Computers and Chemical Engineering 181 (2024) 108542

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Targeting fidelity of pharmaceutical systems models by optimization of precision on parameter estimates

Margherita Geremia [a], Giulio Cisco [a], Samir Diab [b], Gabriele Bano [c], Fabrizio Bezzo [a,*]

[a] Department of Industrial Engineering, CAPE-Lab – Computer-Aided Process Engineering Laboratory, University of Padova, Via Marzolo 9, Padova, PD 35131, Italy
[b] GSK, Park Road, Ware SG12 0DP, United Kingdom
[c] GSK, 1250 S Collegeville Rd, Collegeville, PA 19426, United States

ARTICLE INFO

Keywords:
Systems model
Pharmaceutical development
Parameter estimation
Prediction fidelity
Optimization

ABSTRACT

Quantitative models have gained momentum to drive the development of pharmaceutical processes. The assessment of the prediction fidelity of these models is key to provide interpretability of process phenomena and to enable decision-making. Evaluating parametric uncertainty is paramount when the focus is on systems models, which combine different sub-models together, and, thus, parameters related to previous units may strongly impact the prediction of one final output. A framework is proposed to assess reliability in model predictions, where the precision of parameter estimates is explicitly optimized to target pre-set tolerance requirements on process key performance indicators and product critical quality attributes. A direct compression systems model for the manufacturing of oral solid dosage products is used as a case study. Results show that the proposed methodology is effective at guaranteeing the target model fidelity and at quantifying the maximum acceptable uncertainty in the estimates of model parameters.

1. Introduction

Pharmaceutical manufacturing processes typically involve multiple unit operations, which are connected by material and energy streams, and can be suitably represented by a process flowsheet (Boukouvala et al., 2012). Mathematical models can be used to describe the evolution of physical and chemical phenomena along the manufacturing line as well as to predict the pharmacodynamic and pharmacokinetic behavior of the drug product *in vivo* or *in vitro* (Daryaee and Tonge, 2019). The overall model consisting of all sub-models (i.e., all mathematical equations describing the relevant phenomena occurring in a single unit operation, or in a single functional test unit to assess the effect of the drug product) is typically known as *systems model* (Avraam et al., 1998).

Systems models describing pharmaceutical manufacturing and product performance are extremely useful to support process development, drug design and decision making (Braakman et al., 2022; Destro and Barolo, 2022). Examples of systems modeling approaches for industrial applications have been presented in literature. Bano et al. (2022) streamlined the development of an industrial dry granulation process for immediate release (IR) tablets; Moreno-Benito et al. (2022) proposed an integrated model combining first-principles and data-driven approaches of a continuous direct compression (DC) manufacturing line for the production of IR tablets; White et al. (2022) presented a systems model of a pharmaceutical tablet manufacturing process to assess whether a given drug product be manufactured using dry granulation or DC. Systems models have been successfully used to support manufacturing of active pharmaceutical ingredient (API) in both continuous (Diab et al., 2022a) and batch (Diab et al., 2022b) modes. Monaco et al. (2023) adopted a systems approach to study the impact of operating conditions and material properties in wet granulation manufacturing.

Evaluating the prediction reliability with respect to model key performance indicators (KPIs) and critical quality attributes (CQAs) – which here will be generically called key indicators (KIs) – is paramount in understanding whether the available model is suitable for the intended industrial purpose – where it may be used to predict CQAs that are important for the therapeutic efficacy of a drug and patient safety. The application of quantitative and statistical metrics for the assessment of model prediction uncertainty is fundamental to enhance the systematic use of (systems) models for process development and decision-making (Bai et al., 2019; Zineh, 2019).

The accuracy of predictions of a (systems) model with respect to model KIs is also known as model fidelity (Geremia et al., 2023), and strongly depends on the precision of model parameter estimates. Not only does the prediction fidelity of a selected KI depend on the

* Corresponding author.
E-mail address: fabrizio.bezzo@unipd.it (F. Bezzo).

https://doi.org/10.1016/j.compchemeng.2023.108542
Received 10 October 2023; Received in revised form 23 November 2023; Accepted 1 December 2023
Available online 2 December 2023

Nomenclature

Acronyms

API	active pharmaceutical ingredient
CI	confidence interval
CQA	critical quality attribute
DAE	differential algebraic equations
DC	direct compression
GSA	global sensitivity analysis
HPLC	high performance liquid chromatography
IR	immediate release
KI	key indicator
KPI	key performance indicator
MBDoE	model-based design of experiments
MC	Monte Carlo
OSD	oral solid dosage
PLS	partial least squares
SQP	sequential quadratic programming
USP	Unites States Pharmacopeia
UV	ultraviolet

