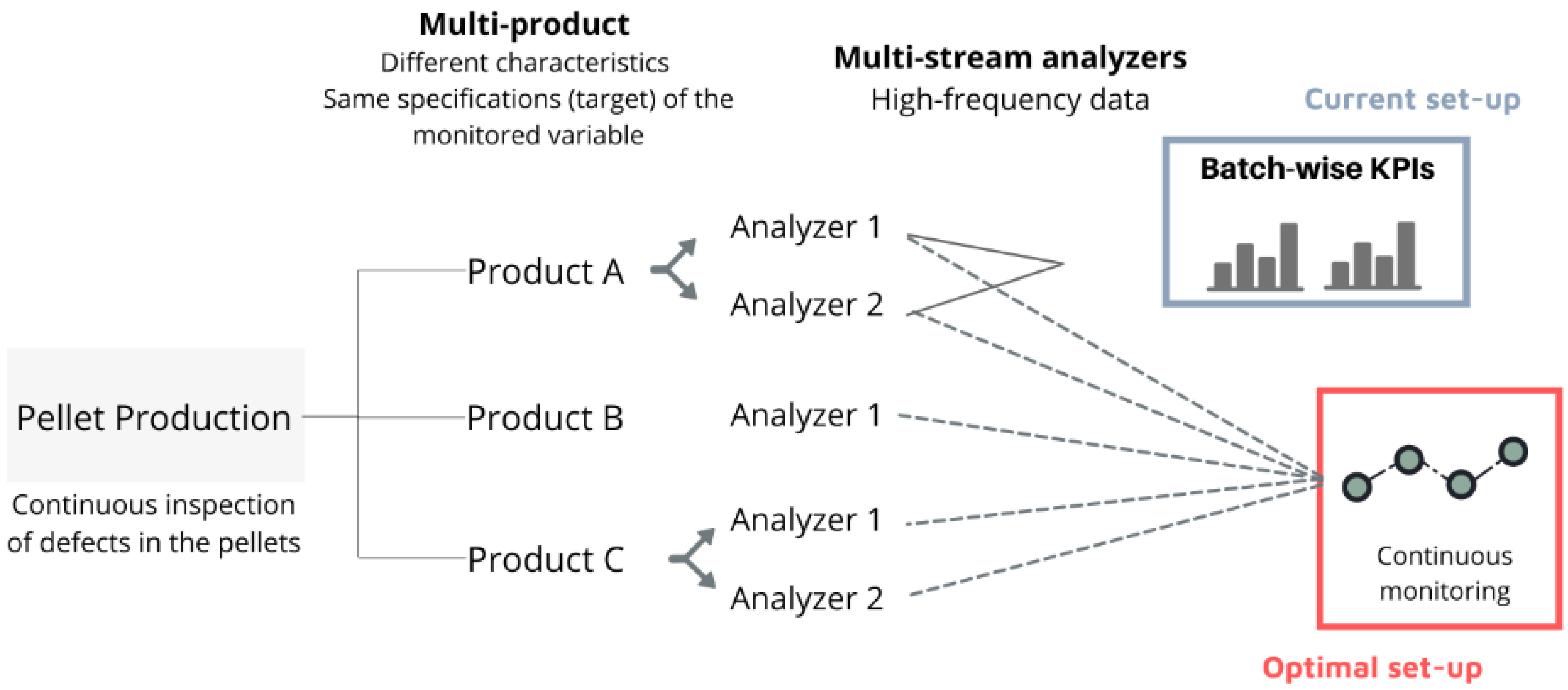
Caterina Rizzo[1,3], Swee-Teng Chin[2], Alessandro Di Bucchianico[3]

