Surrogate modeling based uncertainties analysis for the determination of safe and optimal operating conditions in batch reactors

Lujie Shi, Younes Aoues, Valeria Casson Moreno, Yankai Wang, Sébastien Leveneur

Computers and Chemical Engineering 192 (2025) 108909

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Surrogate modeling based uncertainties analysis for the determination of safe and optimal operating conditions in batch reactors

Lujie Shi [a,d], Younes Aoues [a], Valeria Casson Moreno [c], Yankai Wang [d], Sébastien Leveneur [b,*]

[a] INSA Rouen Normandie, Normandie Univ, LMN, EA3828, F-76000 Rouen, France
[b] INSA Rouen Normandie, Univ Rouen Normandie, Normandie Univ, LSPC UR 4704, F-76000 Rouen, France
[c] Department of Civil and Industrial Engineering, University of Pisa, 56126 Pisa, Italy
[d] School of Power and Energy, Northwestern Polytechnical University, 710129 Xi'an, China

ARTICLE INFO

Keywords:
GVL production
Thermal runaway risk
Surrogate model
Uncertainty propagation
Pareto optimality

ABSTRACT

In chemical process optimization, identifying conditions that balance production rate and thermal risks is crucial. This paper presents a surrogate-assisted optimization methodology that integrates parameters uncertainty, specifically focusing on synthesizing γ-valerolactone (GVL) in adiabatic and batch modes. A surrogate model was established to elucidate the relationships between input variables, production rate and risk index, which reduces the computational burden associated with complex differential equations. The Latin Hypercube Sampling method was employed to assess how uncertainties propagate through the processes. This study formulates a multi-objective optimization model that seeks to find a balance between the highest possible GVL production rate and the lowest probability of failure under deterministic and uncertain scenarios. The results in Pareto charts illustrate the possible operating conditions and determine the optimized initial conditions. This approach serves as a model for optimizing complex chemical processes, balancing production capacity and safety while considering uncertainty management.

1. Introduction

In the current fine chemical production, operations are predominantly conducted in batch and semi-batch modes. The latter is essentially for fast and exothermic reactions (Copelli et al., 2017, 2018). Over 90 % of industrial chemicals are still made in batch reactors (Ashe, 2022). The use of batch processes is favored for manufacturing higher-value products, such as specialty chemicals and pharmaceuticals, due to their versatility (Dimian et al., 2014). With rising costs of raw materials, energy, and increasingly stringent environmental regulations, chemical processes need to be as efficient, sustainable, and cost-effective as possible (Sun et al., 2020). However, 25 % of chemical accidents are caused by the loss of control over reaction heat, i.e., thermal runaway events (Kummer et al., 2021; Dakkoune et al., 2019). Therefore, understanding how to prevent uncontrolled reaction heat release and thereby ensure the safe production in chemical enterprises holds significant importance (Saada et al., 2015; Ni et al., 2016a, 2016b, 2016c).

Generally, there are four ways to prevent accidents (AICE, 2019): inherent safety, passive, active, and procedural strategies. Understanding safe operating conditions can be considered an inherent safety strategy, especially crucial for exothermic chemical systems that involve multiple reactions due to the intricate interplay between selectivity and safety constraints. In the context of batch or semi-batch reactions, each can have several operational variables, such as temperature, pressure, or heat transfer.

Optimizations to obtain safe and productive operating conditions for exothermic reaction systems is a critical challenge in chemical engineering (Soares et al., 2016; Lin et al., 2016), particularly due to the risks associated with thermal runaway (Copelli et al., 2013). Various studies have addressed this issue using different methodologies, emphasizing both safety and productivity. For instance, Kummer et al. (2020) and Zhu et al. (2022) studied the framework to control thermal runaway that consists of the runaway criterion and parameter identification. Casson Moreno et al. (2017) and Casson et al. (2012) explored the application of thermal runaway criteria to predict and manage the stability and safety of reactor operating conditions.

Optimization techniques have also been employed to navigate the trade-offs between competing objectives in chemical processes. Bortz et al. (2014) and Zora et al. (2021) applied Pareto sets for the chemical process, allowing decision-makers to evaluate trade-offs between different process goals, such as maximizing yield while minimizing risk.

* Corresponding author.
E-mail addresses: younes.aoues@insa-rouen.fr (Y. Aoues), sebastien.leveneur@insa-rouen.fr (S. Leveneur).

https://doi.org/10.1016/j.compchemeng.2024.108909
Received 15 May 2024; Received in revised form 30 September 2024; Accepted 27 October 2024
Available online 28 October 2024

Nomenclature

M	Parameters in kinetic model
X	Input parameters
Y	Output parameters
C	Concentration
K	Number of samplings in uncertainty analysis
m	Number of solutions in Pareto front
P	Probability
P_f	Probability of failure, i.e. probability of thermal runaway
P_{ft}	Failure threshold
Pr	Production rate
Ri	Risk index
R^2	Coefficient of determination
t	Time (s)

Subscript

i, j	Ordinal number

Glossary

Amb	Amberlite IR-120
BHP	Butyl Hydroxy Pentanoate
BL	Butyl Levulinate
GVL	γ -Valerolactone
HPA	Hydroxy Pentanoic Acid
NSGA- II	Non-dominated Sorting Genetic Algorithm II
LA	Levulinic Acid
LHS	Latin Hypercube Sampling
ODE	Ordinary Differential Equation
MSE	Mean Squared Error
Ru/C	Ruthenium on activated Carbon

Deb et al. (2004) utilized the NSGA-II algorithm to optimize the epoxy-polymerization process, revealing trade-offs between molecular weight, polydispersity, and reaction time, thus enhancing process understanding and efficiency.

Optimization under uncertainty is an even harder challenge in chemical engineering due to variability in measurements, disturbances, and raw material properties (Shi et al., 2023; Pishvaee et al., 2012). Addressing this challenge often involves complex calculations, which can be computationally intensive. Several strategies have been proposed to alleviate the computational burden of solving differential equations during optimization. One approach involves improving optimization algorithms to enhance efficiency, such as Bayesian optimization (Tadepalli et al., 2023; Manoj et al., 2023) or differential evolution (Chen et al., 2014). While another employs explicit relationships to replace dynamic models, simplifying the computational requirements (Zora et al., 2021). More recent methods (Pantula and Mitra, 2019, 2020; Miriyala et al., 2018) adopt data-driven methodologies that integrate machine learning techniques with optimization, improving solution accuracy. For instance, Miriyala et al. (Miriyala et al., 2016) compared various surrogate modeling approaches, showing that Sobol-assisted ANN can substantially reduce computational costs while maintaining high accuracy in optimizing complex chemical reaction networks.

Hydrogenation of levulinic acid or alkyl levulinates to produce γ-valerolactone (GVL) is an exothermic reaction system. GVL is a promising platform molecule (Horváth et al., 2008; Alonso et al., 2013; Yan et al., 2015), polar aprotic solvent (Baco et al., 2022; Chew et al., 2020) and is also a promising starting material for jet fuel (Bond et al., 2010; Han, 2017). GVL is also an essential intermediate for the production of poly(methyl methacrylate)-substitute (Manzer, 2004; Al-Naji et al., 2019). In fact, surrogate models are invaluable in areas such as optimization, sensitivity analysis, design space exploration, and any domain where direct evaluations of the actual system are prohibitive in terms of time or cost. However, less current research considers parameter uncertainty within the optimization of GVL processes.

In this study, we employed the kinetic model developed by Delgado et al. (2022), which delineates the hydrogenation kinetics of levulinic acid (LA) and butyl levulinate (BL) into GVL using Ru/C and Amberlite IR-120 in adiabatic mode. We introduce a bi-objective optimization method aimed at delineating the safety boundaries of the GVL production process. This method incorporates surrogate modeling and Latin Hypercube Sampling (LHS). By adopting surrogate models, complex differential equations are substituted with higher-order polynomials, significantly reducing the computational load required for iterative calculations during the optimization process. This approach also facilitates a faster computation of uncertainty propagation by LHS. By identifying more favorable operating conditions through stochastic programming optimization, our method enhances the efficiency of the exothermic process, thereby contributing to advancements in the field.