Greek letters

β	total fraction of tensile strength that can be lost due to lubrication
γ	lubrication rate constant [dm^{-1}]
δ	Dirac delta function
$\boldsymbol{\delta}$	vector of pre-set tolerances on the prediction of $\boldsymbol{K}$
ε	average porosity of the swollen product [–]
$\boldsymbol{\varepsilon}$	set of parameter uncertainty of the whole systems model
$\varepsilon_{i,\max,b}$	boundary maximum uncertainty of estimated parameter θ_i maximizing the objective function
$\varepsilon_{i,\max}$	maximum uncertainty of estimated parameter θ_i maximizing the objective function
$\boldsymbol{\varepsilon}_{M_i}$	set of parameter uncertainty of sub-model M_i
$\dot{e}$	tablet erosion rate [m s^{-1}]
$\boldsymbol{\theta}$	set of parameters of the whole systems model
$\boldsymbol{\theta}_{M_i}$	set of parameters of sub-model M_i
λ	swelling rate [s^{-1}]
μ	liquid viscosity [Pa s]
$\boldsymbol{\xi}$	set of relative uncertainty of model parameters $\boldsymbol{\theta}$
ξ_i	relative uncertainty of estimated parameter θ_i
$\boldsymbol{\xi}_{\mathbf{max}}$	set of relative uncertainty of model parameters $\boldsymbol{\theta}$ maximizing the objective function
ρ_p	density of particles [kg m^{-3}]
τ	total stress [MPa]
τ_{or}	average tablet tortuosity [–]
ϕ	shape factor of particles [–]

Latin letters

a_1	parameter of extended Kushner model [MPa]
a_2	parameter of extended Kushner model [–]
a_i^h	auxiliary variable in the optimization problem
a_{sf}	Kawakita model parameter [–]
A_t	tablet surface area [m^2]
b_1	extended Kushner parameter [–]
b_2	extended Kushner parameter [–]
B_{API}	rate of release of API [s^{-1}]
b_{sf}	Kawakita model parameter [MPa^{-1}]
C_2	Peppas and Colombo parameter [MPa]
C_3	Peppas and Colombo parameter [MPa]
c_{API}	bulk concentration of API [kg m^{-3}]
c_{sat}	saturation concentration of API [kg m^{-3}]
d_h	tablet hydraulic diameter [m]
E	elastic constant (1) [–]
$E(\cdot)$	expectation operation
G_0	elastic constant (2) [MPa]
h	stochastic scenario
H_{coat}	thickness of the coating layer [m]
K	extent of lubrication [dm]
k_{API}	mass transfer coefficient of API [$(m^3 \times kg^{-1})^{n_{API}}$ s^{-1}]
$\boldsymbol{K}$	set of key indicators of the whole systems model
$\bar{\boldsymbol{K}}$	vector of the target values of the whole systems model KIs
$\boldsymbol{K}_{M_i}$	set of key indicators of sub-model M_i
l	particle size [–]
$l_{0,API}$	particle size at the beginning of the process [m]
LC	percentage of label content [–]
M	number of sub-models of the systems model
M_t	tablet mass [kg]
$M_{t,0}$	initial tablet mass [kg]
n	swelling parameter [–]
N	number of parameters combinations via MC simulations
n_{API}	order of dissolution [–]
N_K	number of model key indicators
N_θ	number of model parameters
N_{tot}	total number of stochastic scenarios
P	compaction pressure [MPa]
p_c	capillary pressure [Pa]
P_d	water penetration depth [m]
$R_{API,l}$	particle dissolution coefficient [m s^{-1}]
S_i	first-order sensitivity index [–]
S_p	shape factor of pores [–]
$S_{i,\ TOT}$	total effect sensitivity index [–]
t	time [s]
$\boldsymbol{t_{ref}}$	reference t-value
$T_{t/2}$	half tablet thickness [m]
TS	tensile strength [MPa]
TS_0	tensile strength at zero porosity [MPa]
$\boldsymbol{u}_{M_i}$	vector of input variables of sub-model M_i
$V(K_{M_i})$	variance of model KI K_{M_i}
V_c	coating volume [m^3]
V_m	liquid volume in the vessel [m^3]
w_l	liquid content in the tablet [–]
x_{API}	mass fraction of API [–]
$\boldsymbol{x}_{M_i}$	vector of state variables of sub-model M_i
$\boldsymbol{y}_{M_i}$	vector of output measured variables of sub-model M_i

parameters of the specific sub-model, but it is also affected by the precision of parameter estimates related to the previous sub-models impacting the KI of interest. For instance, let us consider the case of a pharmaceutical manufacturing process to produce oral solid dosage (OSD) products via direct compression (Wang et al., 2017). The API and excipients are first fed to a co-mill; lubricant is then added to improve flowability and to stop the powder from sticking to the tablet press die walls and punch; mixing occurs in a continuous blender to improve the homogeneity of the blend, which is finally fed to the tablet press unit. The prediction of tablet hardness is affected by the lubrication extent attained in the blender; therefore, it is expected that uncertainty in the parameters of the sub-models for the previous units will impact the prediction of the tablet hardness.