[1]Dow Inc., Herbert H. Dowweg 5, 4542 NL Hoek, The Netherlands

[2]Dow Inc., 332 SH 332 E, Lake Jackson, TX 77666, USA

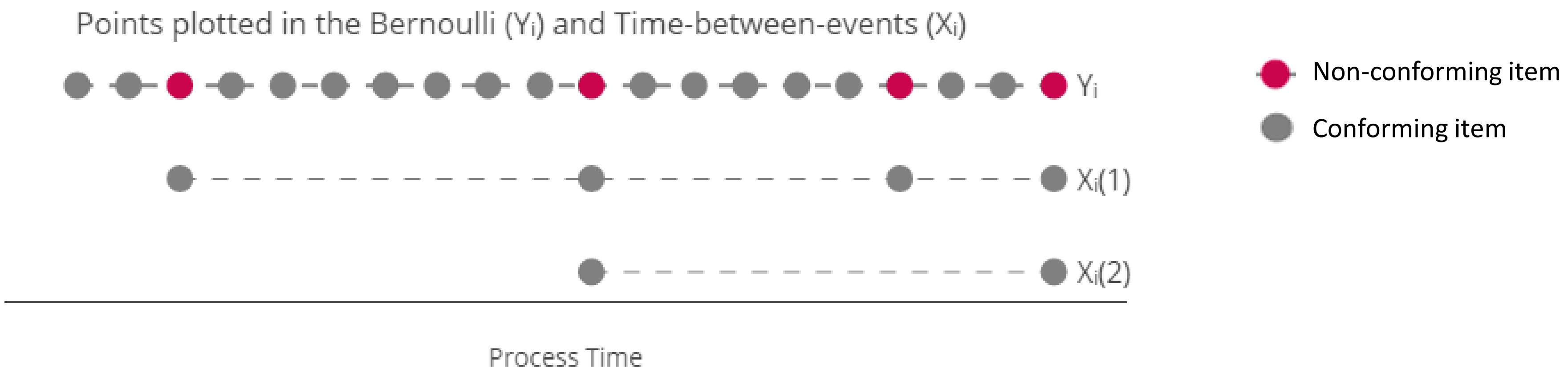
[3]Eindhoven University of Technology, Department of Mathematics, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

The case study: defect monitoring in pellet production



The case study: model and challenges

- Need of monitoring procedures that are **tailored** based on different statistical models and more effective at detecting more realistic out-of-scenarios
- Challenges of aggregating data: process time and control chart time are asynchronous



DEFINITION 1: *Let τ_π the cumulative time from the start at which the change occurs, referring to the process time and let $\tau + 1$ be the first decision point at which it is possible to detect the change on the control chart time, then τ is*

$$\inf\{\tau : \sum_{i=1}^{\tau} x_i \geq \tau_\pi\}$$

Sequential change-point detection under simple hypotheses

- The most widely set of simple hypothesis studied in statistical process control represent an abrupt and **unexpected persistent shift** in the monitored parameter

$$\begin{aligned} H_0: &\quad \theta = \theta_0, \text{ for all } i \\ H_a(\tau): &\quad \begin{cases} \theta_i = \theta_0, \text{ for } i \in (0, \tau] \\ \theta_i = \theta_1, \text{ for } i \in [\tau+1, \infty) \end{cases}, \end{aligned}$$

- **Gamma log-likelihood ratio statistics** The parameter θ_0 is assumed to be well estimated from historical data. The unknown post-change mean is estimated using the maximum likelihood estimation.

$$\ln \Lambda_k^{\Gamma}(\theta_0, \theta_1; \mathbf{x}) = \max_{0 \le \tau < k-1} r(k-\tau) \left[\ln\left(\frac{\theta_0}{\widehat{\theta_1}} \right) + \frac{\widehat{\theta_1} - \theta_0}{\theta_0} \right]$$

- In this context
  - r represent the order of the chart: it is fixed and known a priori
  - The $\ln \Lambda_k^{\Gamma}$ statistics is always defined

