Optimal design of experiment for model discrimination

Oliwer Sliczniuk[a,*], Pekka Oinas[a]

[a]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. The empirical correlations used in the model are based on the set of laboratory experiments performed under multiple constant operating conditions: $30-40°C$, $100-200$ bar, and $3.33-6.67\cdot10^{-5}$ kg/s. The model-based design of the experiments technique is applied in this work to design an experiment to discriminate between two competing process models.

1. Introduction

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. [1]. 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. Aqueous extract of chamomile is widely used to calm nerves and mitigate anxiety, hysteria, nightmares, insomnia and other sleep-related conditions, according to Srivastava [2]. Orav et al. [3] 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.

Evaluating the economic viability of the process is essential when choosing a suitable technology for essential oil extraction. Traditional methods, such as distillation and organic solvent extraction, are commonly employed but have drawbacks. Distillation, for example, involves high temperatures that can lead to the thermal degradation of heat-sensitive compounds. This limitation has led to the increased popularity of alternative techniques such as supercritical fluid extraction. Supercritical carbon dioxide is appealing thanks 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.

*Corresponding author

oliwer.sliczniuk@aalto.fi (O. Sliczniuk)

ORCID(s): 0000-0003-2593-5956 (O. Sliczniuk); 0000-0002-0183-5558 (P. Oinas)

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.

The model proposed by Reverchon et al. [4] is called the hot ball model, as it is based on an analogy to heat transfer and describes an extraction process from solid particles. This model assumes that particles contain low quantities of solute and that solubility is not a limiting factor.

The Broken-and-Intact Cell model, proposed by Sovova [5], assumes that external surfaces of particles are mechanically disrupted, allowing the solvent access to the solute in the broken cells. In contrast, the solute in intact cells remains less accessible due to higher mass transfer resistance.

Reverchon [6] formulated a fluid-solid extraction model where the solute is treated as a single component, governed by internal mass transfer resistance and omitting the effects of external mass transfer, axial dispersion and variations in fluid density and flow rate throughout the bed.

This work builds upon the linear kinetic model suggested by Reverchon [6], 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_{in}) 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.

Design of Experiments (DoE) is a structured approach that examines how various elements influence a particular result. By evaluating multiple factors at once, DoE allows the uncovering of the impacts of each element and their combinations, yielding a comprehensive comprehension of

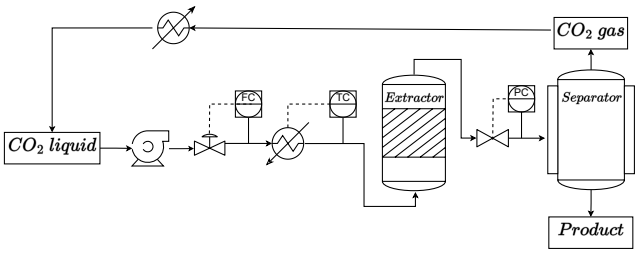


Figure 1: Process flow diagram

the entire system. DoE begins with determining an experiment's objectives and selecting the study's process factors. DoE aims to obtain the maximum information from an experimental apparatus modelled by devising experiments that will yield the most informative data in a statistical sense for use in parameter estimation and model validation.

The first ideas of DoE were introduced by Fisher [7], who described the fundamental problem of experimental design as deciding what pattern of factors combination will best reveal the properties of the response, and this response is influenced by the factors. This type of DoE views an experiment as simply connecting inputs with outputs and is therefore called a "black-box experiment design". It aims to select the combinations of factor values that will provide the most information on the input-output relationship in the presence of variation. This type's main class of statistical design techniques is the so-called factorial method. These methods are created to measure the additive effects on a response for each input factor and investigate the effects of interactions between factors. Factorial methods are unsuitable for situations where constraints exist on the output or internal states. They must also be better suited for dynamic experiments, where inputs and outputs are complex time profiles. However, this group of methods is still widely used due to its simplicity.

The factorial methods can be used for parameter screening, giving main and interaction effects of the considered factors with fewer runs. Ramandi et al. [8] applied a full factorial design for screening the extraction parameters of fatty acids from *Borago officinalis L.* flowers by SFE technique before optimisation using central composition design. Four factors: temperature, pressure, modifier volume and static extraction time were considered independent variables for full factorial design. All these factors were studied at two levels.