Characterizing the fidelity of a model requires a quantification of the prediction uncertainty of model outputs. Methodologies based on Monte Carlo (MC) simulations (Fishman, 1995) have typically been exploited to evaluate the relationship between model parameters and model outputs. They need to simulate the process using a sufficiently high

number of combinations of model parameters, which are typically *pseudo*-random or Sobol sequences (Kucherenko et al., 2015). To mention some recent examples, Briskot et al. (2019) assessed the prediction fidelity of chromatography models by generating samples of parameters values using a Bayesian Markov Chain MC approach, while Demetriades et al. (2022) quantified *in vitro* cancer drug pharmacodynamics via MC modeling. Recently, some approaches have been proposed to evaluate regions of model reliability in terms of prediction error. Quaglio et al. (2018) used a decision function to evaluate the expected model reliability in unexplored regions of the experimental design space; Dasgupta et al. (2021) proposed a kriging interpolation technique to map the mean squared error of model prediction. Cenci et al. (2023) presented an explorative model-based design of experiments (MBDoE) method to reduce model prediction uncertainty by using a mapping of model prediction variance.

With specific reference to the pharmaceutical sector, a general procedure for the evaluation of the prediction fidelity of (systems) models was proposed by Geremia et al. (2023). The framework was successfully demonstrated to: (*i*) assess model prediction fidelity using standardized model evaluation methods, i.e., to quantify the impact of model parameter uncertainty on the selected model KIs, and (*ii*) optimize the experimental campaigns for the parameter estimation in quantitative (systems) models. The methodology relies on a linear partial least-squares (PLS) regression model, which is built considering the relationship between model parameters and outputs. Thus, if the original model is strongly nonlinear with respect to the relationship between parameters and KIs, the uncertainty region may be represented ineffectively.

In this work, we present an alternative approach for the quantification of model parameter impact on the prediction fidelity of KIs. Specifically, we exploit the available systems model discussed in Geremia et al. (2023), and apply an optimization approach to explicitly quantify the required level of precision of parameter estimates that allows targeting a pre-set fidelity on the prediction of model KIs. The proposed procedure relies on an optimization approach and does not introduce any linearization.

The same DC systems model for manufacturing of OSD products considered in Geremia et al. (2023) will be used as a case study. The systems model comprises the following sub-models: (1) tablet press unit operation, (2) tablet disintegration test unit, and (3) *in vitro* dissolution test unit.

The remainder of the article is organized as follows. In Section 2, we introduce the framework for the assessment of model prediction fidelity, and we briefly outline the general scope of each step of the procedure. In Section 3, we thoroughly describe the optimization approach to quantify the parameter impact on the prediction of model KIs. In Section 4, we give details regarding the case study, while in Section 5 we implement the methodology, and critically discuss the results that are obtained. Final remarks will conclude the study.

2. Problem statement

In mathematical terms, a generic systems model is comprised of a number (M) of sub-models representing the different sub-systems (M_1, M_2,… M_M), and the relationship between the unit inputs and outputs can be described by a set of differential and algebraic equations (DAEs). Each sub-model i can be mathematically described by the following DAEs:

$$\begin{cases} \boldsymbol{f}_{M_i}(\boldsymbol{x}_{M_i}(t), \dot{\boldsymbol{x}}_{M_i}(t), (\boldsymbol{\theta}_{M_i} \pm \boldsymbol{\varepsilon}_{M_i}), \boldsymbol{u}_{M_i}(t), t) = 0 \\ \boldsymbol{y}_{M_i}(t) = \boldsymbol{g}_{M_i}(\boldsymbol{x}_{M_i}(t),\ \dot{\boldsymbol{x}}_{M_i}(t), (\boldsymbol{\theta}_{M_i} \pm \boldsymbol{\varepsilon}_{M_i}), \boldsymbol{u}_{M_i}(t), t) \\ \boldsymbol{K}_{M_i}(t) = \boldsymbol{h}_{M_i}(\boldsymbol{x}_{M_i}(t),\ \dot{\boldsymbol{x}}_{M_i}(t), (\boldsymbol{\theta}_{M_i} \pm \boldsymbol{\varepsilon}_{M_i}), \boldsymbol{u}_{M_i}(t), t) \end{cases}, \quad (1)$$

with $i = 1, 2, \ldots, M$ where $\boldsymbol{x}_{M_i}$, $\boldsymbol{u}_{M_i}$, and $\boldsymbol{\theta}_{M_i}$, refer to sub-model i and are (respectively) the vector of state variables, the vector of input variables, and the vector of the model parameter estimates. $\boldsymbol{\varepsilon}_{M_i}$ is the corresponding vector of absolute values of parameter uncertainties, which quantifies the uncertainty in the estimated values of model parameters; in other words, it is a measure of the absolute parameter uncertainties due to incomplete/imperfect knowledge about the values of the estimates (Sin et al., 2009). If model parameters are estimated perfectly, the values of $\boldsymbol{\varepsilon}_{M_i}$ are zero. Also note that here we assume that the model is the perfect representation of the actual system, and therefore uncertainty is only related to precision in the model parameters (i.e. there is no *structural* model mismatch).