Sequential change-point detection under composite hypotheses

Indifference interval model

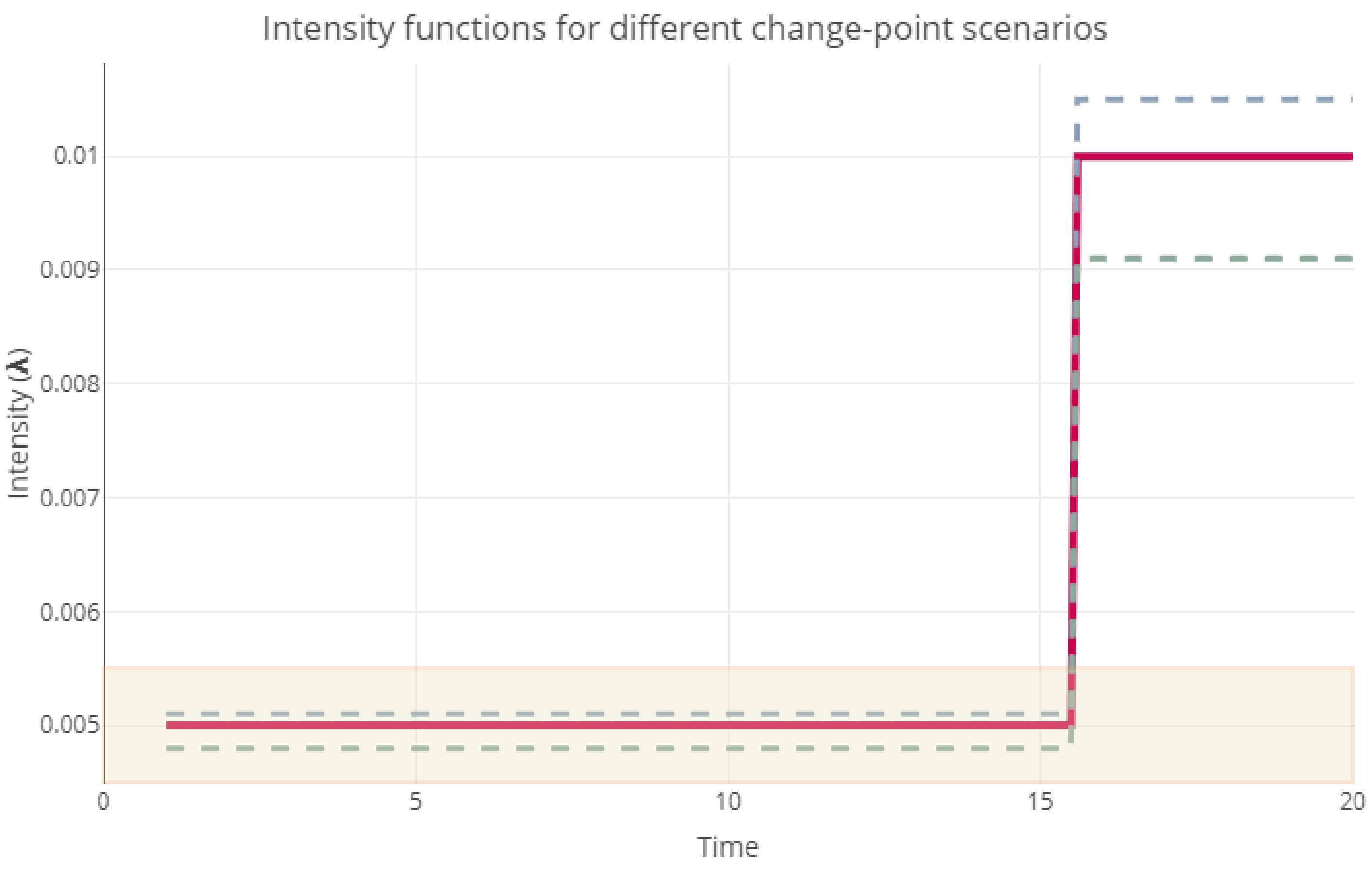
- Simple hypotheses set a hard limit to a decision between one of two possible states of nature; composite hypotheses cover a **set of values** from the parameter space.
- In practical scenarios it is more important to detect a change from a target value allowing a margin, i.e., so that the process mean remains within a certain specified tolerance interval. (**indifference interval model**)