Caldera et al. [9] optimised extraction parameters of SFE to extract antioxidant compounds from rosemary. 2^3 full factorial design was used to select important variables before optimisation of the selected factors by Box-Behnken design. Three factors, namely temperature, pressure and static extraction time, were studied in this experiment.

As opposed to the "black-box" statistical experiment design methods, another form of optimal design has been developed, which takes explicit advantage of some knowledge of the structure underlying the system, represented by a mathematical model, in particular in the form of differential and algebraic equations. What characterises the model-based experiment design approach is :

1. the explicit use of the model equations and current parameters to predict the "information content" of the next experiment

2. the application of an optimisation framework to the solution of the resulting numerical problem

After an initial dataset has been collected and fitted to a mathematical model, the model undergoes further analysis. Additional experiments may be designed and conducted to differentiate between competing models that passed the preliminary tests. Once inadequate models are rejected, the remaining model may undergo another round of experiment design to enhance the precision of its parameters. This paper focuses on the final step of the validation procedure, known as model-based Design of Experiments (DoE), aimed at improving parameter precision. To the authors' knowledge, the model-based design of dynamic experiments has not been applied to any case of supercritical extraction. The literature review below provides examples of this technique used in crystallization and pharmacology processes.

Chung et al. [10] applied model-based experimental design to a batch crystallization process with a cooling jacket. A dynamic programming formulation minimizes the volume of a confidence hyper-ellipsoid for the estimated nucleation and growth parameters over the supersaturation profile and the seed characteristics, namely, the crystal mass, mean size, and width of the seed distribution. As a result, the accuracy of the parameter estimates can be improved by identifying the optimal temperature profile.

Duarte et al. [11] investigated compartment models incorporating Michaelis-Menten elimination kinetics for pharmacological applications. The authors designed both static and dynamic experiments for 2- and 3-compartment models using D-optimality criteria. The dynamic experiments for both models involved determining the initial concentration in the first compartment and optimizing the profile of the mass flow rate of the drug entering this compartment.

2. Materials and methods

2.1. Governing equations

Following the work of Anderson [12], 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 [13], all flows are compressible, but some of them can be treated as incompressible when the Mach number is smaller than 0.3. 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.

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

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

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

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

where ρ_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:

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

where A is the total cross-section of the extractor and ϕ 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

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

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, ϕ and ρ_f:

$$v = \frac{F}{\rho_f A\phi} \tag{6}$$

Similarly, superficial velocity may be introduced:

$$u = v\phi = \frac{F}{\rho_f A} \tag{7}$$

The fluid density ρ_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 ρ_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:

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

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:

$$\frac{\partial c_s}{\partial t} = \underbrace{r_e}_{\text{Kinetics}} \tag{9}$$

2.2.4. Kinetic term

As the solvent flows through the bed, CO_2 molecules diffuse into the pores and adsorb on the particle surface to form an external fluid film around the solid particles due to the 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 and reaches c_{pf} at the pore 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.

Bulley et al. [14] 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 $r_e\left(c_f - c_p^*(c_s)\right)$. In contrast, Reverchon

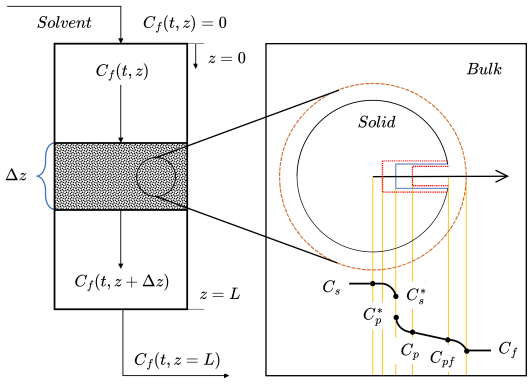


Figure 2: Mass transfer mechanism.

[6] proposes a driving force given by the difference between c_s and c_p^*. Concentration c_p^* is determined by the equilibrium relationship with c_f and the extraction rate given by Equation 10:

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

where μ is sphericity, l a characteristic dimension of particles that can be defined as $l = r/3$, r is the mean particle radius, ρ_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. [14], 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:

$$c_f^* = k_p c_s \tag{11}$$

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. According to Spiro and Kandiah [15], k_p can be expressed through the mass partition coefficient k_m:

$$k_m = \frac{k_p \rho_s}{\rho_f} \tag{12}$$

According to Reverchon [6], the kinetic term becomes

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

2.2.5. Uneven solute's distribution in the solid phase

Following the idea of the Broken-and-Intact Cell (BIC) model (Sovova [16]), 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 γ, as given by Equation 14:

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

where Υ describes the curvature of the decay function. Equation 15 describes the final form of the kinetic term:

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

The γ 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.