The paper is organized as follows: Section 2 gives the GVL synthesis process and kinetic model and introduces inputs and outputs. Section 3 proposes the optimization method, which is calculated based on the surrogate model and uncertainty. The results of the optimization are shown and interpreted in Section 4. Finally, a brief conclusion and future work are given in Section 5.

2. Reaction and kinetic model

2.1. Kinetic model and uncertainties

The GVL production process utilizes a Ru/C catalyst and Amberlite IR-120 (Amb). It unfolds in two primary stages, as illustrated in Fig. 1. The initial stage comprises the hydrogenation of LA and BL into intermediate products, specifically Hydroxy Pentanoic Acid (HPA) and Butyl Hydroxy Pentanoate (BHP). Subsequently, the second stage involves the cyclization of these intermediates to form GVL. For an in-depth understanding of the kinetic model associated with this production process, readers are encouraged to consult the study conducted by Delgado et al. (2022).

According to the previous studies of our group (Delgado et al., 2022; Wang et al., 2020), this system is a two-step reaction comprising a hydrogenation and cyclization step. Hydrogenation is an exothermic step that governs the reaction temperature.

Delgado et al. (2022) suggest that the most robust kinetic model for the process under study is the non-competitive Langmuir-Hinshelwood model. This model posits that levulinates (both BL and LA) and hydrogen are adsorbed at distinct sites. The ordinary differential equations, derived from the material and energy balances in a batch reactor, are a function of time (t). These equations are comprehensively detailed in Table 1.

In this reaction, the input parameters are represented as $X = (X_1,\ X_2, \cdots X_i,\ \cdots, X_I)(I = 7)$ and listed in Table 2. Considering the working conditions, the values of the input parameters are restricted in the lower and higher bounds range.

The parameters in the kinetic model are estimated based on laboratory kinetic experiments (Delgado et al., 2022). There are 18 parameters listed as $M = (M_1,\ M_2, \cdots M_i,\ \cdots, M_I)(I = 18)$ in Table 3.

2.2. Outputs from the model: risk index and production rate

It is essential for all chemical processes to reduce the thermal runaway risk and improve productivity. Hence, two outputs, including

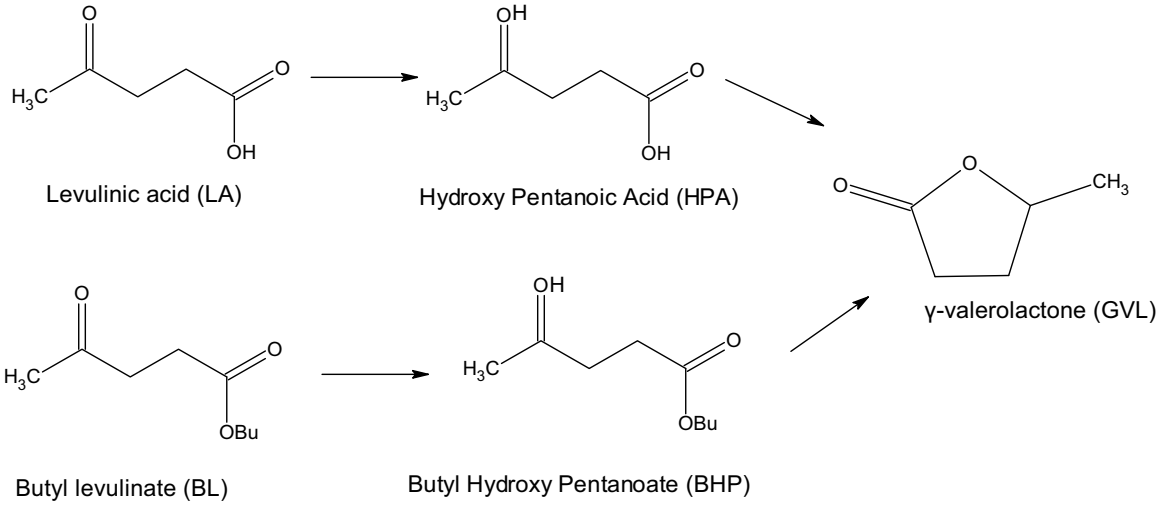

Fig. 1. Reaction steps of LA and BL to GVL over Ru/C and Amberlite IR-120.

Table 1
Mass and energy balance equations in the kinetic model.

Material Balance Equations on the single chemical species
$\frac{dC_{BL}}{dt} = -R_{BL\_hyd}$
$\frac{d[H_2]_{liq}}{dt} = k_{la}^{*}\left([H_2]^{*}_{liq} - [H_2]_{liq}\right) - R_{BL\_hyd} - R_{LA\_hyd}$
$\frac{dC_{BHP}}{dt} = R_{BL\_hyd} - R_{BHP\_noncat} - R_{BHP\_RuC} - R_{BHP\_SO3H} - R_{BHP\_diss}$
$\frac{dC_{buOH}}{dt} = R_{BHP\_noncat} + R_{BHP\_RuC} + R_{BHP\_SO3H} + R_{BHP\_diss}$
$\frac{dC_{LA}}{dt} = -R_{LA\_hyd}$
$\frac{dC_{HPA}}{dt} = R_{LA\_hyd} - R_{HPA\_noncat} - R_{HPA\_RuC} - R_{HPA\_SO3H} - R_{HPA\_diss}$
$\frac{dC_{water}}{dt} = R_{BL\_hyd} - R_{HPA\_noncat} + R_{HPA\_RuC} + R_{HPA\_SO3H} + R_{HPA\_diss}$
$\frac{dC_{GVL}}{dt} = R_{BHP\_noncat} + R_{BHP\_RuC} + R_{BHP\_SO3H} + R_{BHP\_diss} + R_{HPA\_noncat} + R_{HPA\_RuC} + R_{HPA\_SO3H} + R_{HPA\_diss}$
Energy Balance Equation
$\frac{dT_R}{dt} = \frac{\left(-R_{Hydrogenation}\cdot\Delta H_{R,Hydrogenation}\cdot V - R_{Cyclization}\cdot\Delta H_{R,Cyclization}\cdot V\right) + UA\cdot\left(T_j - T_R\right) + \alpha\cdot\left(T_{amb} - T_R\right)}{m_R\cdot C_{PR} + m_{insert}\cdot Cp_{insert} + m_{catalyst}\cdot Cp_{catalyst}}$

Table 2
Definitions of the input parameters.

NO.	Parameters	lower value	higher value	3*Standard deviation	Distribution	Definition	Unit
X_1	T_{j0}	90	140	1	Normal	Initial temperature of the jacket	K
X_2	P_{H2}	20	50	2	Normal	Initial pressure of H_2	bar
X_3	m_{LA0}	0.09	0.11	0.01	Normal	Initial mass of LA	kg
X_4	m_{BL0}	0.19	0.21	0.01	Normal	Initial mass of BL	kg
X_5	m_{GVL0}	0.09	0.11	0.01	Normal	Initial mass of GVL	kg
X_6	m_{Ru}	0.001	0.003	0.0003	Normal	Initial mass of Ru	kg
X_7	m_{Amb}	0.005	0.02	0.002	Normal	Initial mass of Amb	kg

production rate and risk index, were proposed in this study.

2.2.1. Production rate

An essential point for a chemical process is economic attractiveness. In this study, the production rate *Pr* is defined as the ratio of the production of GVL to the time to finish the reaction (t_GVL_{final}).

$$Pr = \frac{c\_GVL_{final} - c\_GVL_0}{t\_GVL_{final}} \quad (1)$$

A high yield difference and short reaction time is advantageous.

2.2.2. Risk index

Two parameters for the thermal runaway risk are defined: adiabatic temperature rise (ΔT_{ad}) and Time to the Maximum Rate (TMR_{ad}) (Pan et al., 2023). Generally, these parameters are determined by using a simplified kinetic model, i.e., zero-order approach in adiabatic conditions (Vernières-Hassimi et al., 2017). ΔT_{ad} is the difference between the maximum and initial reaction temperature and characterizes the severity of the risk. TMR_{ad} defines the time to reach the maximum temperature ratio and characterizes the probability of the thermal runaway risk (Stoessel, 2020), as explained in Appendix A

This paper applies a power function to calculate the risk index (*Ri*) by

Table 3
Definitions of the kinetic model parameters.

