

# Design of experiment for model discrimination: SFE case, part I

Oliwer Sliczniuk<sup>a,\*</sup>

<sup>a</sup>Aalto University, School of Chemical Engineering, Espoo, 02150, Finland

## ARTICLE INFO

**Keywords:**

Supercritical extraction  
Optimal design of experiment  
Model discrimination  
Mathematical modelling

## ABSTRACT

This study investigates the process of chamomile oil extraction from flowers. A parameter-distributed model consisting of a set of partial differential equations is used to describe the governing mass transfer phenomena in a solid-fluid environment under supercritical conditions using carbon dioxide as a solvent. The concept of quasi-one-dimensional flow is applied to reduce the number of spatial dimensions. The flow is assumed to be uniform across any cross-section, although the area available for the fluid phase can vary along the extractor. The physical properties of the solvent are estimated from the Peng–Robinson equation of state. Based on the set of laboratory experiments performed under multiple constant operating conditions: 30–40 °C, 100–200 bar, and  $3.33 - 6.67 \cdot 10^{-5}$  kg/s, two process models are developed. The goal of this work is to design an experiment to discriminate between two competing process models. Statistical methods, such as the Kolmogorov–Smirnov and Mann–Whitney U test, and Jensen–Shannon divergence, are used to discriminate between two models under the parameter uncertainty.

## 1. Introduction

Supercritical CO<sub>2</sub> is defined as carbon dioxide that is pressurized and heated above its critical point (31.1 °C, 74 bar). Depending on the operating conditions, the fluid properties such as viscosity and density can vary, which leads to multiple industrial applications of CO<sub>2</sub>.

One of the most popular applications of supercritical CO<sub>2</sub> is the extraction of essential oils, as described by many researchers, for example, by Sodeifian and Sajadian [1], Reverchon et al. [2] or Sovova [3]. Traditional methods, such as distillation and organic solvent extraction, are commonly employed but have drawbacks. Distillation, involves high temperatures that can lead to the thermal degradation of heat-sensitive compounds. This limitation has increased the popularity of alternative techniques, such as supercritical fluid extraction. Supercritical CO<sub>2</sub> is appealing due to its distinctive properties: it is inflammable, non-toxic and non-corrosive. Supercritical fluids can exhibit both gas- and liquid-like properties, allowing for adjustable dissolving power through changes in operating conditions.

The applications of supercritical carbon dioxide are not limited only to an extraction process but can also be used for impregnation as described by Weidner [4], Machado et al. [5] or Fathi et al. [6]. Impregnation is defined as modifying the properties of bulk substances by physically or chemically binding/adsorbing impregnates to a bulk material or surface, such as the hydrophobization of surfaces. The main advantage of using supercritical CO<sub>2</sub> is that after depressurization, it desorbs from the surface and evaporates, leaving a solvent-free product. On the other hand, the main disadvantage of using carbon dioxide for impregnation is the low solubility of many drugs of interest.

Another application of supercritical CO<sub>2</sub> is nanoparticles formation as investigated by Padrela et al. [7], Franco

and De Marco [8], Saadati Ardestani et al. [9] or Sodeifian et al. [10]. Supercritical carbon-dioxide-assisted technologies enable the production of different morphologies of different sizes, including nanoparticles and nanocrystals, by modulating operating conditions. Supercritical fluid-based processes have advantages over techniques conventionally employed to produce nanosized particles or crystals, such as reduced use of toxic solvents. Moreover, the CO<sub>2</sub> is completely removed from the final product by simple depressurization.

This study investigates the extraction of essential oil from chamomile flowers (*Matricaria chamomilla L.*) via supercritical fluid extraction techniques and the modelling of this process. Chamomile is a medicinal herb widely cultivated in southern and eastern Europe — in countries such as Germany, Hungary, France and Russia. It can be found outside Europe, for instance in Brazil as discussed by Singh et al. [11]. This plant is distinguished by its hollow, bright gold cones, housing disc or tubular florets and surrounded by about fifteen white ray or ligulate florets. Chamomile has been used for its medicinal benefits, serving as an anti-inflammatory, antioxidant, mild astringent, and healing remedy. Extracts of chamomile are widely used to calm nerves and mitigate anxiety, hysteria, nightmares, insomnia and other sleep-related conditions, according to Srivastava [12]. Orav et al. [13] reported that oil yields from dried chamomile samples ranged from 0.7 to 6.7 mL/kg. The highest yields of essential oil, between 6.1 and 6.7 mL/kg, were derived from chamomile sourced from Latvia and Ukraine. In comparison, chamomile from Armenia exhibited a lower oil content of 0.7 mL/kg.

The literature offers various mathematical models to describe the extraction of valuable compounds from biomass. Selecting a process model is case-to-case dependent and requires analysis of each model's specific assumptions about mass transfer and thermodynamic equilibrium.

Goto et al. [14] presented the Shrinking Core (SC) model, which describes a process of irreversible desorption

\*Corresponding author

✉ oliwer.sliczniuk@aalto.fi (O. Sliczniuk)  
ORCID(s): 0000-0003-2593-5956 (O. Sliczniuk)

that is followed by diffusion through the pores of a porous solid. When the mass transfer rate of the solute in the non-extracted inner region is significantly slower than in the outer region, where most of the solute has already been extracted, or when the solute concentration exceeds its solubility in the solvent, a distinct boundary may form between the inner and outer regions. As extraction progresses, the core of the inner region shrinks. The model envisions supercritical  $\text{CO}_2$  extraction as a sharp, inward-moving front, with a completely non-extracted core ahead of the front and a fully extracted shell behind it.

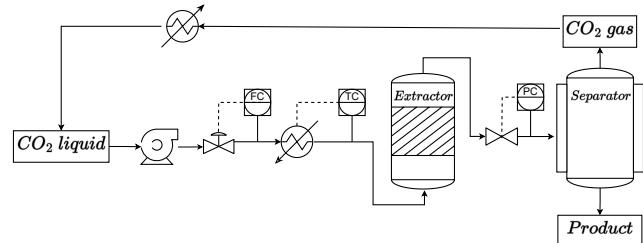
Sovova [3] proposed The Broken-and-Intact Cell (BIC) model, which assumes that a portion of the solute, initially stored within plant structures and protected by cell walls, is released during the mechanical breakdown of the material. The solute located in the region of broken cells near the particle surface is directly exposed to the solvent, while the core of the particle contains intact cells with undamaged walls. This model describes three extraction phases: a fast extraction phase for accessible oil, a transient phase, and a slow phase controlled by diffusion. The model has been successfully applied to the extraction of grape oil (Sovová et al. [15]) and caraway oil (Sovova et al. [16]).

The Supercritical Fluid Extraction (SFE) process can be treated similarly to heat transfer, considering solid particles like hot balls cooling down in a uniform environment. Bartle et al. [17] introduced the hot ball diffusion (HBD) model, where spherical particles with uniformly distributed solute diffuse similarly to heat diffusion. Unlike the BIC model, where solute is readily available on the particle surface, the HBD model is suited for systems with small quantities of extractable materials and is not limited by solubility. The model is particularly relevant when internal diffusion controls mass transfer, allowing results from single particles to be extended to the entire bed under uniform conditions. Reverchon et al. [2] have further elaborated on the HBD model and used it to simulate extraction processes for natural materials.

Reverchon [18] proposed a model for extraction of essential oils, which are mainly located inside the vegetable cells in organelles called vacuoles. Only a small fraction of essential oil might be near the particle surface due to the breaking up of cells during grinding or in epidermal hairs located on the leaf surface. The fraction of oil freely available on the particle surface should not be significant in the case of SFE from leaves. Consequently, SFE of essential oil from leaves should be mainly controlled by the internal mass-transfer resistance. Therefore, the external mass-transfer coefficient was neglected in the development of the model of Reverchon [18]. The mass balances were developed in the additional hypotheses that the axial dispersion can be neglected and that the solvent density and flow rate are constant along the bed.

This work builds upon the linear kinetic model suggested by Reverchon [18], deriving fundamental governing equations to develop a comprehensive model for the chamomile oil extraction process. This model aims for

control-oriented simplicity, assuming a semi-continuous operation within a cylindrical vessel. The process involves a supercritical solvent being pumped through a fixed bed of finely chopped biomass to extract the solute, followed by separation of the solvent and solute in a flush drum to collect the extract. Parameters such as pressure ( $P$ ), feed flow rate ( $F$ ) and inlet temperature ( $T_{in}$ ) are adjustable and measurable, while the outlet temperature ( $T_{out}$ ) and the amount of product at the outlet can only be monitored. Figure 1 presents a simplified process flow diagram.



**Figure 1:** Process flow diagram.

In the study of dynamical systems, selecting the most accurate model to represent a physical process is crucial for effective prediction, control, and understanding of underlying mechanisms. When multiple models are trained on the same dataset but exhibit subtle structural differences, their responses may diverge in nuanced ways. Traditional model selection criteria often fail to capture these subtleties, making it necessary to employ a robust statistical framework for optimal model discrimination.