Alternatively, the decay function γ can be interpreted by referring to the Shrinking Core model presented by Goto et al. [17], where the particle radius changes as the amount of solute in the solid phase decreases. As the particle size decreases due to dissolution, the diffusion path increases, which makes the diffusion slower and reduces the value of the diffusion coefficient. This analogy can be applied to Equation 14 to justify the application of a varying diffusion coefficient.

2.2.6. Empirical correlations

The empirical correlations for D_i and Υ were derived by article 1 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.

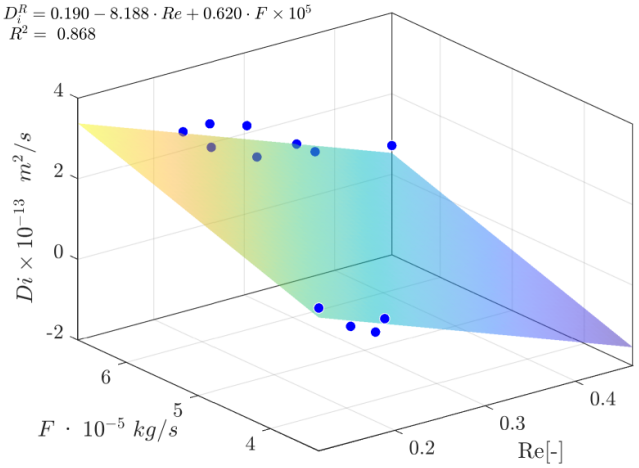


Figure 3: Multiple linear regression $D_i^R = f(Re, F)$

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

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

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

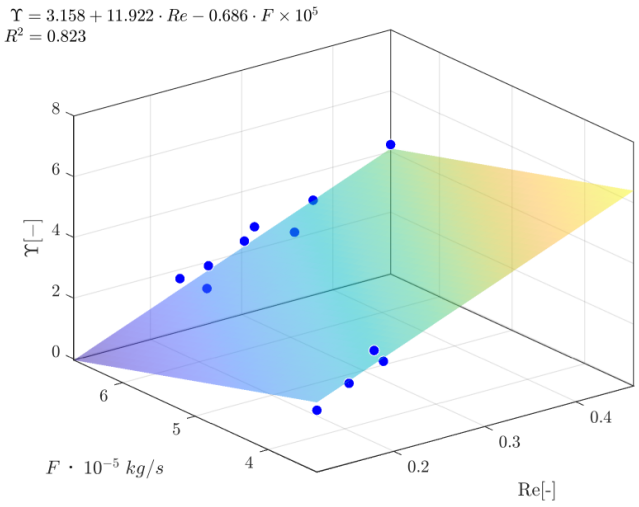


Figure 4: Multiple linear regression $\Upsilon = f(Re, F)$

a mathematical function that characterizes the deviation of a thermodynamic property (enthalpy, entropy, and internal energy) of a real substance from that of an ideal gas at the same temperature and pressure. As presented by Gmehling et al. [18], for the Peng-Robinson equation of state, the enthalpy departure function is defined by Equation 17.

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

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

Equation 17 requires an reference sate, 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 fining procedure to repeated at every time step to find a temperature profile along spatial direction z.

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

2.2.8. Pressure term

As explained in Chapters 2.1, at Low-Mach number conditions, the thermodynamic pressure is nearly constant in space 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, which describes 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. [19]. 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:

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

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:

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

where $x \in \mathbb{R}^{N_x=3N_z+2}$ and $\Theta \in \mathbb{R}^{N_\Theta=N_\theta+N_u}$, N_θ 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.