NO.	Parameter	Estimate	Unit
M_1	$k_{BL\_hyd}(T_{Ref})$	3.02E-06	$m^3.mol^{-1}.s^{-1}.kg.dryRuC^{-1}$
M_2	Ea_{BL_hyd}	3.62E+04	$J.mol^{-1}$
M_3	$k'_{BHP\_cat\_Amb}(T_{Ref})$	4.36E-05	$s^{-1}.kg.dryAmb^{-1}$
M_4	$k_{BHP\_noncat}(T_{Ref})$	5.93E-05	s^{-1}
M_5	Ea_{BHP_noncat}	7.78E+04	$J.mol^{-1}$
M_6	$k_{LA\_hyd}(T_{Ref})$	7.75E-06	$m^3.mol^{-1}.s^{-1}.kg.dryRuC^{-1}$
M_7	Ea_{LA_hyd}	4.61E+04	$J.mol^{-1}$
M_8	$K_{LA^\wedge}$	1.69E-03	$m^3.mol^{-1}$
M_9	$k'_{HPA\_cat\_Amb}(T_{Ref})$	4.79E-04	$s^{-1}.kg.dryAmb^{-1}$
M_{10}	$k_{HPA\_noncat}(T_{Ref})$	1.25E-06	s^{-1}
M_{11}	Ea_{HPA_noncat}	4.15E+05	$J.mol^{-1}$
M_{12}	$k_{BHP\_RuC}(T_{Ref})$	2.41E-05	$s^{-1}.kg.dryRuC^{-1}$
M_{13}	$k_{HPA\_RuC}(T_{Ref})$	5.74E-05	$s^{-1}.kg.dryRuC^{-1}$
M_{14}	Kc	1.59E-04	$m^3.mol^{-1}$
M_{15}	$k_{BHP\_diss}(T_{Ref})$	1.69E-06	$m^3.mol^{-1}.s^{-1}$
M_{16}	Ea_{BHP_diss}	1.09E+05	$J.mol^{-1}$
M_{17}	$k_{HPA\_diss}(T_{Ref})$	4.73E-06	$m^3.mol^{-1}.s^{-1}$
M_{18}	Ea_{HPA_diss}	6.70E+04	$J.mol^{-1}$

ΔT_{ad}, TMR_{ad} values:

$$Ri = \left(0.2449 \times (\Delta T_{ad})^{0.4372}\right) \times \left(-1.4772 \times (TMR_{ad})^{0.2894} + 7.2398\right) \quad (2)$$

Eq. (2) comes from fitting the classical risk matrix (see Appendix A.1), which provides discrete risk values. Generally, $Ri = 5$ is the boundary between acceptable and moderate risk, and $Ri = 10$ is the boundary between a moderate and non-acceptable risk.

In summary, the kinetic model with parameters can be expressed as:

$$\begin{aligned} &Y = f(\mathbf{X}, \mathbf{M}, t) \\ &\mathrm{X} = (X_1,\ X_2, \cdots, X_i,\ \cdots, X_I)\ (I = 7) \\ &\mathrm{M} = (M_1,\ M_2, \cdots, M_i,\ \cdots, M_I)(I = 18) \\ &\mathrm{Y} = (Pr,\ Ri) \end{aligned} \quad (3)$$

where, **X** and **M** consist of initial parameters (as shown in Table 2) and model parameters (as shown in Table 3), respectively. t is reaction time. The outputs **Y** contains production rate (Pr) and risk index (Ri).

3. Optimization method

This section delves into optimizing the GVL production, detailing the surrogate model, uncertainty quantification, and optimization model employed. Through different case studies, we seek to find the best operational parameters and conditions that maximize product yield while maintaining a safe process.

The proposed method is based on three steps, starting with the kinetic model, as shown in Fig. 2.

This framework commences with a kinetic model that serves as the foundational basis. In this model, the input variables $X = (X_1,\ X_2, \cdots, X_i,\ \cdots, X_I)\ (I = 7)$ are processed through a function $\mathbf{Y} = f(\mathbf{X}, \mathbf{M}, t)$ to yield outputs $Y = (Pr,\ Ri)$. These outputs represent the product rate and risk index, respectively.

Subsequently, a surrogate model was utilized to approximate the Ordinary Differential Equations (ODEs) from the kinetic model. The process initiated with Latin Hypercube Sampling (LHS) to facilitate efficient sampling from the predefined ranges of initial parameters. This was followed by constructing a model based on polynomial regression, accompanied by a comprehensive training and testing regimen. The latter was implemented to validate the predictive capabilities of the surrogate model, ensuring its accuracy and reliability in simulating the response of the kinetic model.

In parallel, an uncertainty analysis was conducted to assess the reliability of the model predictions. This includes the evaluation of input aleatory uncertainty from initial condition parameters. Sampling simulations are key in quantifying these uncertainties, leading to a comprehensive understanding of the output variability.

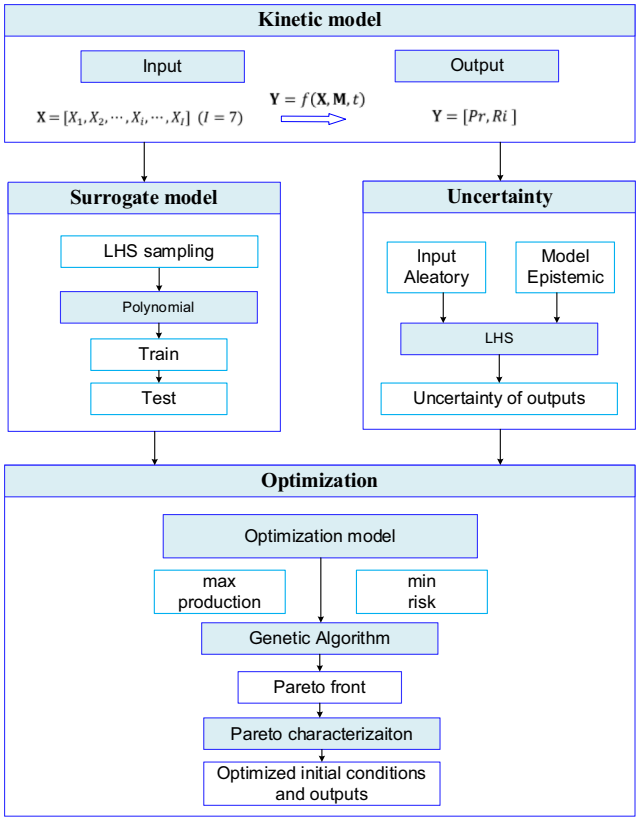

Fig. 2. Illustration of the optimization study with uncertainty and surrogate model.

Finally, an optimization model was formulated to maximize production and minimize risk. This optimization model leverages the insights from the surrogate model and uncertainty analysis to find the optimized input parameters. When addressing the optimization model, a genetic algorithm was employed to identify the Pareto front. The Pareto front presents the solution space to delineate optimal trade-offs between production and risk. The optimal initial conditions were selected for the GVL process with a low probability of failure.

This strategic framework is designed to systematically improve GVL synthesis, and each part is described in the following sections. Generally, this approach can be applied to any chemical process with similar characteristics.

3.1. Surrogate modeling method

The kinetic model, made of complex ODEs, presents significant computational challenges due to its extensive time requirements for solution (Leveneur et al., 2012). In statistical simulations such as Monte Carlo methods, where the accuracy of results is directly tied to the volume of data processed. Typically, $\left[10^{n+2}, 10^{n+3}\right]$ samples (Andrieu-Renaud et al., 2004) are needed to compute accurately the probability of failure of 10^{-n}. In the optimization-solving process, the computational workload escalates exponentially.

In light of this, the use of surrogate models becomes increasingly important. In this study, surrogate models were typically constructed by exploring the input-output relationships of the kinetic model. A schematic mapping between inputs X and outputs Y is illustrated in Fig. 3. which consists of two parts, namely the 7 inputs and 2 outputs.

The main purpose of using surrogate models for optimization is to

Fig. 3. Input parameters and output parameters for the GVL synthesis process.

train a sufficiently accurate approximation to the true kinetic model. This study adopted a fourth-order polynomial for the surrogate regression model to achieve a trade-off between computational efficiency and kinetic model fidelity.