$$
\begin{aligned}
H_0: &\qquad \theta_i = \theta \in [\theta_0 - \delta, \theta_0 + \delta], \text{all } i \\
H_a(\tau): &\quad \begin{cases} \theta_i = \eta \in [\theta_0 - \delta, \theta_0 + \delta], \text{for } i \in (0, \tau] \\ \theta_i = \xi \notin [\theta_0 - \delta, \theta_0 + \delta], \text{for } i \in [\tau + 1, \infty), \end{cases}
\end{aligned}
$$

- This model represents a more appropriate option for **multi-products** processes with different parameters but same specification target.

Sequential change-point detection under composite hypotheses

Different change-point scenarios



Sequential change-point detection under composite hypotheses

Indifference interval

- The likelihood ratio for this model is

$$\Lambda_k(\theta, \eta, \xi; \mathbf{x}) = \max_{0 \le \tau \le k-1} \frac{\sup_{\eta \in \Theta_0} \prod_{i=1}^{\tau} f_\eta(x_i) \sup_{\xi \notin \Theta_0} \prod_{i=\tau+1}^{k} f_\xi(x_i)}{\sup_{\theta \in \Theta_0} \prod_{i=1}^{k} f_\theta(x_i)} = \max_{0 \le \tau \le k-1} \frac{\prod_{i=1}^{\tau} f_{\hat{\eta}}(x_i) \prod_{i=\tau+1}^{k} f_{\hat{\xi}}(x_i)}{\prod_{i=1}^{k} f_{\hat{\theta}}(x_i)}$$

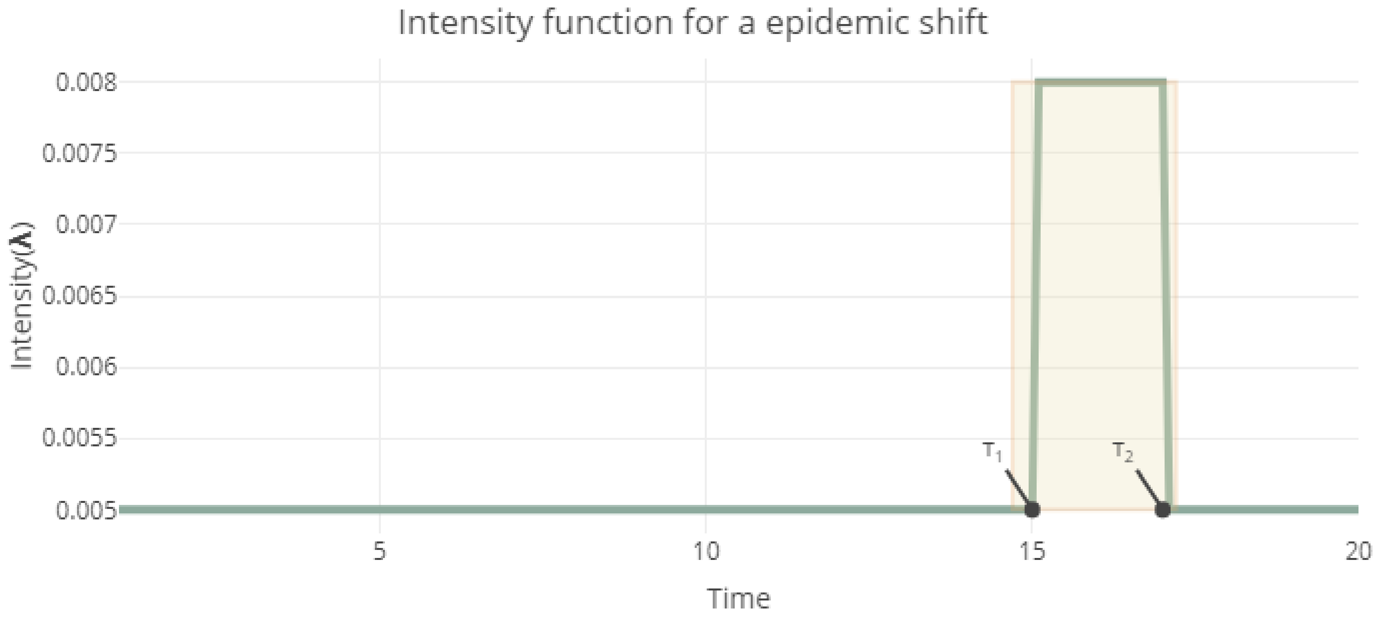
- It must be noted that in these conditions, when the null hypothesis parameters need to be estimated (i.e., composite null hypothesis), the numerator and denominator do not cancel out since $\hat{\theta} \neq \hat{\eta}$.