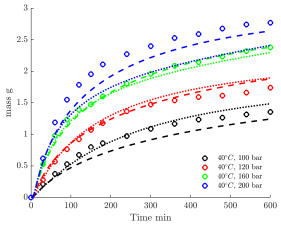
3. General function approximators

3.1. Radial Basis Function

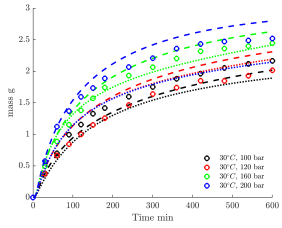
An alternative approach to the first principle modelling of physical process is to apply general function approximators and train them based on a dataset. Such an approach has an advantage of not pre-assuming a structure of a model, which results in higher flexibility of a model. On the other hand, this come with a cost of higher number of parameters to be fitted and choosing appropriate function approximator. In this work, a Radial Basis Function (RBF) is used to define $\frac{dc_s}{dt}$ based on a dataset. RBF is a sum real-valued functions (so called kernels) δ whose value depends only on the distance between the input and some fixed point, called a center c, so that $\delta(x) = \hat{\delta}(||x-c||)$. The distance is usually Euclidean distance, although other metrics are sometimes used. Sums of radial basis functions are typically used to approximate given functions $y(x) = \sum_{i=1}^{N} w_i \delta(||x-c_i||)+b$. Where N corresponds to the number of kernels, w to weight in the summation and b is a bias. The kernel can be defined as a Gaussian, Inverse quadratic, Inverse multi-quadratic, Polyharmonic spline etc. In this work, the two-dimensional Gaussian is used. All the kernels are assumed to have the same shape, which means the all have the same widths in the same direction. Following observation from article 1, the two independent variables are normalized concentration of the solute in the solid phase, defined as $\tilde{c}_s(t) = 1 - \left(\frac{c_s(t)}{c_{s0}}\right)$, and the Reynolds number.

i	1	2	3	4	5	6
$\tilde{c}^c_{si}$	-0.294	1.465	-0.303	1.124	-0.025	-0.424
Re^c_i	0.410	0.281	0.057	0.913	1.185	0.269
w_i	2.001	0.541	6.569	-1.208	-0.663	6.123
σ_1^2	0.433					
σ_1^2	0.051					
b	-0.131					

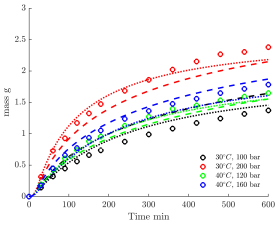
Table 1
Parameters of the RBF



(a) Simulation results at $6.67 \cdot 10^{-5}$ kg/s



(b) Simulation results at $6.67 \cdot 10^{-5}$ kg/s



(c) Simulation results at $3.33 \cdot 10^{-5}$ kg/s

Figure 5: Comparison of models against the dataset

$$\frac{dc_s}{dt} = \sum_{i=1}^{N} w_i \delta(||x-c_i||)+b = \sum_{i=1}^{N} w_i \exp\left(-\frac{\left(\tilde{c}_s(t)-\tilde{c}^c_{si}\right)^2}{2\sigma_1^2} - \frac{\left(Re(t)-Re^c_i\right)^2}{2\sigma_2^2}\right)+b \tag{22}$$

where $\tilde{c}^c_{si}$ and Re^c_i correspond to centrers of each Gaussian kernels as defined above. σ_1 na σ_2 corresponds to the width of each kernel in direction of $\bar{c}_{si}$ and Re_i, respectively. The unknowns of this equation are N, w_i, $\tilde{c}^c_{si}$ and Re^c_i. If N is pre-selected, then the total number of unknowns can calculated as $3N + 3$. The new process model is defined by substituting Equation 9 with 22. The parameter estimation procedure follows article 1. The obtained parameters are presented in Table 1.

Good agreement between the simulation results and the dataset can be observed, as presented by calculated mean square error and standard deviation as presented in Table 2.