$\boldsymbol{y}_{M_i}$ and $\boldsymbol{K}_{M_i}$ are the vector of measured output variables and the vector of KIs for the sub-model i, respectively (vector field $\boldsymbol{f}_{M_i}$ represents all other DAEs in sub-model i). We assume that the vector $\boldsymbol{u}_{M_i}$ of input variables can be manipulated to vary the KIs. A KI can be equal to output y_{j,M_i} or be derived from combinations of outputs. With reference to the whole systems model, we will denote the set of all model parameter estimates, $\boldsymbol{\theta}$, the corresponding set of parameter uncertainty $\boldsymbol{\varepsilon}$, and predicted model key indicators, $\boldsymbol{K}$, as follows:

$$\boldsymbol{\theta} = \left[\boldsymbol{\theta}_{M_1},\ \boldsymbol{\theta}_{M_2},\ldots,\ \boldsymbol{\theta}_{M_M}\right]^T \quad (2)$$

$$\boldsymbol{\varepsilon} = \left[\boldsymbol{\varepsilon}_{M_1},\ \boldsymbol{\varepsilon}_{M_2},\ldots,\ \boldsymbol{\varepsilon}_{M_M}\right]^T, \quad (3)$$

$$\boldsymbol{K} = \left[\boldsymbol{K}_{M_1},\ \boldsymbol{K}_{M_2},\ldots,\ \boldsymbol{K}_{M_M}\right]^T. \quad (4)$$

Assessing the fidelity of a given systems model requires quantifying how the contributions of all model parameters impact the prediction of the KIs, namely, to compute the maximum values of the elements in vector $\boldsymbol{\varepsilon}$ such that the pre-set accuracy on the prediction of model KIs is attained. It is worth highlighting that, in general, not only does model fidelity depend on the parameters of one specific sub-system model, but it also relies on the fidelities of the parameters of all sub-system models impacting the unit being investigated.

The required fidelity on model KIs can be mathematically formulated as:

$$(\boldsymbol{K} - \bar{\boldsymbol{K}})^2 \leq \boldsymbol{\delta}^2, \quad (5)$$

where $\bar{\boldsymbol{K}}$ is the vector of the target values of model KIs, and $\boldsymbol{\delta}$ is the vector of pre-set tolerances on the prediction of $\boldsymbol{K}$, i.e., the maximum error in the prediction of $\boldsymbol{K}$ with respect to the target $\bar{\boldsymbol{K}}$.

The presented work aims at computing the maximum values of elements in vector $\boldsymbol{\varepsilon}$ which guarantee that predicted model KIs fall within the range of pre-set tolerances, namely, determining the maximum uncertainty of each model parameter, θ_i, which ensures the pre-set accuracy requirements.

3. Methodology

In this section, we thoroughly describe the novel optimization approach for the assessment of the parameter impact on the prediction of model KIs. It is worth highlighting that we aim at optimizing the precision on parameter estimates such that we ensure that the parametric uncertainty meets the pre-set tolerance requirements on model KIs. Thus, we purposely exclude a detailed study on preliminary steps, i.e., (*i*) model identifiability and parameter ranking, (*ii*) design and execution of experiments, and (*iii*) parameter estimation, for which the reader may refer to more general articles (e.g., McLean and McAuley 2012, Braakman et al. 2022, Geremia et al. 2023). Techniques used in the presented methodology to fulfil model identifiability and parameters ranking are briefly discussed in Appendix A.

3.1. Parameter impact on the fidelity of model predictions

Let us first introduce the relative parameter uncertainty, ξ_i, as the ratio between the parameter uncertainty, ε_i, and the absolute value of the current estimated parameter, θ_i:

$$\xi_i = \frac{\varepsilon_i}{|\theta_i|} \ . \tag{6}$$

This ensures that uncertainties have a common scale. The objective function is here defined as the product between all the relative parameter uncertainties:

$$obj(\boldsymbol{\xi}) = \prod_{i=1}^{n} \xi_i \ . \tag{7}$$

The optimization problem consists in maximizing $obj(\boldsymbol{\xi})$, while satisfying the fidelity constraint imposed by Eq. (5):

$$\max(obj(\boldsymbol{\xi})) \ s.t. \ (\boldsymbol{K} - \bar{\boldsymbol{K}})^2 \leq \boldsymbol{\delta}^2. \tag{8}$$

Values of $\boldsymbol{\xi}$ maximizing the objective function, which we will call $\boldsymbol{\xi}_{\mathbf{max}}$, are computed through a nonlinear sequential quadratic programming (SQP) optimization (Boggs and Tolle, 1995). Note that the formulation of the optimization problem requires setting upper bounds (UBs) for the maximum values of elements in $\boldsymbol{\xi}$. When the relative maximum uncertainty, ξ_i for parameter θ_i hits the UB, (i.e., $\xi_{\max,i} = \mathrm{UB}_i$), here we assume that parameter θ_i is non-influential towards the prediction of model KIs, and any level of uncertainty on that parameter is acceptable.