The process of discriminating between competing models typically follows a sequential approach, involving two iterative steps. The first is the design phase, where a discrimination criterion is used to select the next experiment. This criterion, which reflects the information content of potential experimental conditions, helps identify the experiment that is expected to provide the most information for distinguishing between models. Once the experiment is conducted, the second step—data analysis—is performed. A stopping rule is then applied to determine if one model can be deemed significantly better than the others. If no clear distinction is made, the process is repeated. Most model discrimination criteria rely on either maximum divergence or maximum entropy principles.

Building on information theory and a Bayesian approach, Box and Hill [19] introduced a sequential design procedure for model discrimination. Since entropy measures the amount of available information about a system, additional experiments should be selected to maximize the expected change in entropy, thereby increasing the information obtained about the system through successive experiments.

Buzzi-Ferraris and Forzatti [20] observed that the performance of Box's sequential design for model discrimination depends on the order in which experimental observations are processed. This is due to the recursive nature of updating model probabilities. As model probabilities should be independent of the data presentation order, the authors

noted that experiments should be selected in regions where differences in prediction variances are large, rather than where differences in model responses are most pronounced. They applied this approach to two cases: ammonia synthesis from nitrogen and hydrogen, and vapor-phase chlorination of tetrachloroethane using activated silica gel at 200 °C.

Chen and Asprey [21] proposed a novel method for designing optimally informative dynamic experiments aimed at discriminating among several competing multiresponse nonlinear dynamic models, which are generally described by systems of differential and algebraic equations. Their approach extends Buzzi-Ferraris and Forzatti [20]’ discrimination criterion to dynamic systems, reformulating the experimental design problem as an optimal control problem. This method was applied to the fed-batch fermentation of Baker’s yeast, with the goal of discriminating between four competing models by identifying optimal profiles for the dilution factor and inlet substrate concentration.

## 2. Materials and methods

### 2.1. Governing equations

Following the work of Anderson [22], the governing equations for a quasi-one-dimensional flow were derived. A quasi-one-dimensional flow refers to a fluid flow scenario assuming that the flow properties are uniformly distributed across any cross-section. This simplification is typically applied when the flow channel’s cross-sectional area changes, such as through irregular shapes or partial filling of an extractor. According to this assumption, velocity and other flow properties change solely in the flow direction.

As discussed by Anderson [23], all flows are compressible, but some of them can be treated as incompressible since the velocities are low. This assumption leads to the incompressible condition:  $\nabla \cdot u = 0$ , which is valid for constant density (strict incompressible) or varying density flow. The assumption allows for removing acoustic waves and large perturbations in density and/or temperature. In the 1-D case, the incompressibility condition becomes  $\frac{du}{dz} = 0$ , so the fluid velocity is constant along the  $z$ -direction.

The set of quasi-one-dimensional governing equations in Cartesian coordinates is described by Equations 1 - 3:

$$\frac{\partial(\rho_f A_f)}{\partial t} + \frac{\partial(\rho_f A_f v)}{\partial z} = 0 \quad (1)$$

$$\frac{\partial(\rho_f v A_f)}{\partial t} + \frac{\partial(\rho_f A_f v^2)}{\partial z} = -A_f \frac{\partial P}{\partial z} \quad (2)$$

$$\frac{\partial(\rho_f e A_f)}{\partial t} + \frac{\partial(\rho_f A_f v e)}{\partial z} = -P \frac{\partial(A_f v)}{\partial z} + \frac{\partial}{\partial z} \left( k \frac{\partial T}{\partial z} \right) \quad (3)$$

where  $\rho_f$  is the density of the fluid,  $A_f$  is the function which describes a change in the cross-section,  $v$  is the velocity,  $P$  is the total pressure,  $e$  is the internal energy of the fluid,  $t$  is time and  $z$  is the spatial direction.

### 2.2. Extraction model

#### 2.2.1. Continuity equation

The previously derived quasi-one-dimensional continuity equation (Equation 1) is redefined by incorporating

the function  $A_f = A\phi$ . This modification distinguishes constant and varying terms, where the varying term accounts for changes in the cross-sectional area available for the fluid. Equation 4 shows the modified continuity equation:

$$\frac{\partial(\rho_f \phi)}{\partial t} + \frac{\partial(\rho_f v A \phi)}{\partial z} = 0 \quad (4)$$

where  $A$  is the total cross-section of the extractor and  $\phi$  describes porosity along the extractor.

Assuming that the mass flow rate is constant in time, the temporal derivative becomes the mass flux  $F$ , and the spatial derivative can be integrated along  $z$  as

$$\int \frac{\partial(\rho_f v A \phi)}{\partial z} dz = F \rightarrow F = \rho_f v A \phi \quad (5)$$

To simplify the system dynamics, it is assumed that  $F$  is a control variable and affects the whole system instantaneously (due to  $\nabla \cdot u = 0$ ), which allows finding the velocity profile that satisfies mass continuity based on  $F$ ,  $\phi$  and  $\rho_f$ :

$$v = \frac{F}{\rho_f A \phi} \quad (6)$$

Similarly, superficial velocity may be introduced:

$$u = v\phi = \frac{F}{\rho_f A} \quad (7)$$

The fluid density  $\rho_f$  can be obtained from the Peng-Robinson equation of state if the temperature and thermodynamic pressure are known along  $z$ . Variation in fluid density may occur due to pressure or inlet temperature changes. In a non-isothermal case, in Equations 6 and 7  $\rho_f$  is considered the average fluid density along the extraction column.

#### 2.2.2. Mass balance for the fluid phase

Equation 8 describes the movement of the solute in the system, which is constrained to the axial direction due to the quasi-one-dimensional assumption. Given that the solute concentration in the solvent is negligible, the fluid phase is described as pseudo-homogeneous, with properties identical to those of the solvent itself. It is also assumed that the thermodynamic pressure remains constant throughout the device. The analysis further simplifies the flow dynamics by disregarding the boundary layer near the extractor’s inner wall. This leads to a uniform velocity profile across any cross-section perpendicular to the axial direction. Thus, the mass balance equation includes convection, diffusion and kinetic terms representing the fluid phase behaviour:

$$\frac{\partial c_f}{\partial t} + \frac{1}{\phi} \frac{\partial(c_f u)}{\partial z} = \frac{1-\phi}{\phi} r_e + \frac{1}{\phi} \frac{\partial}{\partial z} \left( D_e^M \frac{\partial c_f}{\partial z} \right) \quad (8)$$

where  $c_f$  represents the solute concentration in the fluid phase,  $r_e$  is the mass transfer kinetic term and  $D_e^M$  is the axial diffusion coefficient.

#### 2.2.3. Mass balance for the solid phase

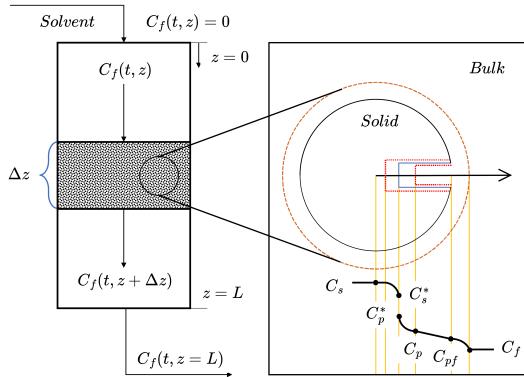
As given by Equation 9, the solid phase is considered stationary, without convection and diffusion terms in the mass balance equation. Therefore, the only significant term in this equation is the kinetic term of Equation 10, which

connects the solid and fluid phases. For simplicity, the extract is represented by a single pseudo-component:

$$\frac{\partial c_s}{\partial t} = \underbrace{r_e}_{\text{Kinetics}} \quad (9)$$

#### 2.2.4. Kinetic term

As the solvent flows through the fixed bed,  $\text{CO}_2$  molecules diffuse into the pores, adsorb on the inner surface and form a film due to solvent-solid matrix interactions. The dissolved solute diffuses from the particle's core through the solid-fluid interface, the pore and the film into the bulk. Figure 2 shows the mass transfer mechanism, where the mean solute concentration in the solid phase is denoted as  $c_s$ , and the equilibrium concentrations at the solid-fluid interface are denoted as  $c_s^*$  and  $c_p^*$  for the solid and fluid phases, respectively. The concentration of the solutes in the fluid phase in the centre of the pore is denoted as  $c_p$ . As the solute diffuses through the pore, its concentration changes, reaching  $c_{pf}$  at the opening. Then, the solute diffuses through the film around the particle and reaches bulk concentration  $c_f$ . The two-film theory describes the solid-fluid interface inside the pore. The overall mass transfer coefficient can be determined from the relationship between the solute concentration in one phase and its equilibrium concentration.



**Figure 2:** Mass transfer mechanism.

Bulley et al. [24] suggest a process where the driving force for extraction is given by the difference between the concentration of the solute in the bulk,  $c_f$ , and in the centre of the pore,  $c_p^*$ . The concentration  $c_p^*$  is in equilibrium with  $c_s$  according to the equilibrium relationship. The rate of extraction is thus  $r_e (c_f - c_p^*(c_s))$ . In contrast, Reverchon [18] proposes a driving force given by the difference between  $c_s$  and  $c_p^*$ . As given by Equation 10, the concentration  $c_p^*$  is determined by the equilibrium relationship with  $c_f$ :

$$r_e = \frac{D_i}{\mu l^2} (c_s - c_p^*) \quad (10)$$

where  $\mu$  is sphericity,  $l$  a characteristic dimension of particles that can be defined as  $l = r/3$ ,  $r$  is the mean particle radius,  $\rho_s$  is the solid density,  $D_i$  corresponds to the overall

diffusion coefficient and  $c_p^*$  is the concentration at the solid-fluid interface (which according to the internal resistance model is supposed to be at equilibrium with the fluid phase).