Experiment	1	2	3	4	5	6	7	8	9	10	11	12
Mean squared error	0.0049	0.0100	0.0042	0.1255	0.0401	0.0368	0.0061	0.1292	0.0084	0.0109	0.0028	0.0130
Standard deviation of error	0.0705	0.0619	0.0334	0.1422	0.0791	0.0704	0.0779	0.1313	0.0466	0.1069	0.0287	0.0477

Table 2
Error between experimental data and model predictions

4. Optimal design of experiment for model discrimination

Let's consider two probability distributions $p(y_1)$ and $p(y_2)$ where each represents the Gaussian probability density function for a model.

$$p(Y|y_1) = \prod_{i=1}^{n_t} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{\sum \left(Y - y_1(t,x,p)\right)^2}{2\sigma_1^2}\right) \tag{23}$$

If the ratio of two probability distributions is considered to indicate the measure of similarity, then $\ln\left(\frac{p(Y|y_1)}{p(Y|y_2)}\right)$ becomes a measure of the odds in favour of choosing hypothesis H_1 ($p(Y|y_1)$ is a true model) over hypothesis H_2 ($p(Y|y_2)$ is a true model). Alternatively, the ratio can be interpreted as the information in favour of hypothesis H_1 as opposed to the hypothesis H_2. The so-called 'weight of evidence' or expected information in favour of choosing H_1 over H_2 can be defined through the Kullback–Leibler divergence and is represented by:

$$I(1:2) = \int_{-\infty}^{\infty} p(Y|y_1) \ln\left(\frac{p(Y|y_1)}{p(Y|y_2)}\right) dY \tag{24}$$

The above equation can be written more explicitly as

$$\begin{aligned} I(1:2) &= \int_{-\infty}^{\infty} p(Y|y_1) \left[\sum_{i=1}^{n_t} \left(\ln\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{(Y_i - y_{1i})^2}{2\sigma_1^2} + \frac{(Y_i - y_{2i})^2}{2\sigma_2^2}\right)\right] dY \\ &= \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \ln\left(\frac{\sigma_2}{\sigma_1}\right)\right) dY - \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \frac{(Y_i - y_{1i})^2}{2\sigma_2^2}\right) dY \\ &+ \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \frac{(Y_i - y_{2i})^2}{2\sigma_2^2}\right) dY \end{aligned} \tag{25}$$

As the expected error is constant for all the measurements: $\mathbb{E}[(Y_i - y_{1i})^2] = \mathbb{E}[\sigma_1^2]$, which simplifies the equation:

$$\begin{aligned} I(1:2) &= \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \ln\left(\frac{\sigma_2}{\sigma_1}\right)\right) dY - \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(\frac{1}{2} p(Y|y_1)\right) dY \\ &+ \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \frac{(Y_i - y_{2i})^2}{2\sigma_2^2}\right) dY \end{aligned} \tag{26}$$

The first two terms can be simplified by taking a constant in front of integrals and by noticing that $\int p(x)dx = 1$.

$$I(1:2) = n_t \ln\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{n_t}{2} + \sum_{i=1}^{n_t} \frac{1}{2\sigma_2^2} \int_{-\infty}^{\infty} \left(p(Y|y_1)\left(Y_i^2 - 2Y_i y_{2i} + y_{2i}^2\right)\right) dY \tag{27}$$

$$\begin{aligned} I(1:2) &= n_t \ln\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{n_t}{2} + \sum_{i=1}^{n_t} \frac{1}{2\sigma_2^2} \int_{-\infty}^{\infty} \left(p(Y|y_1) \times Y_i^2\right) dY \\ &- \sum_{i=1}^{n_t} \frac{2y_{2i}}{2\sigma_2^2} \underbrace{\int_{-\infty}^{\infty} \left(p(Y|y_1) \times Y_i\right) dY}_{\text{expected value}=y_{1i}} + \sum_{i=1}^{n_t} \frac{y_{2i}^2}{2\sigma_2^2} \underbrace{\int_{-\infty}^{\infty} p(Y|y_1) dY}_{=1} \end{aligned} \tag{28}$$

The remaining integral can be solved by recognizing that $\sigma^2 = \int_{-\infty}^{\infty} X^2 p(X) dX - \mathbb{E}[(X)]^2$, which leads to $\int_{-\infty}^{\infty} Y_i^2 p(Y|y_{1i}) dY = y_{1i}^2 + \sigma_1^2$.