The choice of a fourth-order polynomial is motivated by its robust capacity for encapsulating the nonlinear relationships inherent to the ODEs involved in GVL synthesis. The coefficients of this polynomial are computed through a disciplined training regimen, using the Ordinary Least Squares method to ensure that the surrogate model is well-adjusted to the underlying data obtained from the original kinetic model.

3.2. Uncertainty analysis

Two key considerations were made in the uncertainty analysis: (1) We employed the LHS method for stratified sampling, ensuring comprehensive coverage of the parameter space. (2) Based on prior laboratory measurements and empirical studies (Delgado et al., 2022), we assumed the parameters follow a normal distribution with corresponding means and standard deviations. Aleatory uncertainty, due to measurement errors or parameter variations, is represented by normally distributed input parameters. The specific distributions, including the mean and standard deviation for the seven initial parameters, are detailed in Table 2.

The 7 parameters are represented as $X = [X_1, X_2, \cdots X_i, \cdots, X_I]$ $(I = 7)$. By applying the LHS sampling method (Shi et al., 2023), the sampling number of simulations is set to K in this study. The sampled data are:

$$\overline{\mathbf{X}} = [\overline{X_1}, \overline{X_2}, \cdots \overline{X_i}, \cdots, \overline{X_I}] = \begin{bmatrix} x_1^{(1)} & \cdots & x_i^{(1)} & \cdots & x_I^{(1)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(k)} & \cdots & x_i^{(k)} & \cdots & x_I^{(k)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(K)} & \cdots & x_i^{(K)} & \cdots & x_I^{(K)} \end{bmatrix} \tag{4}$$

where $\overline{\mathbf{X}}$ is sampled data. The subscript i is the ordinal number for initial parameters and the superscript k donates k^{th} sampling by LHS for uncertainty. This allows us to propagate the input uncertainties through the kinetic or the surrogate models, thereby generating a distribution of the output parameters (risk indicator and production rate) as $\overline{\mathbf{Y}}$:

$$\overline{\mathbf{Y}} = [\overline{Pr}, \overline{Ri}] = \begin{bmatrix} Pr^{(1)} & Ri^{(1)} \\ & \vdots \\ Pr^{(k)} & Ri^{(k)} \\ & \vdots \\ Pr^{(K)} & Ri^{(K)} \end{bmatrix} \# \tag{5}$$

Through this probabilistic sampling, we can assess the uncertainty in model outputs, which is crucial for understanding the probability of thermal runaway risk in our optimizations.

3.3. Optimization models

In the quest for an optimized GVL synthesis process, this study builds a multi-objective optimization to maximize production yield and minimize operational risk simultaneously. This sophisticated optimization paradigm falls under the auspices of Pareto optimization. In the operational environment of GVL synthesis, this translates into a scenario where the interplay between yield maximization and probability of failure minimization is delicately balanced.

The primary objective in optimizing this GVL production process is to maximize production. The constraints are created based on the risk index to keep the risk in the acceptable zone. Hence, the first optimization model is defined as following:

$$\begin{aligned} &\max Pr = f_1(\mathbf{X}, \mathbf{M}, t) \\ &\min Ri = f_2(\mathbf{X}, \mathbf{M}, t) \\ &s.t.\ \boldsymbol{X}_l \le \boldsymbol{X}_i \le \boldsymbol{X}_u \end{aligned} \tag{6}$$

where the $\mathbf{X}$ is the decision variable, $\mathbf{M}$ is the kinetic model parameters, and t is the time parameter. In the constraints, the lower bound $\boldsymbol{X}_l$ and upper bound $\boldsymbol{X}_u$ are listed in Table 2.

Moreover, it is necessary to consider the aleatory uncertainty from input parameters $\mathbf{X}$, which are seen as random variables that are defined by mean and standard deviation $(\boldsymbol{\mu}_\mathbf{X},\ \boldsymbol{\sigma}_\mathbf{X})$. The second optimization model considering input uncertainty is defined:

$$\begin{aligned} &\max Pr = f_1(\boldsymbol{\mu}_\mathbf{X}, \mathbf{M}, t) \\ &\min P_f = f_2(\boldsymbol{\mu}_\mathbf{X}, \boldsymbol{\sigma}_\mathbf{X}, \mathbf{M}, t) \# \\ &s.t.\ \boldsymbol{X}_l \le \boldsymbol{\mu}_\mathbf{X} \le \boldsymbol{X}_u \end{aligned} \tag{7}$$

In this equation, the $\boldsymbol{\mu}_\mathbf{X}$ is decision variable, $\boldsymbol{\sigma}_\mathbf{X}$ is fixed as shown in Table 2. By using LHS uncertainty method in Section 3.3, $P_f = P(\overline{\boldsymbol{Ri}} > P_{ft})$. P stands for probability. P_f is probability of failure, i.e. the probability of having a thermal runaway, meant as the probability of Ri values greater than the threshold P_{ft}, as reported in Appendix A

In this paper, uncertainty was treated as random variables, and solving these two optimization models delineates the stochastic programming and the Pareto optimality. A decision variable vector $\boldsymbol{X}^*$ is Pareto optimal if there is no other vector $\boldsymbol{X}$ such that $f_1(\boldsymbol{X}, \mathbf{M}, t) \ge f_1(\boldsymbol{X}^*, \mathbf{M}, t)$ (higher production rate) and $f_2(\boldsymbol{X}, \mathbf{M}, t) \le f_2(\boldsymbol{X}^*, \mathbf{M}, t)$ (lower risk index), with at least one strict inequality for one objective.

All the Pareto-optimized solutions $\boldsymbol{X}^*$ made up the Pareto front $\boldsymbol{\Phi}^*$, which represents the set of all non-dominated solutions, providing a spectrum of optimal trade-offs between the dual objectives of productivity and probability of failure. By employing this strategy, we can give decision-makers diverse options, allowing for informed choices that align with specific operational priorities and risk appetites.

4. Results and discussion

4.1. Surrogate model and validation

The Latin Hypercube Sampling (LHS) was applied to train the surrogate model to generate a diverse and representative training dataset. Drawing from a 7-dimensional input space, LHS was used to produce 10^4 samples. This stratified sampling method ensured that each of the 7 input parameters was uniformly sampled across its entire range. We incorporated a k-fold cross-validation approach with $k = 5$ to ensure robustness and avoid overfitting. In this technique, the dataset was partitioned into 5 equally sized folds. The model was trained five times, using 4 of the folds for training and the remaining fold for validation. The training process (Pedregosa et al., 2011) took almost 22 mins for this model on a workstation with an Intel(R) Xeon(R) CPU E5–2660 CPU, 256 GB RAM.

To assess the performance of our surrogate model, we used two widely recognized statistical metrics: Mean Squared Error (MSE) and Coefficient of Determination (R^2). In the following, y represents the simulated data points and $\hat{y}$ denotes the predicted data points from the surrogate model, n is the data number.

The MSE measures the average squared difference between the simulated and predicted values. It is given by the formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{8}$$

A lower MSE indicates a better fit of the surrogate model to the simulated data.

The R^2 value captures the proportion of variance in the simulated dataset that is predictable from the predicted dataset. The formula for R^2 is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - E(y))^2} \tag{9}$$

A R^2 value close to 1 suggests that the model explains a large portion of the variance in the dependent variable, whereas a R^2 value close to 0 indicates that the model does not explain much of the variance.

MSE quantifies the difference between the simulated value and the predicted value, R^2 values of the K-S tests. In the surrogate model adopted in this study, both isothermal and adiabatic reaction conditions were taken into consideration. In isothermal conditions, the risk index *Ri* equals zero; thus, only results obtained under adiabatic conditions are presented. The results of each training are listed in Table 4.

From Table 4, we choose the model with minimum MSE as the final regression model. The correlation between the predicted and the simulated values are shown in Fig. 4(a) and (b), respectively.

Fig. 4 shows the correlation plots proving that the predicted values are mostly conformed to the simulated values of the outputs (Pr and Ri). A large R^2 (close to 1) implies that the distributions are not statistically different. Hence, the polynomial fit in this study has strived to construct a reliable and robust regression model.