According to Bulley et al. [24], a linear equilibrium relationship (Equation 11) can be used to find the equilibrium concentration of the solute in the fluid phase  $c_f^*$  based on the concentration of the solute in the solid phase  $c_s$ :

$$c_f^* = k_p c_s \quad (11)$$

The volumetric partition coefficient  $k_p$  acts as an equilibrium constant between the solute concentration in one phase and the corresponding equilibrium concentration at the solid-fluid interphase. Spiro and Kandiah [25] proposed to define the mass partition coefficient  $k_m$  as

$$k_m = \frac{k_p \rho_s}{\rho_f} \quad (12)$$

According to Reverchon [18], the kinetic term becomes

$$r_e = -\frac{D_i}{\mu l^2} \left( c_s - \frac{\rho_s c_f}{k_m \rho_f} \right) \quad (13)$$

#### 2.2.5. Uneven solute's distribution in the solid phase

Following the idea of the Broken-and-Intact Cell (BIC) model (Sovova [26]), the internal diffusion coefficient  $D_i$  is considered to be a product of the reference value of  $D_i^R$  and the exponential decay function  $\gamma$ , as given by Equation 14:

$$D_i = D_i^R \gamma(c_s) = D_i^R \exp \left( \Upsilon \left( 1 - \frac{c_s}{c_{s0}} \right) \right) \quad (14)$$

where  $\Upsilon$  describes the curvature of the decay function. Equation 15 describes the final form of the kinetic term:

$$r_e = -\frac{D_i^R \gamma}{\mu l^2} \left( c_s - \frac{\rho_s c_f}{k_m \rho_f} \right) \quad (15)$$

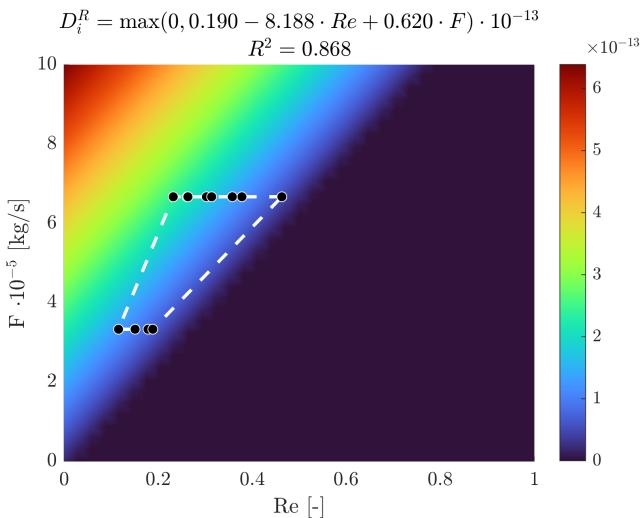
The  $\gamma$  function limits the solute's availability in the solid phase. Similarly to the BIC model, the solute is assumed to be contained in the cells, some of which are open because the cell walls were broken by grinding, with the rest remaining intact. The diffusion of the solute from a particle's core takes more time than the diffusion of the solute close to the outer surface. The same idea can be represented by the decaying internal diffusion coefficient, where the decreasing term is a function of the solute concentration in the solid.

An alternative interpretation of the decay function  $\gamma$  involves considering the porous structure of the solid particles, where the pores are initially saturated with the solute. During extraction, the solute within these pores gradually dissolves into the surrounding fluid. Initially, the solute molecules near the pore openings dissolve and diffuse rapidly due to the short diffusion paths. As the extraction progresses, the dissolution front moves deeper into the pore structure, and solute from the inner regions of the pores begins to dissolve. The diffusion of solute molecules from the interior of the pores to the external fluid becomes progressively slower because the effective diffusion path length increases. This lengthening of the diffusion path enhances the mass transfer resistance, reducing the overall diffusion rate.

In an extreme case, this model could be compared with the Shrinking Core Model presented by Goto et al. [14], where the particle radius decreases as the solute content in the solid phase diminishes. In the SC model, the reduction in particle size leads to significant changes in both the diffusion path length and the surface area available for mass transfer. The diminishing particle size increases the diffusion path within the remaining solid core and decreases the external surface area, both of which contribute to a slower extraction rate. By comparing this to the varying diffusion coefficient in our model, some conceptual similarities can be noticed.

### 2.2.6. Empirical correlations

The empirical correlations for  $D_i$  and  $\Upsilon$  were derived by Sliczniuk and Oinas [27] and validated for temperatures between  $30 - 40$  °C, pressures between 100 – 200 bar, and mass flow rates between  $3.33 - 6.67 \cdot 10^{-5}$  kg/s. Figures 3 and 4 show the results of multiple linear regression applied to solutions of parameter estimation and selected independent variables. The region marked with the white dashed line represents the confidence region, where the model has been tested. Both correlations should be equal of greater than zero, to avoid unphysical behaviour such as the reverse mass transfer. The multiple linear regression functions are combined with the rectifier function to ensure the non-negativity.



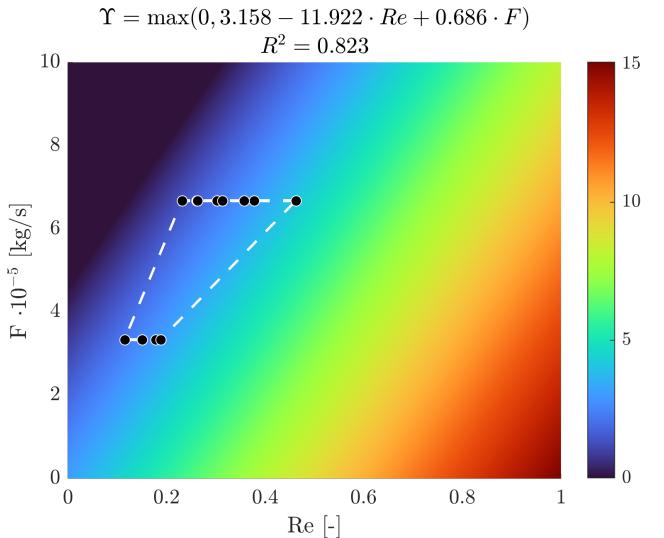
**Figure 3:** Multiple linear regression  $D_i^R = f(Re, F)$

### 2.2.7. Heat balance

The heat balance equation describe the evolution of the enthalpy in the system and it is given by Equation 16

$$\frac{\partial(\rho_f h A_f)}{\partial t} = -\frac{\partial(\rho_f h A_f v)}{\partial z} + \frac{\partial(P A_f)}{\partial t} + \frac{\partial}{\partial z}\left(k \frac{\partial T}{\partial z}\right) \quad (16)$$

If the value of enthalpy  $h$  is known from the time evolution of the energy equation, and pressure  $P$  is known from measurement, then the temperature  $T$  can be reconstructed based on the departure function. The departure function is a mathematical function that characterizes the deviation of a thermodynamic property (enthalpy, entropy, and internal



**Figure 4:** Multiple linear regression  $\Upsilon = f(Re, F)$

energy) of a real substance from that of an ideal gas at the same temperature and pressure. As presented by Gmehling et al. [28], for the Peng-Robinson equation of state, the enthalpy departure function is defined by Equation 17.

$$h - h^{id} = RT \left[ T_r(Z - 1) - 2.078(1 + \kappa)\sqrt{\alpha(T)} \ln \left( \frac{Z + (1 + \sqrt{2})B}{Z + (1 - \sqrt{2})B} \right) \right] \quad (17)$$

where  $\alpha$  is defined as  $(1 + \kappa)(1 - \sqrt{T_r})^2$ ,  $T_r$  is the reduced temperature,  $P_r$  is the reduced pressure,  $Z$  is the compressibility factor,  $\kappa$  is a quadratic function of theacentric factor and  $B$  is calculated as  $0.07780 \frac{P_r}{T_r}$ .

Equation 17 requires a reference state, which is assumed to be  $T_{ref} = 298.15$  K and  $P_{ref} = 1.01325$  bar.

A root-finder can be used to find a value of temperature, which minimizes the difference between the value of enthalpy coming from the heat balance and the departure functions. The root finding procedure is repeated at every time step to find a temperature profile along spatial direction  $z$ .

$$\min_T \left( \underbrace{h(t, x)}_{\text{Heat balance}} - \underbrace{h(T, P, \rho_f(T, P))}_{\text{Departure function}} \right)^2 \quad (18)$$

### 2.2.8. Pressure term

As explained in Chapters 2.1, this low velocity region, pressure is, nearly constant due to the small pressure wave propagation that occurs at the speed of sound. Under such conditions, the term  $\partial P / \partial t$  can be approximated by a difference equation, describing the pressure change in the whole system. The pressure  $P$  in the system is considered a state variable, while the pressure in the new time-step  $P_{in}$  is considered a control variable.

$$\frac{\partial P}{\partial t} \approx \frac{P_{in} - P}{\Delta t} \quad (19)$$

Such a simplified equation allows for instantaneous pressure change in the system but does not consider a gradual pressure build-up and the effects of pressure losses. In a real system, the dynamics of pressure change would depend on a pump and a back-pressure regulator.