The final form of the Kullback–Leibler divergence becomes:

$$\begin{aligned} I(1:2) &= n_t \ln\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{n_t}{2} + \sum_{i=1}^{n_t} \frac{y_{1i}^2 + \sigma_1^2}{2\sigma_2^2} \sum_{i=1}^{n_t} \frac{2y_{2i}y_{1i}}{2\sigma_2^2} + \sum_{i=1}^{n_t} \frac{y_{2i}^2}{2\sigma_2^2} \\ &= n_t \ln\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{n_t}{2} + \frac{n_t}{2\sigma_1^2} + \sum_{i=1}^{n_t} \frac{1}{\sigma_2^2} \left(y_{1i} - y_{2i}\right)^2 \end{aligned} \tag{29}$$

$$I(2:1) = n_t \ln\left(\frac{\sigma_1}{\sigma_2}\right) - \frac{n_t}{2} + \frac{n_t}{2\sigma_2^2} + \sum_{i=1}^{n_t} \frac{1}{2\sigma_1^2} \left(y_{1i} - y_{2i}\right)^2 \tag{30}$$

While Kulback-Liebler divergence is a statistical distance, it is not a metric on the space of probability distributions. While metrics are symmetric and generalize linear distance, satisfying the triangle inequality, divergences are asymmetric in general and generalize squared distance. In general, $I(1:2) \neq I(2:1)$. By taking into account that the Kullbacl-Liebler divergence is additive for independent distribution, the function J for model discrimination can be defined as

$$\begin{aligned} J(1,2) = I(1:2) + I(2:1) &= \int_{-\infty}^{\infty} [p(Y|y_2) - p(Y|y_1)] \ln \frac{p(Y|y_1)}{p(Y|y_2)} dy \\ &= \frac{n_t(\sigma_1^2 - \sigma_2^2)}{2\sigma_1^2\sigma_2^2} + \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2} \times \sum_{i=1}^{n_t} \left(y_{1i} - y_{2i}\right)^2 \end{aligned} \tag{31}$$

The first term of J is independent of changes in y_1 and y_2, while the second term is equivalent to the sum of squared differences between two model outputs. By maximizing J, the y_1 and y_2 are spread apart. Although both models were fitted with the same dataset, they employ structurally different extraction kinetic terms. These structural differences lead to different outputs, particularly in regions not covered by the dataset.

5. Results

6. Conclusions

References

[1] 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.

[2] 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.

[3] A. Orav, A. Raal, and E. Arak. Content and composition of the essential oil ofchamomilla recutita(l.) rauschert from some european countries. Natural Product Research, 24(1):48–55, January 2010. ISSN 1478-6427. doi: 10.1080/14786410802560690.

[4] 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.

[5] 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.

[6] 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.

[7] R.A Fisher. The Design of Experiments. Olivier & Boyed, 1935.

[8] N.F. Ramandi, N.M. Najafi, F. Raofie, and E. Ghasemi. Central composite design for the optimization of supercritical carbon dioxide fluid extraction of fatty acids from borago officinalis l. flower. Journal of Food Science, 76(9), October 2011. ISSN 1750-3841. doi: 10.1111/j.1750-3841.2011.02394.x.

[9] G. Caldera, Y. Figueroa, M. Vargas, and D.T. Santos. Optimization of supercritical fluid extraction of antioxidant compounds from venezuelan rosemary leaves. International Journal of Food Engineering, 8(4), October 2012. ISSN 1556-3758. doi: 10.1515/1556-3758.1953.

[10] S.H. Chung, D.L. Ma, and R.D. Braatz. Optimal model-based experimental design in batch crystallization. Chemometrics and Intelligent Laboratory Systems, 50(1):83–90, January 2000. ISSN 0169-7439. doi: 10.1016/s0169-7439(99)00049-0.

[11] B. Duarte, A. Atkinson, J. Granjo, and N. Oliveira. Calculating d-optimal designs for compartmental models with a michaelis–menten elimination rate. Journal of Process Control, 83:88–101, November 2019. ISSN 0959-1524. doi: 10.1016/j.jprocont.2019.09.001.

[12] J. D. Jr Anderson. Computational fluid dynamics the basic with applications. McGraw-Hill, 1995. ISBN 9780071132107.

[13] J. D. Jr Anderson. Fundamentals of Aerodynamics. McGraw-Hill Education, 2023. ISBN 9781264151929.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] J. Gmehling, M. Kleiber, B. Kolbe, and J. Rarey. Chemical Thermodynamics for Process Simulation. Wiley, mar 2019. doi: 10.1002/9783527809479.

[19] 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.

[20] E. Walter and L. Pronzato. Identification of parametric models from experimental data. Communications and control engineering. Springer, London, 2010. ISBN 9781849969963.