$\varepsilon_{i,\max,\ b} = \xi_{i,\ \max} |\theta_i|$ is the maximum parameter uncertainty, where $\xi_{i,\ \max}$ is computed by solving the optimization problem (8), which ensures the pre-set requirement on the prediction fidelity of the KI, K_{M_j} (i.e., the pre-set tolerance δ_j^2). $\varepsilon_{i,\max,\ b}$ must be both subtracted and added to the current estimate of parameter θ_i in order to define the boundaries of parametric uncertainty, i.e., the lower and upper bounds within which the prediction fidelity is acceptable. Note that the relationship between model parameters and tolerance on KIs might be nonlinear, and, thus, real boundaries might be asymmetric. However, here we adopt a symmetric approach to conservatively compute $\varepsilon_{i,\max,\ b}$ as the minimum value between lower and upper uncertainties at boundaries with respect to the current estimate θ_i. This is also coherent with the way in which uncertainty in the estimated θ_i is typically evaluated, i.e., referring to the corresponding confidence interval (CIs), which represents the symmetric range of values with respect to the current estimate θ_i within which that parameter is expected to fall (Dekking et al., 2005).[1]

When N_θ model parameters simultaneously affect the prediction fidelity of the KIs, parameter uncertainties can be combined in 2^{N_θ} ways, based on the two possible signs (+ or −) of the contribution $\varepsilon_{i,\max,b}$ to the current estimate of θ_i, i.e., the lower and upper bounds for each model parameter. Practically, 2^{N_θ} scenarios must be evaluated during the optimization routine in Eq. (8). Let us consider one generic scenario (h) among the 2^{N_θ} possibilities, with $1 \leq h \leq 2^{N_\theta}$. Boundary values of each model parameter in scenario h, θ_i^h, can be expressed as:

$$\theta_i^h = \theta_i\left(1 + \ \xi_i a_i^h\right), \tag{9}$$

where a_i^h is an auxiliary variable used to account for the two possible signs (+ or −) of the contribution $|\varepsilon_{i,\max,b}|$ to the current estimate of θ_i:

$$a_i^h = \{-1; +1\} \ s.t. \ \boldsymbol{a}^h \neq \boldsymbol{a}^{k \neq h}, \quad i = 1, 2, \ldots, N_\theta; h = 1, 2, \ \ldots, 2^{N_\theta}. \tag{10}$$

If the relationship between model parameters and KIs is non-monotonic, the $\varepsilon_{i,\max,b}$ values that are computed by solving Eq. (8) may not represent the actual boundaries. Fig. 1 exemplifies such a condition, under the hypothesis that the prediction requirement of a certain model KI, K_{M_j}, depends on one parameter only (e.g., θ_i). Note that this is one-dimensional example is just shown for visualization, not a limitation of the method.

To tackle this potential issue, we simulate the model for a number $N_{tot} - 2^{N_\theta}$ of additional scenarios, where model parameter θ_i is still computed using Eq. (9), while auxiliary variables in Eq. (10) are redefined as follows:

$$a_i^h = \begin{cases} \{-1; +1\} & if \ 1 \leq h \leq 2^{N_\theta} \\ uniform(-1, \ 1) & if \ 2^{N_\theta} < h \leq N_{tot} \end{cases} \ s.t. \ \boldsymbol{a}^h \neq \boldsymbol{a}^{k \neq h}, \tag{11}$$

where $uniform(-1, +1)$ indicates a uniform distribution of values in the range $[-1; +1]$, and N_{tot} is the total number of scenarios which are evaluated. N_{tot} should be selected such that a sufficiently high number of scenarios are evaluated (Kucherenko et al., 2015); in this work, $N_{tot} = 10^4$. Practically, the evaluation of additional scenarios is aimed at checking that the constraints on KIs are satisfied for any scenario h such that $\theta_i - \ \varepsilon_{i,\max,\ b} \leq \theta_i^h \leq \ \theta_i + \varepsilon_{i,\max,\ b} \ \forall \ i = 1, \ 2, \ \ldots N_\theta$. If the constraints are not satisfied for at least one stochastic scenario, a new $\varepsilon_{i,\max} < \varepsilon_{i,\max,b}$ is set. This is selected as the maximum uncertainty obtained for the maximum scenario, h^*, satisfying the tolerance requirements on model KIs for any value of parameter θ_i^h in the range $\theta_i - |\xi_i a_i^{h^*}| \leq \theta_i^h \leq \theta_i + |\xi_i a_i^{h^*}|$.

From now on, we will generically indicate with $\varepsilon_{i,\max}$ the maximum allowable uncertainty for parameter θ_i.