### 2.2.9. Extraction yield

The process yield is calculated according to Equation 20 as presented by Sovova et al. [16]. The measurement equation evaluates the solute's mass at the extraction unit outlet and sums it up. The integral form of the measurement (Equation 20) can be transformed into the differential form (Equation 21) and augmented with the process model.

$$y = \int_{t_0}^{t_f} \frac{F}{\rho_f} c_f \Big|_{z=L} dt \quad (20)$$

$$\frac{dy}{dt} = \frac{F}{\rho_f} c_f \Big|_{z=L} \quad (21)$$

### 2.2.10. Initial and boundary conditions

It is assumed that the solvent is free of solute at the beginning of the process  $c_{f0} = 0$ , that all the solid particles have the same initial solute content  $c_{s0}$ , and that the system is isothermal, hence the initial state is  $h_0$ . The fluid at the inlet is considered not to contain any solute. The initial and boundary conditions are defined as follows:

$$\begin{aligned} c_f(t=0, z) &= 0 & c_s(t=0, z) &= c_{s0} & h(t=0, z) &= h_0 \\ c_f(t, z=0) &= 0 & h(t, z=0) &= h_{in} & \frac{\partial c_f(t, z=L)}{\partial x} &= 0 \\ \frac{\partial h(t, z=L)}{\partial x} &= 0 & c_s(t, z=\{0, L\}) &= 0 & y(0) &= 0 & P(0) &= P_0 \end{aligned}$$

### 2.2.11. Discretization methods

The method of lines is used to transform the process model equations into a set of ODEs denoted as  $G(x; \Theta)$ . The backward finite difference is used to approximate the first-order derivative, while the central difference scheme approximates the second-order derivative  $z$  direction. The length of the fixed bed is divided into  $N_z$ , i.e. equally distributed points in the  $z$  direction. The state-space model after discretization is denoted as  $x$  and defined as follows:

$$\dot{x} = \frac{dx}{dt} = \begin{bmatrix} \frac{dc_{f,1}}{dt} \\ \vdots \\ \frac{dc_{f,N_z}}{dt} \\ \hline \frac{dc_{s,1}}{dt} \\ \vdots \\ \frac{dc_{s,N_z}}{dt} \\ \hline \frac{dh_1}{dt} \\ \vdots \\ \frac{dh_{N_z}}{dt} \\ \hline \frac{dP}{dt} \\ \hline \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} G_1(c_f, c_s, h; \Theta) \\ \vdots \\ G_{N_z}(c_f, c_s, h; \Theta) \\ \hline G_{N_z+1}(c_f, c_s, h; \Theta) \\ \vdots \\ G_{2N_z}(c_f, c_s, h; \Theta) \\ \hline G_{2N_z+1}(c_f, c_s, h; \Theta) \\ \vdots \\ G_{3N_z}(c_f, c_s, h; \Theta) \\ \hline G_{3N_z+1}(c_f, c_s, h; \Theta) \\ \hline G_{3N_z+2}(c_f, c_s, h; \Theta) \end{bmatrix} \underbrace{G(x; \Theta)}$$

where  $x \in \mathbb{R}^{N_x=3N_z+2}$  and  $\Theta \in \mathbb{R}^{N_\Theta=N_\theta+N_u}$ ,  $N_\theta$  is the number of parameters,  $N_u$  is the number of control variables.

For a derivative to be conservative, it must form a telescoping series. In other words, only the boundary terms should remain after adding all terms coming from the discretization over a grid, and the artificial interior points should be cancelled out. Discretization is applied to the conservative form of the model to ensure mass conservation.

## 3. Model discrimination

Multiple alternative models are frequently proposed to describe the same physical phenomenon. Discriminating between these models requires performing a new set of experiments. Each model predicts the response  $y$  as a function of experimental conditions  $\Xi$  and model parameters  $\theta$  (here,  $\theta$  refers to parameters describing extraction kinetics and constitutes a subset of the parameter space  $\Theta$ ). The models differ in their mathematical formulation and the parameters involved, although some parameters appearing in different models may share the same physical interpretation.

Mathematical models can be developed from multiple hypotheses about physical phenomena and subsequently tested against experimental data. Alternatively, universal function approximators can be employed when limited information about the system is available. Depending on prior knowledge, model developers can choose between purely theoretical and data-driven approaches. The following section presents an alternative formulation to the model described in Section ??, specifically substituting the extraction kinetic term presented in Section 2.2.4. The details on the challenger model can be found in Appendix A.1. This combination leverages general knowledge about the system while acknowledging that extraction kinetics may be difficult to

formulate solely from mass transfer theory. The fundamental source of modelling uncertainty lies in the fact that solute mass transfer between solid and fluid phases depends on multiple factors and phenomena that vary between plant materials. Consequently, no unified model exists that can accurately describe all SFE processes.

### 3.1. Statistical analysis for the single-point model discrimination

When developing a process model, several alternative modelling solutions are typically proposed and subsequently compared. Discrepancies in the outputs of competing models under identical inputs are expected due to differences in model structure. Identifying where models differ most substantially can reveal model or data deficiencies.

Following the work of Box and Hill [19], Himmelblau [29], and Bard [30], the ratio of two probability distributions can serve as a measure of evidence favouring one model over another. The quantity  $\ln\left(\frac{p_1(y)}{p_2(y)}\right)$  represents the odds in favour of hypothesis  $H_1$  (that  $p_1(y)$  is the true model) over hypothesis  $H_2$  (that  $p_2(y)$  is the true model). Alternatively, this ratio can be interpreted as the information supporting hypothesis  $H_1$  relative to hypothesis  $H_2$ . The ‘weight of evidence’ or expected information favouring  $H_1$  over  $H_2$  can be defined through the Kullback–Leibler divergence (KL divergence):

$$D_{\text{KL}}(p_1(y)\|p_2(y)) = \int_{-\infty}^{\infty} p_1(y) \ln\left(\frac{p_1(y)}{p_2(y)}\right) dy \quad (22)$$

In essence,  $D_{\text{KL}}(p_1(y)\|p_2(y))$  measures how distinctly data drawn from  $p_1$  would differ from data drawn from  $p_2$ . A large value indicates that observations from  $p_1$  are unlikely to be mistaken for those from  $p_2$ . Conversely,  $D_{\text{KL}}(p_2(y)\|p_1(y))$  quantifies the difficulty of confusing data from  $p_2$  with data from  $p_1$ . These divergences can be viewed as asymmetric measures of distance between probability distributions. If model 1 is correct, it is desirable to conduct an experiment  $\Xi$  likely to confirm this by producing a large value of  $D_{\text{KL}}(p_1(y)\|p_2(y))$ . Conversely, if model 2 is correct, the experiment  $\Xi$  should yield a large value of  $D_{\text{KL}}(p_2(y)\|p_1(y))$ .

The probability distributions  $p_1(y)$  and  $p_2(y)$  can be obtained analytically under certain assumptions (Section ??) or numerically via Monte Carlo (MC) methods. In this section, the Monte Carlo approach is used to evaluate the probability density function of model  $i$  output at fixed operating conditions under parameter uncertainty:

$$p_i(y(t)) = \int p_i(y(t)|\theta_i) p_i(\theta_i|\hat{\theta}_i, \Sigma_i) d\theta_i \quad (23)$$

Equation 23 marginalises over the parameter uncertainty, where  $p_i(\theta_i|\hat{\theta}_i, \Sigma_i)$  represents the posterior distribution of model parameters given estimates  $\hat{\theta}_i$  and covariance  $\Sigma_i$ . This integral can be approximated by several methods such as Kernel Density Estimation (KDE) with a Gaussian kernel. Based on MC simulations, a collection of independent and identically distributed model output samples is

obtained. These samples are fitted using KDE to estimate the probability density function  $p_i(y)$ . This procedure is repeated for both models, and the resulting densities are combined to compute the KL divergence, which is evaluated by numerical integration. It is noted that a degenerate distribution of  $p(y_i)$  is not considered an issue due to the marginalisation over the parameter uncertainty.

Although, KL divergence measures difference between two distributions it suffers from asymmetry ( $KL(P|L) \neq KL(L|P)$ ), and is unbound. Consequently, the Jensen–Shannon divergence (J-S D) is introduced (Equation 24). When using logarithm base 2, J-SD is bounded between 0 and 1 for discrete distributions. Otherwise, the upper bound of  $D_J$  is  $\ln(2)$ .

$$D_J = D_{\text{KL}}(p_1(y)\|p_2(y)) + D_{\text{KL}}(p_2(y)\|p_1(y)) \quad (24)$$

Except for J-S D, the Kolmogorov–Smirnov test (K-S) is applied to compare empirical distributions obtained from the models. K-S is a non-parametric test of the equity between one-dimensional probability distributions. To be more precise, the test checks if the analysed samples come from the same distribution. As there is no reference distribution, the two-sample version of the test is used. Based on the analogy to the binary classification problem, alternative measures, such as the Mann–Whitney U or AUC (Area Under Curve) tests, can be introduced. The idea is to treat samples from both models as if they came from different classes. The AUC becomes probability that a randomly drawn Power model samples exceeds a randomly drawn Linear model samples, plus half the probability of a tie. This yields a value in between 0 and 1, where 0.5 indicates no discrimination (complete overlap) and values closer to 1 indicate strong separation. The last proposed way of measuring the discrimination potential is to analyse the distribution of the difference between both models, and check if zero lies within the 95% confidence interval.