Moreover, a comparison was conducted between the computational times of the kinetic model and its surrogate counterpart. In this comparison, the ODEs for the material and energy balances were solved out by using the solver "solve_ivp" (Virtanen et al., 2020) based on Backward Differentiation Formulas (BDF). For a total of 10^4 runs, the kinetic model required a substantial 13.92 min. In contrast, the surrogate model demonstrated efficiency, completing the same number of runs in a mere 0.11 s. This significant reduction in computational time underscores the advantages of replacing computationally expensive ODE simulations of kinetic systems with a cheaper and simpler surrogate model.

The good accuracy and less computing time of the trained surrogate model encourage us to use it to predict the output parameters of the kinetic model in optimization design. As we delve deeper into the results and insights derived from this analysis, it becomes evident how surrogate models can significantly enhance our ability to understand and predict complex systems.

4.2. Influence of the uncertainty

To investigate the variability and assess the robustness of *Pr* and *Ri* under different scenarios, we employed the uncertainty propagation approach. The simulation model was parameterized with input values $X = [100,\ 22,\ 0.1,\ 0.2,\ 0.1,\ 0.0015,\ 0.01]$. These inputs are subject to the normal distribution and the standard deviations are listed in Table 2. To capture the uncertainty and variability in these parameters, we conducted $K = 5000$ simulation runs. This extensive simulation effort provides a robust statistical foundation to evaluate the potential range and distribution of both the production rate and the risk index.

Table 4
The results of 5-fold test.

NO.	MSE	R^2 (Pr)	R^2(Ri)
1	1.829e-6	0.9991	0.9998
2	1.872e-6	0.9999	0.9997
3	1.879e-6	0.9991	0.9998
4	1.921e-6	0.9990	0.9909
5	1.958e-6	0.9991	0.9997

Analysis of the simulation data yielded the scatter plot depicted in Fig. 5, which illustrates the stochastic relationship between the production rate and the risk index.

Fig. 5 presents a scatter plot with marginal histograms illustrating the relationship between production rate and risk index. Each point on the scatter plot represents an observed pair of production rate and the risk index. A discernible upward trend suggests a positive correlation between increased production and higher risk levels. The marginal histograms reveal the underlying distributions of each variable, with the production rate exhibiting a roughly normal distribution and the risk index showing a propensity for lower values. These patterns indicate operational dynamics where risk considerations become more pronounced at higher production rates.

4.3. Pareto optimization results and decision analysis

This research employs the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to derive optimal solutions. Within the NSGA-II framework, we have chosen 100 particles, leading to 100 distinct solutions. These outcomes collectively form the Pareto front, illustrating the trade-offs among competing objectives without favoring any single objective. Specifically, this analysis concentrates on the adiabatic mode, and the threshold value of the failure probabilities are selected as $P_{ft} = 10,\ 9.8,$ and 9.6, respectively (details in Appendix A).

Fig. 6 presents the Pareto optimization results; the horizontal axis represents the production rate, while the vertical axis represents the risk index. It illustrates the total 100 optimal solutions where no single objective can be improved without compromising the other. The spread of points along the Pareto front indicates the range of possible solutions, highlighting the diversity in optimal strategies under consideration. This visualization aids decision-making by clearly delineating the best possible compromises between the competing objectives.

In addition, it can be observed that the production rate is more concentrated in the area below $Pr = 1.0$. Combining the Pareto front analysis with the histograms presented in Fig. 6, it becomes apparent that within the range of $Pr \leq 1.0$, the rate of increase in risk outpaces that of the production yield. Conversely, when $Pr > 1.0$, the growth rates of both parameters tend to converge. Consequently, decision-makers can weigh both production yield and risk to select the optimal reaction conditions. In this deterministic case, the best initial conditions can be found as $\boldsymbol{X}^* = [139.98,\ 20.03,\ 9.01 \times 10^{-2},\ 2.05 \times 10^{-1},\ 1.10 \times 10^{-1},\ 2.01 \times 10^{-3},\ 1.98 \times 10^{-2}]$, the outputs are $Pr = 1.39$, and $Ri = 9.99$.

Furthermore, we examined the results of Pareto optimization under parameter uncertainty, as depicted in Fig. 7. For this analysis, the uncertainty sampling size was set to $K = 1 \times 10^6$. Three distinct Pareto fronts emerge, representing optimal solutions at failure probabilities of 10 %, 1 %, and 0.1 %, respectively.

Fig. 7 presents a scatter plot that illustrates the relationship between the production rate and the probability of failure for the GVL production system. There are three distinct series represented by different colors and symbols, the blue circles ($P_{ft} = 10$), the orange stars ($P_{ft} = 9.8$), and the green crosses ($P_{ft} = 9.6$). Each series indicates that the probability of failure is low at lower production rates, increases rapidly at intermediate rates, and then plateaus at higher rates, suggesting that once a certain production threshold is surpassed, the probability of failure does not increase significantly. The isolated red dots could be outliers or special observations that were recorded but fall outside the expected trend.

After obtaining the optimized results of these two cases, a comparative analysis of the relationship between production rate and risk index under deterministic and uncertain scenarios, as depicted in Fig. 8.

Select the optimized results with $P_f < 0.1\%$, the optimized initial

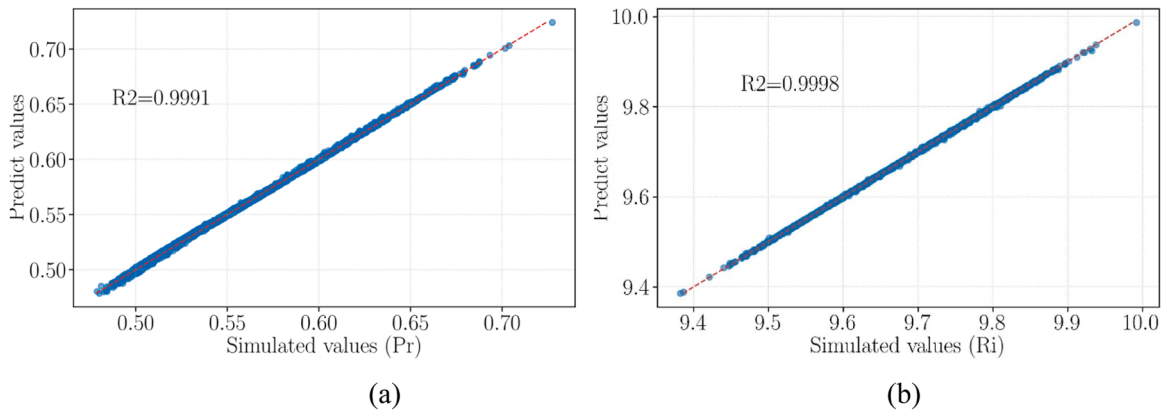

Fig. 4. Correlation plot between the simulated and the predicted values of (a) Production rate and (b) Risk index.

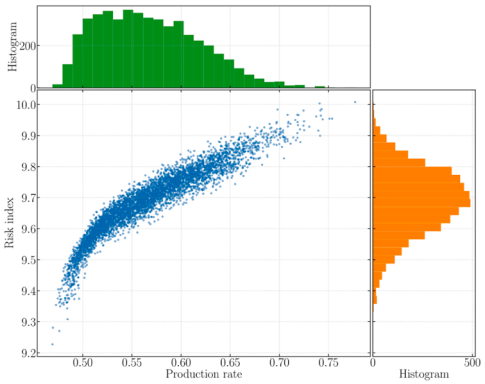

Fig. 5. Scatter of *Pr* and *Ri* in the adiabatic conditions.

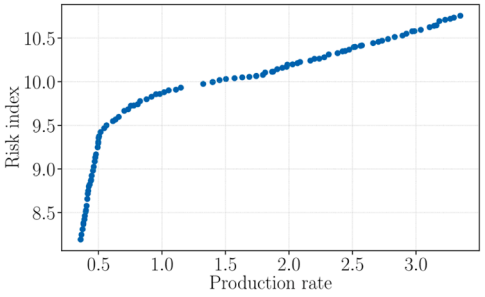

Fig. 6. Pareto front in the deterministic case.

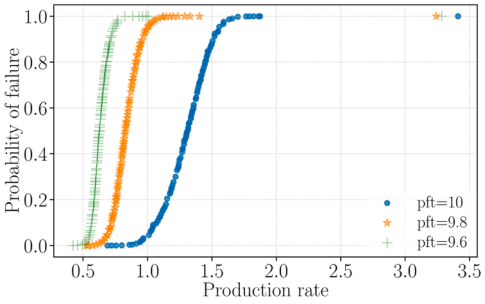

Fig. 7. Pareto front in the uncertain case.