3.2. Assessing the impact of parameter uncertainty on the fidelity of model predictions

Model parameters are estimated using a maximum likelihood estimator (Bard, 1974). To assess whether the current precision of model parameter estimates is sufficient to guarantee all the requirements on model KIs, we quantify uncertainty on parameter estimates by referring to their corresponding 95 % CIs, which is a threshold value commonly adopted to define the variability of an estimate (Dekking et al., 2005). Since the estimated value, θ_i, is the central value in the range defined by its CI, optimization results $\varepsilon_{i,\max}$ should be compared to the half of the corresponding CIs, i.e., (95 % CI_i)/2. For clarity purpose, the flow diagram in Fig. 2 shows the sub-steps used in the optimization approach to assess whether the current precision of parameter estimates is sufficient to attain the preset requirements on model KIs.

As an example, let us consider the case in Fig. 3a, where the value of $\varepsilon_{i,\max}$ is lower than the (95 % CI_i)/2 of parameter θ_i. Practically, the attained precision of parameter θ_i is not sufficient, and a new iteration of the procedure in Fig. 2 is required. The stop criterion is attained when the absolute value of $\varepsilon_{i,\max}$ is higher than the (95 % CI_i)/2 for each model parameter estimate θ_i, and therefore no further experimental effort is needed (Fig. 3b).

4. Case study

In this work, we consider the same DC systems model presented by Geremia et al. (2023). All equations of the systems model are reported in Appendix B. The model is applied to assess how changes in process operation will impact product performance, i.e., how varying the extent of lubrication and tablet press operation will impact the tablet's disintegration time and API dissolution profile. Assumptions under which the systems model has been built are as follows.

1. *Consistent/perfect blending.* Excipients and API powders are perfectly mixed. Therefore, the feeding and blending unit operations are not considered, and blend, content uniformity and tablet weight variability are not considered as KIs in this study.
2. *Dissolution test method.* The analytical method used to measure the *in vitro* dissolution profile of the API is discriminatory, i.e., the method can capture changes in factors that could impact the dissolution performance (i.e., different input setpoints lead to different

[1] The implementation of an asymmetric approach is feasible, but requires doubling the number of optimization variables in order to associate each model parameter θ_i with two different maximum uncertainties at the boundaries (lower and upper).

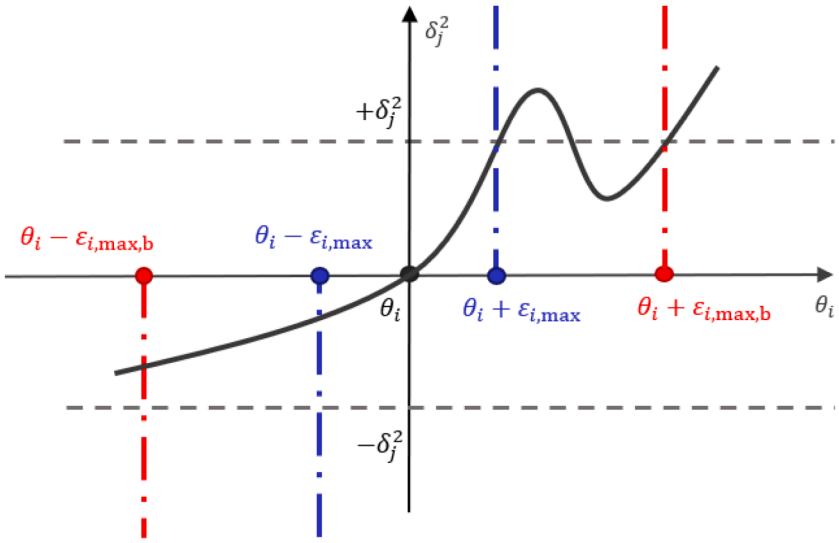

Fig. 1. Maximum allowable parameter uncertainty $\varepsilon_{i,\max}$ on θ_i to satisfy the pre-set tolerance requirement on K_{M_j} in presence of a non-monotonic dependence between θ_i and K_{M_j}.

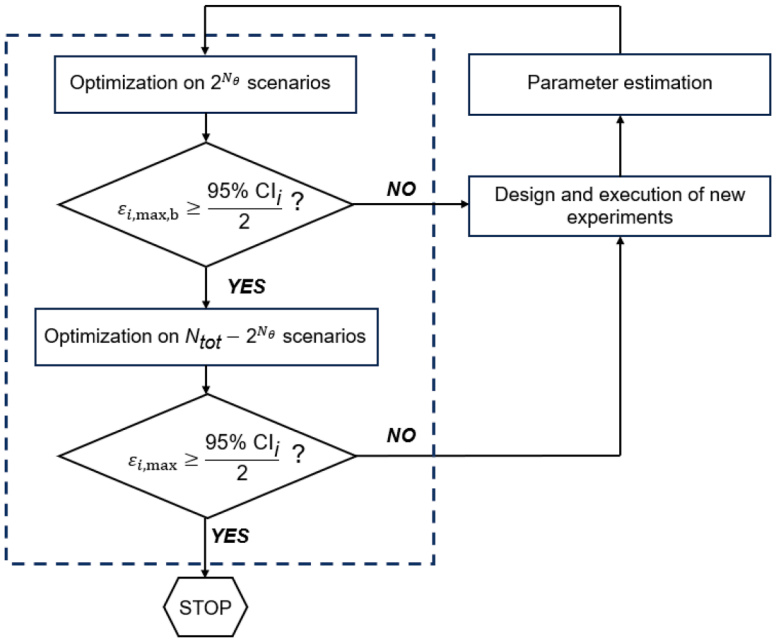

Fig. 2. Sub-steps used in the optimization approach to assess whether the current precision of parameter estimates is sufficient. The dashed box represents the optimization procedure presented in this work.