### 3.2. Results

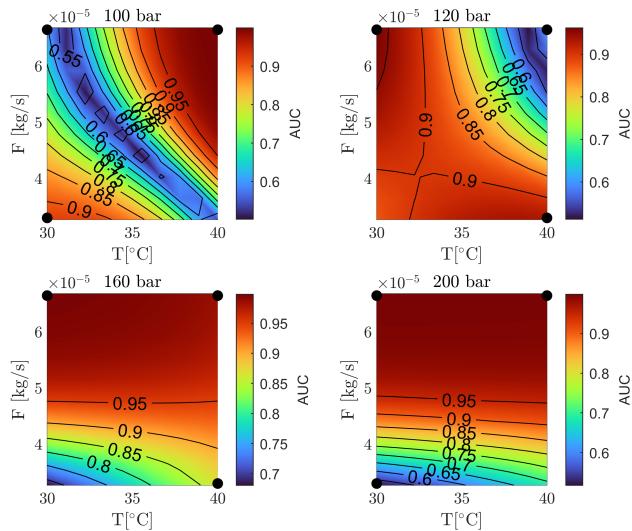
Figures 5a and 5b show values of JS-D and AUC based on the cumulative yield at the final time (600 min) and various operating conditions.

Both figures show different similarities in terms of identified regions of model discrimination. AUC and J-S D suggest that at high pressures (160 and 200 bar), the discrimination potential increases with the mass flow-rate. The same is not true for 120 bar, where models become difficult to discriminate from each other in the region of high mass flow-rate and elevated temperature. The 100 bar regions shows a distinct zone of low model separability, and two regions of high model discrimination.

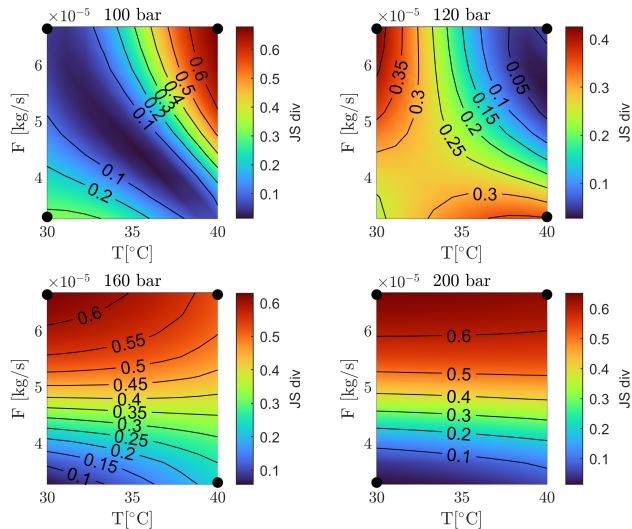
Figures 8 and 9 show temporal evolution of model outputs (cumulative differentiated yield) at high and low discrimination operating conditions.

Comparison between Figures 8 and 9 shows the main drawback of using point estimates for the model discrimination. This method doesn't take into account the previous observations, so might be misleading if both models coincidentally shows similar results at the measurement time. It is

## ODOE-DM



(a) AUC results



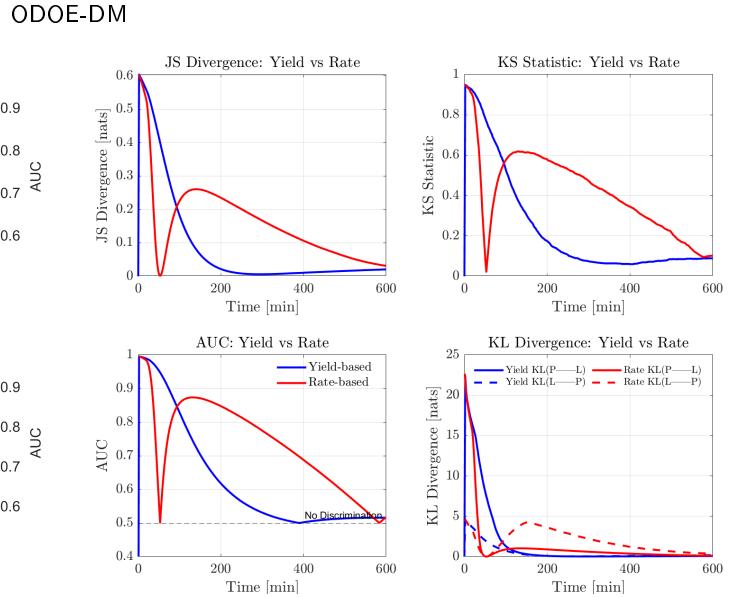
(b) JS-D results

**Figure 5:** Discrimination indicators at final time and under various operating conditions

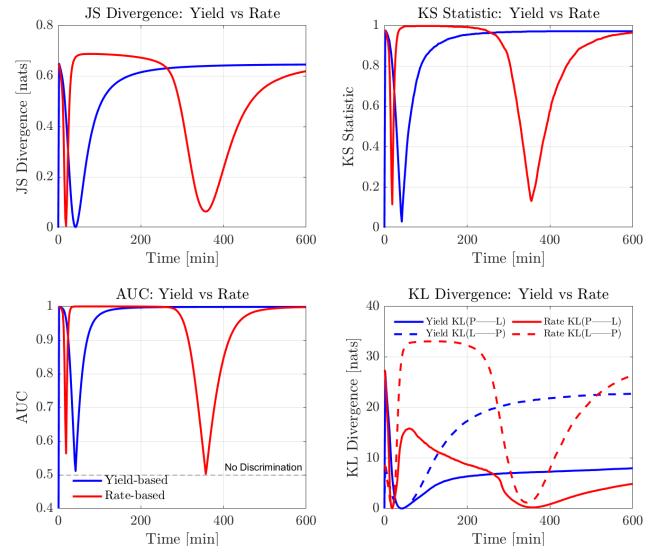
noted that two dynamical systems might reach the same state but on different path, which wouldn't be recognised by the single-point estimates.

The selection of appropriate model output should be consider given the system dynamic, as they might lead to different conclusion. The single-point estimates from the cumulative and differentiated yield are compared in Figure 6.

Detailed analysis of Figures 6a and 6b showed different patterns depending on the discrimination potential. Due to its nature, the cumulative yield absorbs the fluctuations in the measurements which leads to more stable results, but might miss informative local changes in the model output. It is evident on from the plots that indicators computed from the differentiated varies more than the one for cumulative yield,



(a)  $F=3.33e-5\text{kg}/\text{s}$



(b)  $F=6.67e-5\text{kg}/\text{s}$

**Figure 6:** Temporal evolution of discrimination indices at  $T=30^\circ \text{C}$ ,  $P=200 \text{ bar}$

which would affect the conclusion on the model discrimination. Figure 6a shows that the rate-based indices dropped (overlaid distributions) between 300 and 400 min, while the indicators based on the circulative yield has not. If a rate-based index from that period would be used, then

Although, J-S D, AUC and K-S exhibit similar patterns, the analysis of K-L D plots reveal asymmetric behaviour. K-L D can be understood as a measure of how much information is lost when one distribution approximates another. The large difference between  $\text{KL}(\text{L}|\text{P})$  and  $\text{KL}(\text{P}|\text{L})$ , provide information that one model return extreme results which are 'unlikely' accordingly to the second model. This interpenetration aligns with comparison of distributions presented in Figure 9.

## 4. Conclusions

This work focus on model discrimination, which is a part of model development process. After several challenger models are developed, the champion model needs to be selected based on case-specific criteria. These might include model accuracy, complexity or predictive power. The predictive power can be tested on validation dataset, which it self can be a random subset of the development dataset or a new experiment to be yet performed. In the second case, it is beneficial to select such operating conditions that both models differ the most. Such design of an experiment

$$y(t_k) = Y_{true}(t_k) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, \sigma_{meas}^2)$$

where  $Y_{true}$  is whatever the real process produces, and  $\sigma_{meas}$  captures instrument precision (scale repeatability, sampling error, etc.). The noise at different times is assumed independent (the measurement device doesn't remember its previous error).

Given a specific model  $M$  with specific parameters, the model predicts a trajectory  $YM(t; \theta)$ . The likelihood of observing  $y_{obs}$  given that model and those parameters is:

$$p(y_{obs} | \theta, M) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi}, \sigma_{meas}} \exp\left(-\frac{(y(t_k) - YM(t_k; \theta))^2}{2, \sigma_{meas}^2}\right)$$

The product is valid because the measurement errors  $e_k$  are independent. Note: the model trajectories themselves are correlated in time, and that's fine — it's only the measurement noise that needs to be independent.