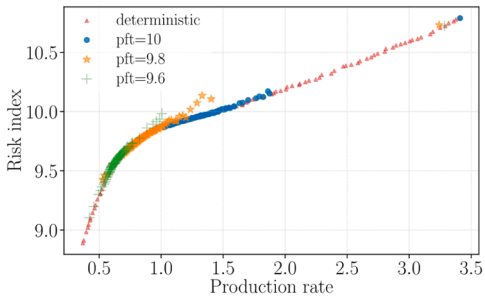

Fig. 8. Pareto fronts in the deterministic and uncertain cases.

conditions ($\boldsymbol{X}^*$) and outputs (*Pr* and *Ri*) can be determined as listed in Table 5. As evidenced by Table 5, it is observed that with demands on failure threshold reduction, there necessitates a decrease in the initial temperature (T_{j0}) of the reaction. The risk index range is reduced from [9.6551, 9.8019] to [9.1042, 9.3337], while the production rate significantly decreases from [0.6890, 0.8828] to [0.4224, 0.5140]. This adjustment correspondingly leads to a diminution in both production rate (*Pr*) and associated risk index (*Ri*). From Table 5, one can observe:

- A lower T_{j0} results in decreased *Pr* and *Ri*. Conversely, lower P_{ft} is achieved with a lower T_{j0} (i.e., R_i), which in turn leads to reduced P_r.
- A decrease in P_{H2} leads to an increase in both *Pr* and *Ri*, indicating that higher P_{H2} results in lower P_{ft}.
- An increase in the initial concentration of LA (i.e., m_{LA0}) corresponds to a decrease in both productivity and associated risk.
- The initial concentrations of BL and GVL (i.e., m_{BL0} and m_{GVL0}) do not affect productivity or risk levels.
- An increase in the concentration of Ru catalyst enhances productivity and raises the risk index. Conversely, lower P_{ft} is associated with a lower m_{Ru}. In the case of Am, no significant influence on *Pr* or *Ri* is observed.

Through Figs. 7, 8 and Table 5, decision-makers can see the potential paths for finding a balance between two objectives and understand the uncertainties associated with input parameters, providing a scientific basis for final decisions.

5. Conclusion

This study introduces the novel application of ML techniques to process safety, specifically targeting chemical reaction hazards and thermal runaway detection in batch exothermic reactions. By introducing a bi-objective optimization method that combines surrogate modeling and LHS sampling, we navigate the intricate balance between

Table 5
The optimized initial conditions and outputs with $P_f < 0.1$.

Probability of Failure (i.e. probability of thermal runaway)	Initial conditions							Outputs	
	T_{j0}	P_{H2}	m_{LA0}(× 10^{-2})	m_{BL0}(× 10^{-1})	m_{GVL0}(× 10^{-1})	m_{Ru}(× 10^{-3})	m_{Amb}(× 10^{-2})	*Pr*	*Ri*
$P_{ft} = 10$	98.15	20.34	9.10	1.94	1.09	2.41	1.99	0.6890	9.6551
	98.52	20.09	9.09	1.90	1.10	2.55	2.00	0.7209	9.6840
	100.51	20.08	9.02	1.90	1.10	2.59	2.00	0.7558	9.7097
	105.16	20.34	9.02	1.94	1.10	2.41	2.00	0.8005	9.7496
	106.71	20.10	9.03	1.91	1.10	2.56	2.00	0.8603	9.7888
	105.53	20.09	9.01	1.90	1.10	2.80	1.99	0.8828	9.8019
$P_{ft} = 9.8$	90.70	20.60	9.05	2.02	0.91	1.73	1.89	0.5284	9.4252
	91.21	20.59	9.05	2.02	0.91	1.79	1.89	0.5422	9.4604
	90.71	20.19	9.05	1.90	1.08	2.21	1.99	0.5733	9.4680
	101.51	21.07	9.01	1.90	1.08	1.65	2.00	0.6014	9.5352
	104.63	20.72	9.02	1.90	1.09	1.61	2.00	0.6238	9.5640
$P_{ft} = 9.6$	91.35	20.47	9.26	1.92	1.08	1.20	1.98	0.4224	9.1042
	90.20	21.32	9.08	1.90	1.10	1.42	2.00	0.4573	9.1981
	92.39	20.47	9.00	1.90	1.08	1.60	2.00	0.4950	9.3001
	91.27	20.27	9.04	1.90	1.07	1.76	2.00	0.5140	9.3337

operational safety and productivity under different failure probabilities. The proposed approach provides a detailed application of the surrogate model and uncertainty propagation in the optimization of the chemical initial conditions.

This study focuses on the hydrogenation of levulinic acid (LA) and butyl levulinate (BL) into γ-valerolactone (GVL) in adiabatic batch conditions, a process characterized by its exothermic nature. We utilized the numerical kinetic model developed by Delgado et al. (2022) to analyze this reaction. For this specific process and kinetic model, seven input variables that influence the initial conditions, such as the jacket temperature, hydrogen pressure, masses of LA, BL, and GVL, along with the mass of the catalysts Ru/C and Amberlite IR-120, were assessed. These inputs were examined in relation to their effects on two critical outcomes: the production rate and the risk index.

Our findings underscore optimization techniques in identifying optimal conditions that minimize thermal risks while maximizing production yields in exothermic reactions. The surrogate-assisted optimization approach, enhanced by the precision of LHS in assessing uncertainty propagation, provides a robust framework for addressing the challenges posed by the nonlinear, time-dependent nature of chemical processes. The results, shown as the Pareto chart in our study, delineated a set of optimal conditions, revealing a trade-off between maximizing the production rate and minimizing the probability of failure. This approach allowed us to significantly reduce (almost 10^4 times) the computational time with respect to conventional kinetic modelling.

This assessment could serve as a guideline for determining the safe operating conditions in GVL production processes. Future work will continue to refine these optimization strategies, exploring broader applications and further integrating kinetic-parameter-free approaches to minimize reliance on expensive and time-consuming kinetic data acquisition.

CRediT authorship contribution statement

Lujie Shi: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Younes Aoues:** Methodology, Investigation, Funding acquisition. **Valeria Casson Moreno:** Writing – review & editing, Validation, Conceptualization. **Yankai Wang:** Conceptualization, Writing – review & editing. **Sébastien Leveneur:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was carried out in the framework of the ARBRE project (Risk Analysis to processes valorizing 2nd generation biomass and using Renewable energies), which is cofounded by European Union through the European Regional Development Fund (00130305) and by Normandy Region (21E05304).

Appendix A

A1 Exponential function fitting for continuous risk index

From the literature (Stoessel, 2020; Pan et al., 2023), a classical risk matrix is built by ΔT_{ad} and TMR_{ad} as shown in Fig. A1.

Level			Negligible	Medium	Critical	Catastrophic
	Parameter	ΔT	<50	50-200	200-400	>400
	TMR	Factor	1	2	3	4
Impossible	>100h	1	1	2	3	4
Remote	50-100h	2	2	4	6	8
Seldom	24-50h	3	3	6	9	12
Occasional	8-24h	4	4	8	12	16
Probable	1-8h	5	5	10	15	20
Frequency	<1h	6	6	12	18	24

Fig. A1. Classical risk matrix for thermal runaway.

The ΔT_{ad} and TMR_{ad} are divided by different ranges and have corresponding factors. The risk matrix's indicators are obtained by multiplying the two factors. Current design regulations indirectly address this limit-state by specifying a threshold value for the risk index. Then, the risk indexes are

divided into a non-acceptable risk zone (red), a moderate risk zone (blue), and an acceptable risk zone (green).

A drawback in this classical risk matrix is that the risk index can only take specific values and is not continuous. This risk matrix effectively assesses the thermal runaway risk in practice, but it may lead to numerical errors when using some statistical methods for numerical simulation. To avoid this problem, this paper applies a power function to fit ΔT_{ad}, TMR_{ad} values and factors.

$$Ri = \left(0.2449 \times (\Delta T_{ad})^{0.4372}\right) \times \left(-1.4772 \times (TMR_{ad})^{0.2894} + 7.2398\right) \tag{A1}$$

As Eq. (1) shows, two factors follow a power-law relationship with ΔT_{ad} and TMR_{ad} values separately. Moreover, the final risk index Ri is also obtained by multiplying the two factors, but the values are continuous. By plotting TMR_{ad}versus ΔT_{ad}, the comparison of risk indexes between the classical risk matrix and fitted method are shown in Fig. A2.