dissolution profiles). This implies that an analytical method has been developed and calibrated for dissolution testing of the considered API, such as high-performance chromatography (HPLC) or ultraviolet (UV) spectroscopy.

Based on these assumptions, the only sub-models that are considered are (*i*) the tablet press unit operation, (*ii*) the tablet disintegration test unit, and (*iii*) the *in vitro* dissolution test unit. Note that the tablet press unit operation is the only sub-system representing a unit operation in the manufacturing line. The other sub-systems concern experimental tests for the assessment of product CQAs, and require information from the tablet press sub-model, i.e., the lubrication extent K in the upstream powder blending, and the compaction pressure exerted by the press P, which are time-invariant variables.

The framework is assessed by means of an *in silico* case study. The *process* is represented by the systems model with parameters at nominal values (Table 1) as retrieved from literature (Peppas and Colombo, 1989; Nassar et al., 2021; Markl et al., 2017). The *model* is represented by the systems model with initial guesses for the parameters as in Table 1. Initial parameter guesses are randomly chosen inside a range equal to ±50 % their nominal values, and UBs of relative uncertainties ξ_i are set equal to 0.500. This is a clear simplification; however, the information needed for an exact description of parameters uncertainty is rarely available in industrial practice, and an approximate approach is often the only possible strategy. If more information is available on the actual distribution of parameter uncertainties, more rigorous approaches could be applied (Schenkendorf et al., 2018).

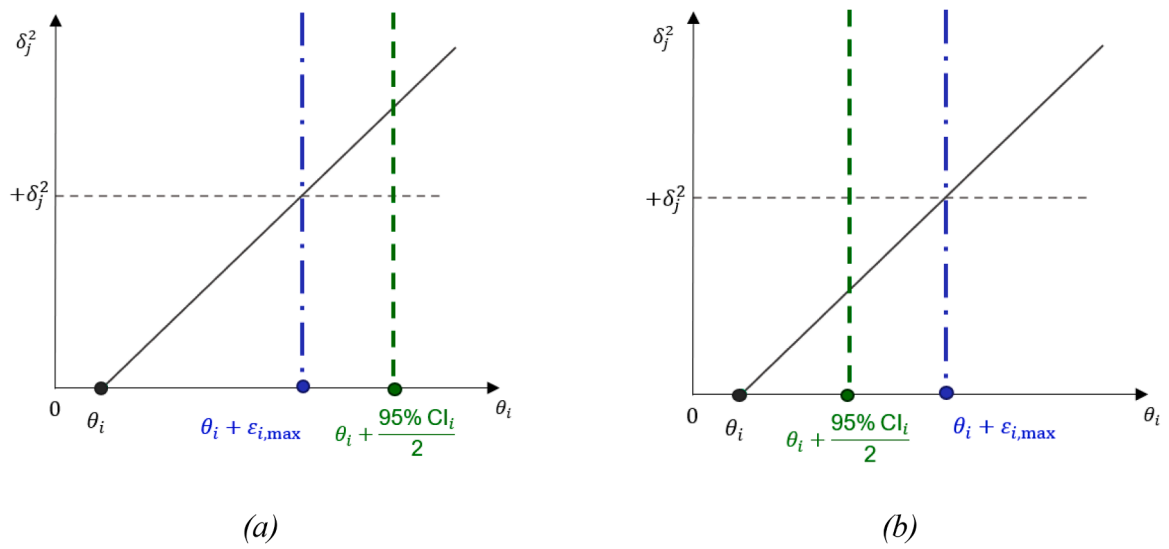

Fig. 3. Comparison between maximum uncertainty $\varepsilon_{i,max}$ allowable on parameter θ_i, and half of its 95% CI: (a) insufficient parameter precision, and (b) sufficient parameter precision.

Table 1
Nominal and initial guess values of model parameters.