Sequential change-point detection under composite hypotheses

Epidemic shift model

The **epidemic shift model** represents a temporary change of parameters, which is a characteristic situation where a feedback controller is active.

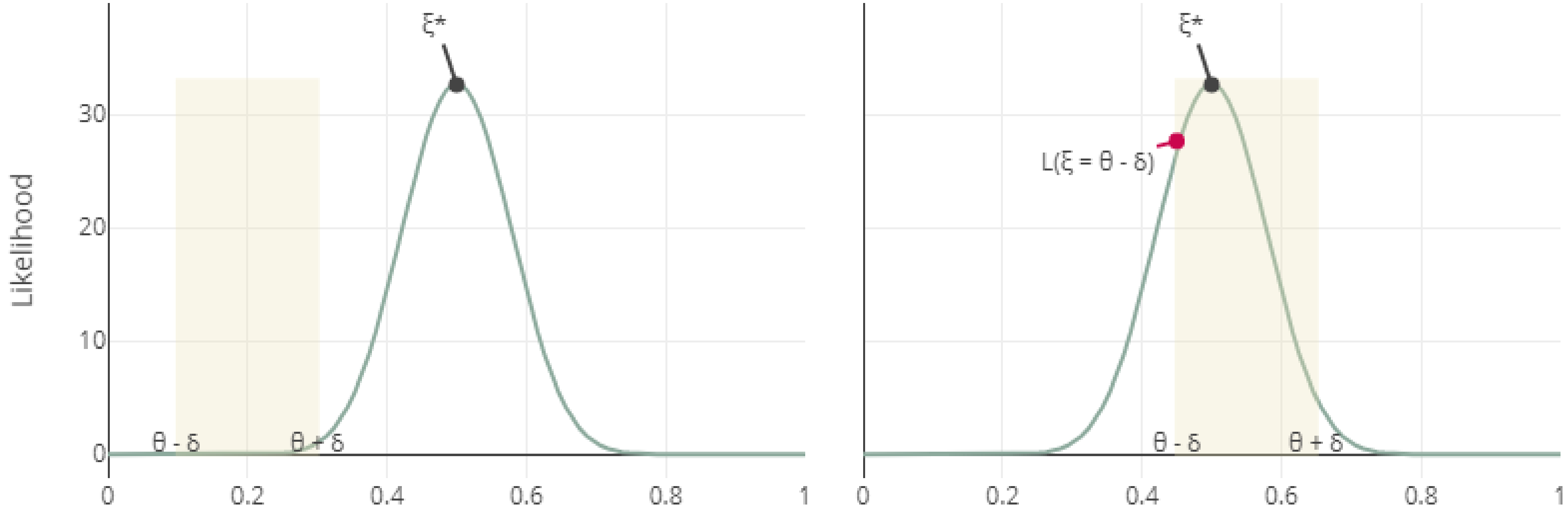
- A more generic version of the classical change-point problem since there are *multiple* change-points with $0 \leq \tau_1 \leq \tau_2$.