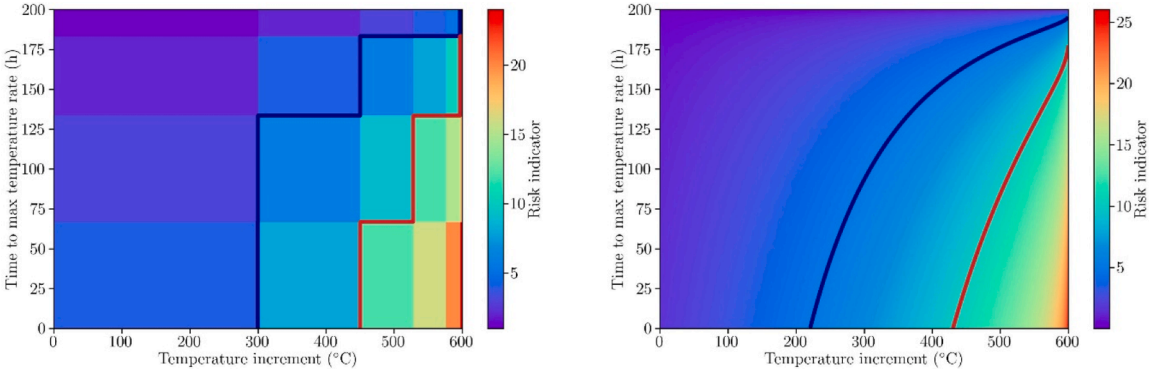

Fig. A2. Comparison of risk index between (a) classical risk matrix and (b) fitted method.

Within the same range of values, the risk indexes in Fig. A2(b) are smoother compared to those in Fig. A2(a), without any abrupt changes. The blue lines in both figures correspond to $Ri = 5$, the boundary for the acceptable and moderate risk zones. Similarly, the red lines correspond to $Ri = 10$, which is the boundary for a moderate and non-acceptable risk zone.

The newly fitted risk index (in Eq. (A1)) remains the same considerations and boundaries of risk assessment in chemical engineering processes. However, it provides more reasonable numerical values for more complex numerical analyses. It is preferable to have a low Ri, which can be achieved by minimizing ΔT_{ad} and maximizing TMR_{ad}.

Data availability

Data will be made available on request.

References

Al-Naji, M., Puértolas, B., Kumru, B., Cruz, D., Bäumel, M., Schmidt, B.V.K.J., Tarakina, N.V., Pérez-Ramírez, J., 2019. Sustainable Continuous Flow Valorization of γ-Valerolactone with Trioxane to α-Methylene-γ-Valerolactone over Basic Beta Zeolites. ChemSusChem. 12, 2628–2636. https://doi.org/10.1002/cssc.201900418.

Alonso, D.M., Wettstein, S.G., Dumesic, J.A., 2013. Gamma-valerolactone, a sustainable platform molecule derived from lignocellulosic biomass. Green. Chem. 15, 584. https://doi.org/10.1039/c3gc37065h.

American Institute of Chemical Engineers, 2019. Guidelines For Inherently Safer Chemical processes: a Life Cycle Approach, 3rd edition. John Wiley & Sons, Inc., Hoboken, NJ, USA : New York, NY. American Institute of Chemical Engineers.

Andrieu-Renaud, C., Sudret, B., Lemaire, M., 2004. The PHI2 method: a way to compute time-variant reliability. Reliab. Eng. Syst. Saf. 84, 75–86. https://doi.org/10.1016/j.ress.2003.10.005.

Ashe, R., 2022. Progress in flow chemistry. Chim. Oggi/Chem. Today 40.

Baco, S., Klinksiek, M., Ismail Bedawi Zakaria, R., Antonia Garcia-Hernandez, E., Mignot, M., Legros, J., Held, C., Casson Moreno, V., Leveneur, S., 2022. Solvent effect investigation on the acid-catalyzed esterification of levulinic acid by ethanol aided by a Linear Solvation Energy Relationship. Chem. Eng. Sci. 260, 117928. https://doi.org/10.1016/j.ces.2022.117928.

Bond, J.Q., Alonso, D.M., Wang, D., West, R.M., Dumesic, J.A., 2010. Integrated Catalytic Conversion of γ-Valerolactone to Liquid Alkenes for Transportation Fuels. Science (1979) 327, 1110–1114. https://doi.org/10.1126/science.1184362.

Bortz, M., Burger, J., Asprion, N., Blagov, S., Böttcher, R., Nowak, U., Scheithauer, A., Welke, R., Küfer, K.-H., Hasse, H., 2014. Multi-criteria optimization in chemical process design and decision support by navigation on Pareto sets. Comput. Chem. Eng. 60, 354–363. https://doi.org/10.1016/j.compchemeng.2013.09.015.

Casson, V., Lister, D.G., Milazzo, M.F., Maschio, G., 2012. Comparison of criteria for prediction of runaway reactions in the sulphuric acid catalyzed esterification of acetic anhydride and methanol. J. Loss. Prev. Process. Ind. 25, 209–217. https://doi.org/10.1016/j.jlp.2011.09.002.

Chen, X., Du, W., Qian, F., 2014. Multi-objective differential evolution with ranking-based mutation operator and its application in chemical process optimization. Chemometr. Intell. Lab. Syst. 136, 85–96. https://doi.org/10.1016/j.chemolab.2014.05.007.

Chew, A.K., Walker, T.W., Shen, Z., Demir, B., Witteman, L., Euclide, J., Huber, G.W., Dumesic, J.A., Van Lehn, R.C., 2020. Effect of mixed-solvent environments on the selectivity of acid-catalyzed dehydration reactions. ACS Catal. 10, 1679–1691. https://doi.org/10.1021/acscatal.9b03460.

Copelli, S., Barozzi, M., Maestri, F., Rota, R., 2018. Safe optimization of potentially runaway reactions: from fedbatch to continuous stirred tank type reactor. J. Loss. Prev. Process. Ind. 55, 289–302. https://doi.org/10.1016/j.jlp.2018.07.003.

Copelli, S., Barozzi, M., Maestri, F., Rota, R., 2017. Safe intensification of potentially runaway reactions: from semibatch to continuous processes. Chem. Eng. Trans. 57, 1687–1692. https://doi.org/10.3303/CET1757282.

Copelli, S., Derudi, M., Sala Cattaneo, C., Nano, G., Torretta, V., Rota, R., 2013. Classification and optimization of potentially runaway processes using topology tools. Comput. Chem. Eng. 56, 114–127. https://doi.org/10.1016/j.compchemeng.2013.05.012.

Dakkoune, A., Vernières-Hassimi, L., Leveneur, S., Lefebvre, D., Estel, L., 2019. Analysis of thermal runaway events in French chemical industry. J. Loss. Prev. Process. Ind. 62, 103938. https://doi.org/10.1016/j.jlp.2019.103938.

Deb, K., Mitra, K., Dewri, R., Majumdar, S., 2004. Towards a better understanding of the epoxy-polymerization process using multi-objective evolutionary computation. Chem. Eng. Sci. 59, 4261–4277. https://doi.org/10.1016/j.ces.2004.06.012.

Delgado, J., Vasquez Salcedo, W.N., Bronzetti, G., Casson Moreno, V., Mignot, M., Legros, J., Held, C., Grénman, H., Leveneur, S., 2022. Kinetic model assessment for the synthesis of gamma-valerolactone from n-butyl levulinate and levulinic acid hydrogenation over the synergy effect of dual catalysts Ru/C and Amberlite IR-120. Chem. Eng. J. 430, 133053. https://doi.org/10.1016/j.cej.2021.133053.

Dimian, A.C., Bildea, C.S., Kiss, A.A., 2014. Batch Processes. Computer Aided Chemical Engineering. Elsevier, pp. 449–488. https://doi.org/10.1016/B978-0-444-62700-1.00011-5.