Taking the log:

$$\log p(y_{obs} | \theta, M) = -\frac{K}{2} \log(2\pi\sigma_{meas}^2) - \frac{1}{2\sigma_{meas}^2} \sum_{k=1}^K (y(t_k) - YM(t_k; \theta))^2$$

This is just the sum of squared residuals, scaled by measurement noise. Smaller residuals higher likelihood.

$$p(y_{obs} | M) = \int p(y_{obs} | \theta, M); p(\theta | M); d\theta$$

This integral has no closed-form solution because  $YM(t; \theta)$  comes from a nonlinear ODE. But you can estimate it with Monte Carlo: you've already drawn  $N$  samples  $\theta(1), \dots, \theta(N)$  from  $p(\theta | M)$  and simulated each one. The MC estimate is:

$$p(y_{obs} | M) \approx \frac{1}{N} \sum_{i=1}^N p(y_{obs} | \theta^{(i)}, M)$$

This is a simple average of likelihoods across your MC samples. It works because the samples are drawn from the prior — this is called a prior predictive Monte Carlo estimator.

The Bayes factor is the ratio of marginal likelihoods:

$$B_{PL} = \frac{p(y_{obs} | M_P)}{p(y_{obs} | M_L)}$$

In log form:

$$\log B_{PL} = \log p(y_{obs} | M_P) - \log p(y_{obs} | M_L)$$

$\log B_{PL} > 0$  - evidence favours Power  $\log B_{PL} < 0$  - evidence favours Linear  $\log B_{PL} = 0$  - data cannot distinguish Step 4: Numerical Computation (Log-Sum-Exp) The naive computation fails because the individual likelihoods  $p(y_{obs} | \theta(i), M)$  involve  $\exp(-SSR / 2\sigma^2)$  where SSR is a sum of  $K$  squared terms. With  $K = 120$  time points and  $\sigma = 0.05$ , these exponentials are astronomically small numbers that underflow to zero in double precision.

The solution is to work entirely in log space. For each MC sample  $i$ :

$$\ell_i = \log p(y_{obs} | \theta^{(i)}, M) = -\frac{K}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^K (y(t_k) - YM^{(i)}(t_k))^2$$

Then the log marginal likelihood is:

$$\log p(y_{obs} | M) = \log \frac{1}{N} \sum_{i=1}^N e^{\ell_i}$$

This is computed via the log-sum-exp trick:

$$\log \sum_{i=1}^N e^{\ell_i} = \ell_{max} + \log \sum_{i=1}^N e^{\ell_i - \ell_{max}}$$

where  $\ell_{max} = \max_i(\ell_i)$ . The subtraction ensures the largest exponential is  $\exp(\ell_{max}) = 1$ , and the rest are  $< 1$ , avoiding overflow. Only samples with  $\ell_i$  very far below  $\ell_{max}$  contribute to zero, but those contribute negligibly to the sum anyway.

Step 5: What This Estimator Actually Does Geometrically Think of each MC sample  $\theta(i)$  as producing a predicted trajectory curve  $YM(i)(t)$ . The observation  $y_{obs}$  is another curve. The likelihood  $\ell_i$  measures the squared distance between these two curves, weighted by  $1/\sigma^2$ .

The marginal likelihood then asks: across all plausible parameter settings, how close does this model's trajectory come to the observation?

If the model's trajectory bundle (your  $Y$  power valid ensemble) passes close to  $y_{obs}$ , many samples have high likelihood → large marginal likelihood If the observation falls outside the bundle, all samples have low likelihood → small marginal likelihood The Bayes factor compares which model's bundle "reaches" the observation more easily.

Why This Differs From Your Pointwise Metrics Your current approach does this at each time point independently:

$$\text{Current: } KL_{integrated} = \int_0^T KL(p(y, t | M_P); ||; p(y, t | M_L)), dt$$

This treats the yield distribution at t=100 min and t=105 min as separate, unrelated comparisons. But in reality, a trajectory that is high at t=100 is almost certainly high at t=105 (strong temporal correlation).

The trajectory Bayes factor operates on the joint distribution:

$$\text{Trajectory BF: } B_{PL} = \frac{p(y(t_1), y(t_2), \dots, y(t_K) | M_P)}{p(y(t_1), y(t_2), \dots, y(t_K) | M_L)}$$

The MC samples encode the joint structure automatically — each row of Y power valid is a complete, physically consistent trajectory from the ODE. You never had to specify the cross-time covariance; it emerges from the simulation.

This is precisely why the trajectory BF can give different answers: if two models have the same marginal distribution at every time point but differ in their correlation structure (e.g., one produces convex curves, the other concave), the pointwise KL would be zero everywhere, but the trajectory BF would still detect the difference.

## References

- [1] Gholamhosseini Sodeifian and Seyed Ali Sajadian. Investigation of essential oil extraction and antioxidant activity of echinophora platyloba dc. using supercritical carbon dioxide. *The Journal of Supercritical Fluids*, 121:52–62, March 2017. ISSN 0896-8446. doi: 10.1016/j.supflu.2016.11.014.
- [2] E. Reverchon, G. Donsi, and L.S. Osseo. Modeling of supercritical fluid extraction from herbaceous matrices. *Industrial & Engineering Chemistry Research*, 32(11):2721–2726, nov 1993. doi: 10.1021/ie00023a039.
- [3] H. Sovova. Rate of the vegetable oil extraction with supercritical co2. modelling of extraction curves. *Chemical Engineering Science*, 49 (3):409–414, 1994. doi: 10.1016/0009-2509(94)87012-8.
- [4] Eckhard Weidner. Impregnation via supercritical co2—what we know and what we need to know. *The Journal of Supercritical Fluids*, 134: 220–227, April 2018. ISSN 0896-8446. doi: 10.1016/j.supflu.2017.12.024.
- [5] Noelia D. Machado, José E. Mosquera, Raquel E. Martini, María L. Goñi, and Nicolás A. Gafán. Supercritical co<sub>2</sub>-assisted impregnation/deposition of polymeric materials with pharmaceutical, nutraceutical, and biomedical applications: A review (2015–2021). *The Journal of Supercritical Fluids*, 191:105763, December 2022. ISSN 0896-8446. doi: 10.1016/j.supflu.2022.105763.
- [6] Mostafa Fathi, Gholamhosseini Sodeifian, and Seyed Ali Sajadian. Experimental study of ketoconazole impregnation into polyvinyl pyrrolidone and hydroxyl propyl methyl cellulose using supercritical carbon dioxide: Process optimization. *The Journal of Supercritical Fluids*, 188:105674, September 2022. ISSN 0896-8446. doi: 10.1016/j.supflu.2022.105674.
- [7] Luís Padrela, Miguel A. Rodrigues, Andreia Duarte, Ana M.A. Dias, Mara E.M. Braga, and Hermínio C. de Sousa. Supercritical carbon dioxide-based technologies for the production of drug nanoparticles/nanocrystals – a comprehensive review. *Advanced Drug Delivery Reviews*, 131:22–78, June 2018. ISSN 0169-409X. doi: 10.1016/j.addr.2018.07.010.
- [8] Paola Franco and Iolanda De Marco. Nanoparticles and nanocrystals by supercritical co<sub>2</sub>-assisted techniques for pharmaceutical applications: A review. *Applied Sciences*, 11(4):1476, February 2021. ISSN 2076-3417. doi: 10.3390/app11041476.
- [9] Nedasatad Saadati Ardestani, Gholamhosseini Sodeifian, and Seyed Ali Sajadian. Preparation of phthalocyanine green nano pigment using supercritical co<sub>2</sub> gas antisolvent (gas): experimental and modeling. *Heliyon*, 6(9):e04947, September 2020. ISSN 2405-8440. doi: 10.1016/j.heliyon.2020.e04947.
- [10] Gholamhosseini Sodeifian, Seyed Ali Sajadian, and Reza Derakhsheshpour. Co<sub>2</sub> utilization as a supercritical solvent and supercritical antisolvent in production of sertraline hydrochloride nanoparticles. *Journal of CO<sub>2</sub> Utilization*, 55:101799, January 2022. ISSN 2212-9820. doi: 10.1016/j.jcou.2021.101799.
- [11] O. Singh, Z. Khanam, N. Misraand, and M.K. Srivastava. Chamomile (*matricaria chamomilla l.*): An overview. *Pharmacognosy Reviews*, 5 (9):82, 2011. ISSN 0973-7847. doi: 10.4103/0973-7847.79103.
- [12] J. Srivastava. Extraction, characterization, stability and biological activity of flavonoids isolated from chamomile flowers. *Molecular and Cellular Pharmacology*, 1(3):138–147, August 2009. ISSN 1938-1247. doi: 10.4255/mcpharmacol.09.18.
- [13] A. Orav, A. Raal, and E. Arak. Content and composition of the essential oil of *chamomilla recutita(l.) rauschert* from some european countries. *Natural Product Research*, 24(1):48–55, January 2010. ISSN 1478-6427. doi: 10.1080/14786410802560690.
- [14] M. Goto, B.C. Roy, and T. Hirose. Shrinking-core leaching model for supercritical-fluid extraction. *The Journal of Supercritical Fluids*, 9 (2):128–133, jun 1996. doi: 10.1016/s0896-8446(96)90009-1.
- [15] H. Sovová, J. Kučera, and J. Jež. Rate of the vegetable oil extraction with supercritical co<sub>2</sub>—ii. extraction of grape oil. *Chemical Engineering Science*, 49(3):415–420, 1994. ISSN 0009-2509. doi: 10.1016/0009-2509(94)87013-6.