Parameter estimation

Indifference Interval Model

The parameters in the composite alternative hypothesis model are subjected to interval restrictions.



$$\hat{\xi}(\tau;\mathbf{x}) = \begin{cases} \hat{\xi}^* & \text{if } \hat{\xi}^* \notin \Theta_0 \\ \theta_0 - \delta & \text{if } \mathcal{L}(\xi = \theta_0 - \delta | \hat{\eta}, \hat{\theta}, \tau, \mathbf{x}) > \mathcal{L}(\xi = \theta_0 + \delta | \hat{\eta}, \hat{\theta}, \tau, \mathbf{x}) \\ \theta_0 + \delta & \text{otherwise} \end{cases}$$

where $\hat{\xi}^*$ is the unrestricted maximum likelihood estimator of the out-of-control parameter.

Parameter estimation

A side note

- In simple-hypotheses models the parameters might be also subjected to interval restriction (e.g., one-sided control charts)
- However, one needs to take it into account also in the definition of the log-likelihood statistic
  - In literature, the estimate in the log-likelihood statistics is often replaced by the *unrestricted* MLE.

Conclusions

- Monitoring schemes in chemical industry needs practical and tailored solutions to be effective and accepted;
- Generalized likelihood ratio-based control charts are known to outperform competitors in detecting wide range of parameters and to offer flexibility for more complex out-of-control scenarios;
- In this study we
  - Explored indifference interval and epidemic shift models as more appropriate out-of-control scenarios in the context of high-purity processes;
  - Highlight the complexity and pitfalls of composite hypothesis change-point detection.

References

1. A. Tartakovsky, I. Nikiforov, and M. Basseville. Sequential Analysis: Hypothesis Testing and Change-point Detection. Taylor & Francis Group, Boca Raton, 2015.
2. J.T. Chang and R.D. Fricker Jr. Detecting when a monotonically increasing mean has crossed a threshold.Journal of Quality Technology, 31(2):217–234, 1999.
3. E. Çınlar. Introduction to Stochastic Processes. Prentice-Hall, 1975.
4. A. Di Bucchianico, M. Husková, P. Kláasterecký, and W.R. van Zwet. Performance of control charts for specific alternative hypotheses. In J. Antoch, COMPSTAT 2004 Symposium, 903–910, Heidelberg, 2004.
5. W. Huang, M.R. Reynolds Jr., and S. Wang. A binomial GLR control chart for monitoring a proportion.Journal of Quality Technology, 44(3):192–208, 2012.
6. F. Kaminsky, J.C. Benneyan, R.D. Davis, and R.J. Burke. Statistical control charts based on a geometric distribution.Journal of Quality Technology, 24(2):64–69, Jan 1992.
7. A. KazemiNia, B.S. Gildeh, and Z. Abbasi Ganji. The design of geometric generalized likelihood ratio control chart.Quality and Reliability Engineering International, 34(5):953–965, 2018.
8. T.L. Lai. Information bounds and quick detection of parameter changes in stochastic systems. IEEE Transactions on Information Theory, 44(7):2917–2929, 1998.
9. J. Lee and J.Park. Poisson GLR control charts. The Korean Journal of Applied Statistics, 27(5):787–796, 2014.
10. J. Lee and W. H. Woodall. A note on GLR charts for monitoring count processes.Quality and Reliability Engineering International, 34(6):1041–1044, 2018.
11. J. Lee, Y. Peng, N. Wang, and M.R. Reynolds Jr. A GLR control chart for monitoring a multinomial process. Quality and Reliability Engineering International, 33(8):1773–1782, 2017.

Thank you for your attention

Do you have any question? Contact me at crizzo@dow.com!