Han, J., 2017. Integrated process for simultaneous production of jet fuel range alkenes and N -methylformanilide using biomass-derived gamma-valerolactone. J. Ind. Eng. Chem. 48, 173–179. https://doi.org/10.1016/j.jiec.2016.12.036.

Horváth, I.T., Mehdi, H., Fábos, V., Boda, L., Mika, L.T., 2008. Gamma-Valerolactone—A sustainable liquid for energy and carbon-based chemicals. Green. Chem. 10, 238–242. https://doi.org/10.1039/B712863K.

Kummer, A., Nagy, L., Varga, T., 2020. NMPC-based control scheme for a semi-batch reactor under parameter uncertainty. Comput. Chem. Eng. 141, 106998. https://doi.org/10.1016/j.compchemeng.2020.106998.

Kummer, A., Varga, T., 2021. What do we know already about reactor runaway? – A review. Process Saf. Environ. Protect. 147, 460–476. https://doi.org/10.1016/j.psep.2020.09.059.
Leveneur, S., Thönes, M., Hébert, J.-P., Taouk, B., Salmi, T., 2012. From kinetic study to thermal safety assessment: application to peroxyformic acid synthesis. Ind. Eng. Chem. Res. 51, 13999–14007. https://doi.org/10.1021/ie3017847.
Lin, C.-P., Li, J.-S., Tseng, J.-M., Mannan, M.S., 2016. Thermal runaway reaction for highly exothermic material in safe storage temperature. J. Loss. Prev. Process. Ind. 40, 259–265. https://doi.org/10.1016/j.jlp.2016.01.006.
Manoj, A., Miriyala, S.S., Mitra, K., 2023. Multi-objective optimization through a novel Bayesian approach for industrial manufacturing of Polyvinyl Acetate. Mater. Manuf. Process. 38, 1955–1963. https://doi.org/10.1080/10426914.2023.2195915.
Manzer, L.E., 2004. Catalytic synthesis of α-methylene-γ-valerolactone: a biomass-derived acrylic monomer. Appl. Catal. A: General 272, 249–256. https://doi.org/10.1016/j.apcata.2004.05.048.
Miriyala, S.S., Mittal, P., Majumdar, S., Mitra, K., 2016. Comparative study of surrogate approaches while optimizing computationally expensive reaction networks. Chem. Eng. Sci. 140, 44–61. https://doi.org/10.1016/j.ces.2015.09.030.
Miriyala, S.S., Subramanian, V.R., Mitra, K., 2018. TRANSFORM-ANN for online optimization of complex industrial processes: casting process as case study. Eur. J. Oper. Res. 264, 294–309. https://doi.org/10.1016/j.ejor.2017.05.026.
Casson Moreno, V., Russo, V., Tesser, R., Di Serio, M., Salzano, E., 2017. Thermal risk in semi-batch reactors: the epoxidation of soybean oil. Process Saf. Environ. Protect. 109, 529–537. https://doi.org/10.1016/j.psep.2017.05.001.
Ni, L., Jiang, J., Mebarki, A., Zhang, M., Dou, Z., Pensee, V., 2016a. Thermal risk in batch reactors: case of peracetic acid synthesis. J. Loss. Prev. Process. Ind. 39, 85–92. https://doi.org/10.1016/j.jlp.2015.11.015.
Ni, L., Mebarki, A., Jiang, J., Zhang, M., Dou, Z., 2016b. Semi-batch reactors: thermal runaway risk. J. Loss. Prev. Process. Ind. 43, 559–566. https://doi.org/10.1016/j.jlp.2016.07.024.
Ni, L., Mebarki, A., Jiang, J., Zhang, M., Pensee, V., Dou, Z., 2016c. Thermal risk in batch reactors: theoretical framework for runaway and accident. J. Loss. Prev. Process. Ind. 43, 75–82. https://doi.org/10.1016/j.jlp.2016.04.004.
Pan, Y., Ren, C., Wang, G., Wang, Y., Zhang, X., Jiang, J., Shu, C.-M., 2023. Thermal hazard evaluation for gamma-valerolactone production by using formic acid as hydrogen donor. J. Loss. Prev. Process. Ind. 81, 104951. https://doi.org/10.1016/j.jlp.2022.104951.
Pantula, P.D., Mitra, K., 2020. Towards efficient robust optimization using data based optimal segmentation of uncertain space. Reliab. Eng. Syst. Saf. 197, 106821. https://doi.org/10.1016/j.ress.2020.106821.
Pantula, P.D., Mitra, K., 2019. A data-driven approach towards finding closer estimates of optimal solutions under uncertainty for an energy efficient steel casting process. Energy 189, 116253. https://doi.org/10.1016/j.energy.2019.116253.
Pedregosa, F., Varoquaux, G., Gramfort, A., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Pishvaee, M.S., Torabi, S.A., Razmi, J., 2012. Credibility-based fuzzy mathematical programming model for green logistics design under uncertainty. Comput. Ind. Eng. 62, 624–632. https://doi.org/10.1016/j.cie.2011.11.028.
Saada, R., Patel, D., Saha, B., 2015. Causes and consequences of thermal runaway incidents—Will they ever be avoided? Process Saf. Environ. Protect. 97, 109–115. https://doi.org/10.1016/j.psep.2015.02.005.
Shi, L., Delgado, J., Aoues, Y., Leveneur, S., 2023. Thermal risk assessment with mixed uncertainty propagation in gamma-valerolactone production. In: 2023 7th International Conference on System Reliability and Safety (ICSRS). Presented at the 2023 7th International Conference on System Reliability and Safety (ICSRS). IEEE, Bologna, Italy, pp. 350–354. https://doi.org/10.1109/ICSRS59833.2023.10381043.
Soares, R.M., Pinto, J.C., Secchi, A.R., 2016. An optimal control-based safety system for cost efficient risk management of chemical processes. Comput. Chem. Eng. 91, 471–484. https://doi.org/10.1016/j.compchemeng.2016.04.029.
Stoessel, F., 2020. Thermal Safety of Chemical Processes: Risk Assessment and Process Design, 1st ed. Wiley. https://doi.org/10.1002/9783527696918.
Sun, Q., Jiang, L., Li, M., Sun, J., 2020. Assessment on thermal hazards of reactive chemicals in industry: state of the Art and perspectives. Prog. Energy Combust. Sci. 78, 100832. https://doi.org/10.1016/j.pecs.2020.100832.
Tadepalli, A., Pujari, K.N., Mitra, K., 2023. A crystallization case study toward optimization of expensive to evaluate mathematical models using Bayesian approach. Mater. Manuf. Process. 38, 2127–2134. https://doi.org/10.1080/10426914.2023.2238051.
Vernières-Hassimi, L., Dakkoune, A., Abdelouahed, L., Estel, L., Leveneur, S., 2017. Zero-order versus intrinsic kinetics for the determination of the time to maximum rate under adiabatic conditions (TMR_{ad}): application to the decomposition of hydrogen peroxide. Ind. Eng. Chem. Res. 56, 13040–13049. https://doi.org/10.1021/acs.iecr.7b01291.
Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2.
Wang, Y., Plazl, I., Vernières-Hassimi, L., Leveneur, S., 2020. From calorimetry to thermal risk assessment: gamma-Valerolactone production from the hydrogenation of alkyl levulinates. Process Saf. Environ. Protect. 144, 32–41. https://doi.org/10.1016/j.psep.2020.07.017.
Yan, K., Yang, Y., Chai, J., Lu, Y., 2015. Catalytic reactions of gamma-valerolactone: a platform to fuels and value-added chemicals. Appl. Catal. B: Environ. 179, 292–304. https://doi.org/10.1016/j.apcatb.2015.04.030.
Zhu, J., Hao, L., Bai, W., Zhu, Z., Zhang, B., Wei, H., 2022. A design framework for optimized economic and inherently safe operation conditions for isoperibolic semi-batch reactors. Process Saf. Environ. Protect. 168, 166–179. https://doi.org/10.1016/j.psep.2022.09.048.
Zora, N., Rigaux, T., Buvat, J.-C., Lefebvre, D., Leveneur, S., 2021. Influence assessment of inlet parameters on thermal risk and productivity: application to the epoxidation of vegetable oils. J. Loss. Prev. Process. Ind. 72, 104551. https://doi.org/10.1016/j.jlp.2021.104551.