- [16] H. Sovova, R. Komers, J. Kucuera, and J. Jezu. Supercritical carbon dioxide extraction of caraway essential oil. *Chemical Engineering Science*, 49(15), 1994. doi: 10.1016/0009-2509(94)e0058-x.
- [17] Keith D. Bartle, Anthony A. Clifford, Steven B. Hawthorne, John J. Langenfeld, David J. Miller, and Robert Robinson. A model for dynamic extraction using a supercritical fluid. *The Journal of Supercritical Fluids*, 3(3):143–149, September 1990. ISSN 0896-8446. doi: 10.1016/0896-8446(90)90039-o.
- [18] E. Reverchon. Mathematical modeling of supercritical extraction of sage oil. *AIChE Journal*, 42(6):1765–1771, June 1996. ISSN 1547-5905. doi: 10.1002/aic.690420627.
- [19] G. E.P. Box and W. J. Hill. Discrimination among mechanistic models. *Technometrics*, 9(1):57–71, February 1967. ISSN 1537-2723. doi: 10.1080/00401706.1967.10490441.
- [20] Guido Buzzu-Ferraris and Pio Forzatti. A new sequential experimental design procedure for discriminating among rival models. *Chemical Engineering Science*, 38(2):225–232, 1983. ISSN 0009-2509. doi: 10.1016/0009-2509(83)85004-0.
- [21] Bing H. Chen and Steven P. Asprey. On the design of optimally informative dynamic experiments for model discrimination in multiresponse nonlinear situations. *Industrial & Engineering Chemistry Research*, 42(7):1379–1390, February 2003. ISSN 1520-5045. doi: 10.1021/ie0203025.
- [22] J. D. Jr Anderson. *Computational fluid dynamics the basic with applications*. McGraw-Hill, 1995. ISBN 9780071132107.
- [23] J. D. Jr Anderson. *Fundamentals of Aerodynamics*. McGraw-Hill Education, 2023. ISBN 9781264151929.
- [24] N. R. Bulley, M. Fattori, A. Meisen, and L. Moyls. Supercritical fluid extraction of vegetable oil seeds. *Journal of the American Oil Chemists' Society*, 61(8):1362–1365, aug 1984. doi: 10.1007/bf02542243.
- [25] M. Spiro and M. Kandiah. Extraction of ginger rhizome: partition constants and other equilibrium properties in organic solvents and in supercritical carbon dioxide. *International Journal of Food Science & Technology*, 25(5):566–575, jun 2007. doi: 10.1111/j.1365-2621.1990.tb01116.x.
- [26] H. Sovova. Broken-and-intact cell model for supercritical fluid extraction: Its origin and limits. *The Journal of Supercritical Fluids*, 129:3–8, nov 2017. doi: 10.1016/j.supflu.2017.02.014.
- [27] Oliwer Sliczniuk and Pekka Oinas. Mathematical modelling of essential oil supercritical carbon dioxide extraction from chamomile flowers. *The Canadian Journal of Chemical Engineering*, 2024. doi: 10.1002/cjce.25557.
- [28] J. Gmehling, M. Kleiber, B. Kolbe, and J. Rarey. *Chemical Thermodynamics for Process Simulation*. Wiley, mar 2019. doi: 10.1002/9783527809479.
- [29] David Mautner Himmelblau. *Process analysis by statistical methods*. Wiley, New York, 1970. ISBN 047139985X. [by] David M. Himmelblau., Includes bibliographical references.
- [30] Yonathan Bard. *Nonlinear parameter estimation*. Acad. Press, New York [u.a.], 1974. ISBN 0120782502.
- [31] John Mandel. Fitting a straight line to certain types of cumulative data. *Journal of the American Statistical Association*, 52(280):552–566, dec 1957. doi: 10.1080/01621459.1957.10501413.

Parameter	Estimate	Std Error	95% CI Low	95% CI High
$k_{w0}$	1.222524	0.053510	1.116804	1.328244
$a_w$	4.308414	0.277930	3.759309	4.857520
$b_w$	0.972739	0.090562	0.793816	1.151661
$n_k$	3.428618	0.176948	3.079022	3.778213

**Table 1**

Error between experimental data and the challenger model

Exp.	T (°C)	P (bar)	F (kg/s)	MSE Cum.	MSE Indep.	Std Error Indep.
1	40	100	6.7e-5	0.0060	0.0007	0.0243
2	40	120	6.7e-5	0.0062	0.0011	0.0327
3	40	160	6.7e-5	0.0235	0.0019	0.0444
4	40	200	6.7e-5	0.1330	0.0046	0.0648
5	30	100	6.7e-5	0.0296	0.0011	0.0324
6	30	120	6.7e-5	0.0191	0.0014	0.0358
7	30	160	6.7e-5	0.0054	0.0013	0.0373
8	30	200	6.7e-5	0.0046	0.0027	0.0541
9	30	100	3.3e-5	0.0144	0.0002	0.0063
10	30	200	3.3e-5	0.0268	0.0013	0.0331
11	30	120	3.3e-5	0.0343	0.0011	0.0306
12	30	160	3.3e-5	0.0004	0.0005	0.0225

**Table 2**

Error between experimental data and the challenger model

	$\sigma(k_{w0})$	$\sigma(a_w)$	$\sigma(b_w)$	$\sigma(n_k)$
$\sigma(k_{w0})$	0.0029	0.0066	0.0000	0.0054
$\sigma(a_w)$	0.0066	0.0772	0.0020	0.0009
$\sigma(b_w)$	0.0000	0.0020	0.0082	-0.0004
$\sigma(n_k)$	0.0054	0.0009	-0.0004	0.0313

**Table 3**

Variance-covariance matrix, challenger model

## A. Appendix

### A.1. Parameter estimation of the challenger model

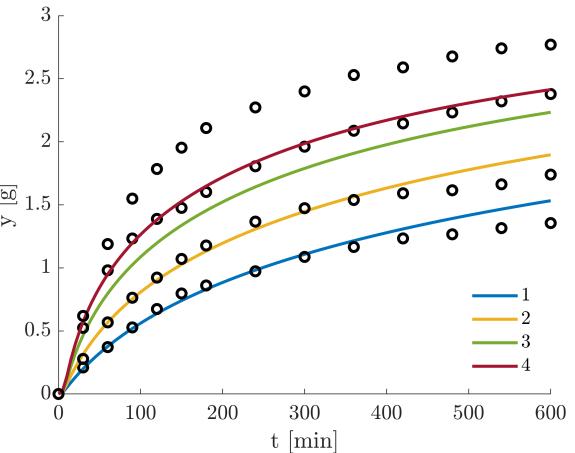
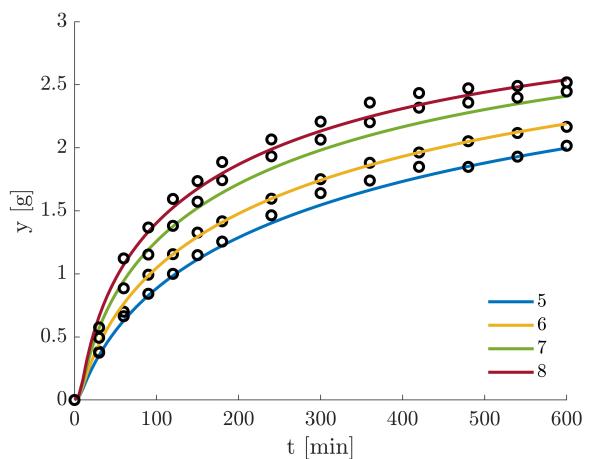
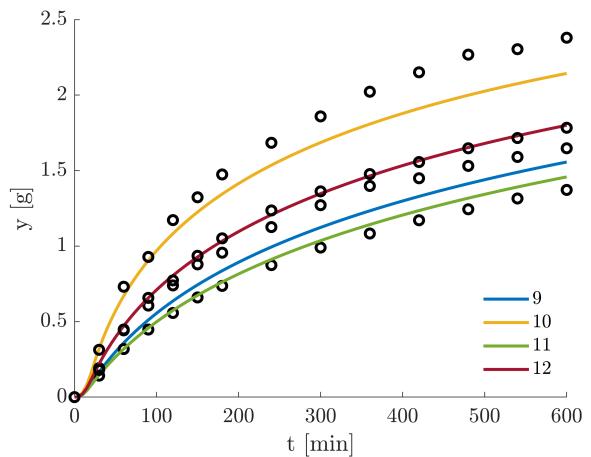
Unlike the champion model presented in Section ??, the challenger model was developed based on the same assumption regarding the kinetic and lack of saturation. Unlike the champion mode, the challenger model introduce the power low-type of structure with the polynomial decay of the overall mass transfer coefficient, as presented by Equations 25 - 27.

$$\frac{dc_s}{dt} = k_w \left( \frac{\rho_f}{800} \right)^{a_w} \left( \frac{F}{5e-5} \right)^{b_w} 10^{-4} \quad (25)$$

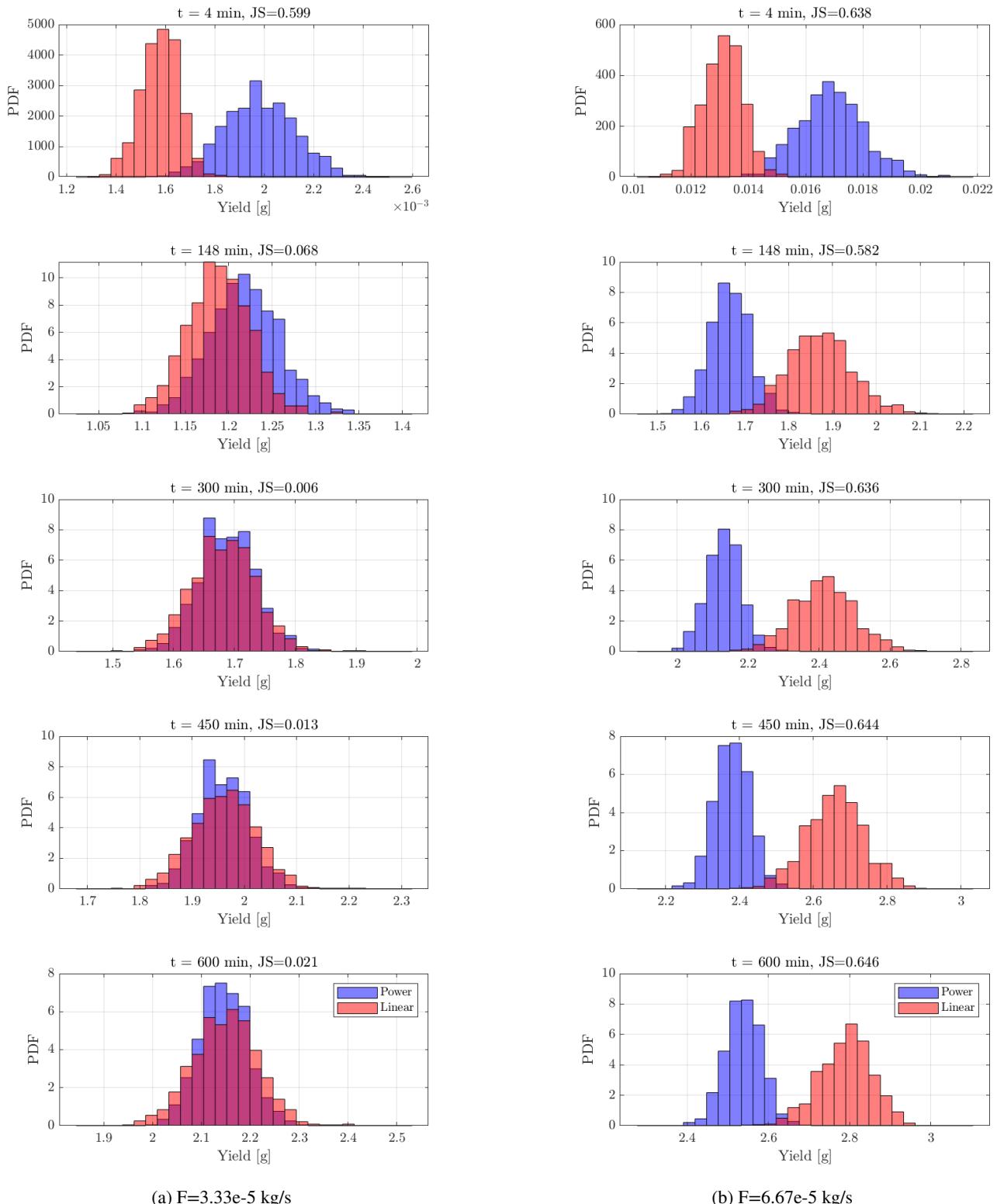
$$k_w = k_{w0} \frac{1}{(1 + \alpha_k)^{n_k}} \quad (26)$$

$$\alpha_k = 1 - \frac{c_s}{c_{s0}} \quad (27)$$

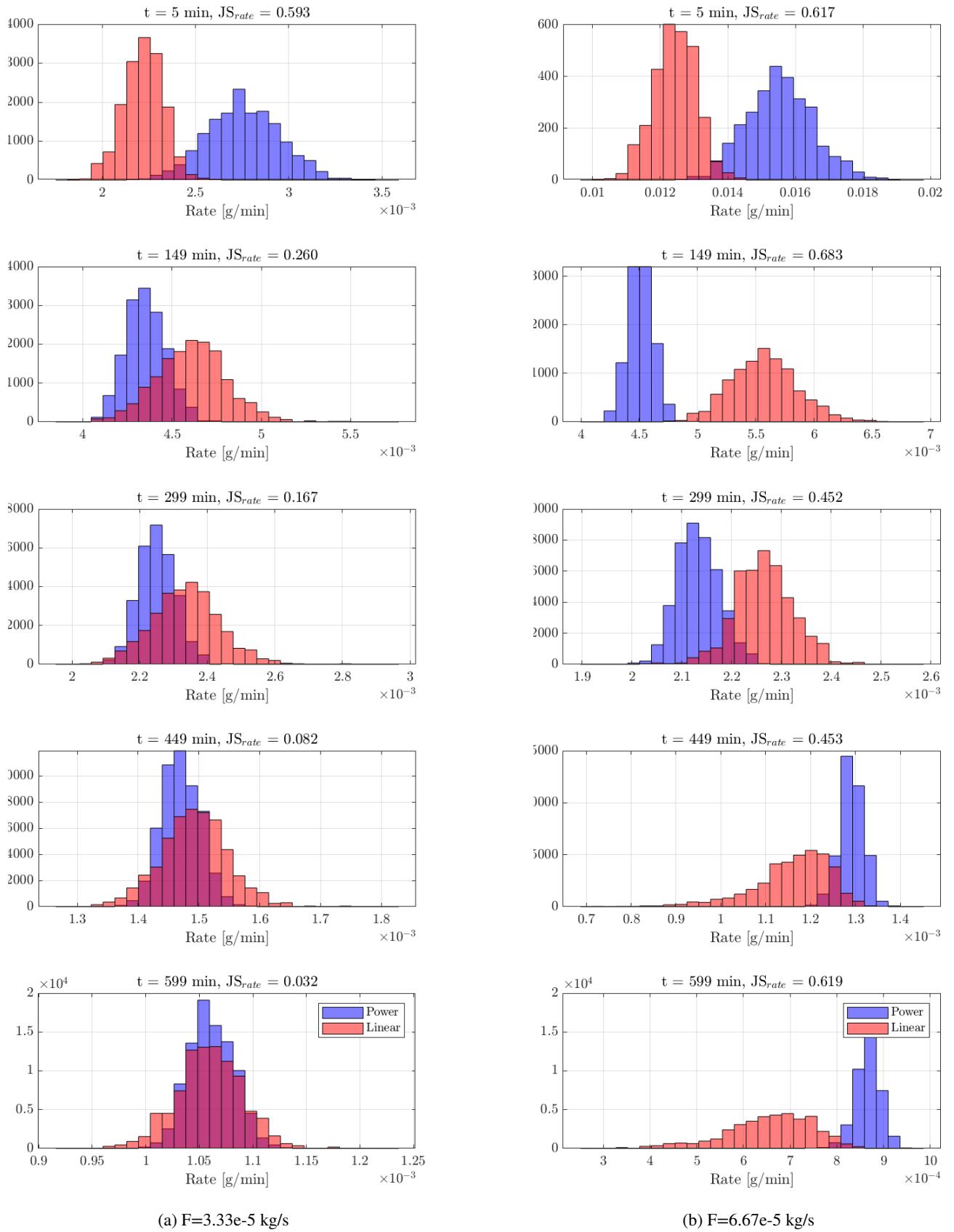
The model was fit on the entire development sample at once hence, no additional correlations are needed. The simulation results demonstrate good agreement with experimental data (Figure 7), as evidenced by the low mean squared error and standard deviation values reported in Table ?? An unbiased estimate of the error variance was computed from the residuals between predicted and observed values. The variance-covariance matrix of coefficient estimates, presented in Table 3, was obtained by combining this variance with the inverse Gram matrix of the design matrix. The p-value obtained from the Jarque-Bera test (0.064), proves that residuals appear to be normally distributed.

(a) Results at  $6.67 \cdot 10^{-5}$  kg/s and 40 °C(b) Results at  $6.67 \cdot 10^{-5}$  kg/s and 30 °C(c) Results at  $3.33 \cdot 10^{-5}$  kg/s**Figure 7:** Simulation results, challenger model

## A.2. Temporal evolution of output output distributions



**Figure 8:** Temporal evolution of cumulative distribution at  $T=30^\circ \text{C}$ ,  $P=200 \text{ bar}$

**Figure 9:** Temporal evolution of extraction rate distribution at  $T=30^\circ \text{ C}$ ,  $P=200 \text{ bar}$