Parameter	Units	Nominal	Initial guess
Tablet press unit operation			
a_1	MPa	11.04	14.81
a_2	–	1.091	1.433
a_{sf}	–	0.463	0.394
b_1	–	−8.202	−6.287
b_2	–	0.326	0.242
b_{sf}	MPa^{-1}	2.460×10^{-2}	1.710×10^{-2}
γ	dm^{-1}	1.211×10^{-3}	7.368×10^{-4}
Tablet disintegration test unit			
C_2	MPa	1.000×10^{2}	63.00
C_3	MPa	1.000×10^{2}	1.410×10^{2}
$\dot{\epsilon}$	$m\ s^{-1}$	1.000×10^{-3}	1.300×10^{-3}
n	–	0.900	1.019
S_P	–	0.524	0.688
In vitro dissolution test unit			
k_{API}	$(m^3 \times kg^{-1})^{(n_{API})}\ s^{-1}$	2.300×10^{-12}	2.996×10^{-12}
n_{API}	–	1.00	0.762

4.1. Performance targets

Three KIs are considered in this study, which correspond to each sub-model output: tablet tensile strength (which is predicted by the model equations and can be related to tablet hardness, which is often measured as part of in-process testing), tablet disintegration time, and API dissolution profile. Acceptability requirements on the KIs are defined as follows (Geremia et al., 2023):

- we consider an IR tablet, with a target *TS* of 2 MPa. We set ±0.2 MPa as the admissible tolerance with respect to the *TS* target value (i.e., ±10 % the target value), which is a typical acceptable specification range (Nassar et al., 2021);
- the target disintegration time is assumed to be 4 min. According to the USP <701> (2011) disintegration test specifications, the time limit for the IR tablet to completely disintegrate is 5 min; therefore, we set ±1 min as the admissible tolerance with respect to the target value of the disintegration time;
- the dissolution profile is monitored through the prediction of the percentage label content of the tablet (%*LC*), where %*LC* = 100 % means that all API within the tablet has dissolved into solution.%*LC* = 80 % at t = 25 min is a possible specification value for an IR tablet; however, the actual specification will depend on the specific product. We set − 15%*LC* as the admissible tolerance with respect to the target value; no overestimation is accepted for a conservative analysis.

Different workflows can be implemented when the objective is to quantify the influence of model parameters on the prediction of the KIs of a systems model (Geremia et al., 2023). We will consider: *(i)* a *modular approach*, in which the KIs of all units are targeted sequentially, and model parameters are estimated on a sub-system basis; *(ii)* a *global approach*, in which the KIs of all units are targeted simultaneously, and the parameters of all sub-system models impacting the KIs are considered at the same time.

4.2. Computational details

All activities were performed on an Intel Core I7-11850H CPU@2.50 GHz processor with 64.0 GB RAM. We used gPROMS v.7.0.7 for all tasks. Performance of the optimization routine is the most time-demanding step of the entire procedure, and depends on the model complexity and on the number of parameters that are considered. The required time to perform the optimization step using the modular approach was: (*i*) few seconds for the model for the tablet press unit operation, (*ii*) few minutes for the model for the tablet disintegration test unit, (*iii*) few minutes for the model for the *in vitro* dissolution test unit. The required time to perform the optimization step using the global approach was ~15 min. Therefore, even if the modular approach consists of more iterations than the global one (14 vs. 8), its total computational time is shorter (~35 min vs. ~120 min).

5. Results

In this section, we discuss the results regarding the parameter impact on the fidelity of model predictions by applying the optimization approach described in Section 3.

Results from preliminary model identifiability and parameters ranking using Sobol's global sensitivity analysis (GSA) are summarized in Table 2; details on the criteria for identifying the most and least influential parameters are reported in Appendix A.

Table 2
Results from preliminary model identifiability and parameters ranking.

Sub-model	Most influential parameters	Less influential Parameters
Tablet press unit operation	a_{sf}, b_1	a_2, γ
Tablet disintegration test unit	a_{sf}, b_1, b_{sf}, n, S_P	a_2, γ, C_2, C_3, $\dot{\epsilon}$
In vitro dissolution test unit	a_{sf}, b_1, b_{sf}, n, S_P, k_{API}	a_2, b_2, γ, C_2, C_3, $\dot{\epsilon}$

5.1. Parameter impact on the fidelity of model predictions

5.1.1. Modular approach

All the KIs are targeted sequentially, while model parameters are estimated on a unit operation basis.

We first focus on the model for the tablet press unit operation, where the KI of interest is *TS*. MBDoE (Franceschini and Macchietto, 2008) is first performed to organize the experimental campaign to maximize the information for the most relevant model parameters (i.e., a_{sf} and b_1) by acting on the operating variables P and K. In this work, we use the A-optimal criterion, which minimizes the dimensions of the enclosing box around the joint confidence region. Optimal values of design variables are computed through a SQP optimization. Data of *TS* are then used to estimate all the parameters for the current unit. Estimated values of model parameters together with their 95 % CIs and *t*-values (5 % confidence level) – which are used for assessing the precision in their estimation – are reported in Table 3.

To determine whether the current precision of model parameter estimates is sufficient to guarantee the fidelity requirements on *TS*, we refer to their corresponding (95 % CIs)/2 and compare the computed values with results of the optimization routine, as shown in Table 4. Recalling that the stopping criterium is attained if all (95 % CI_i)/2 are lower than the absolute value of maximum uncertainties $\varepsilon_{i,\,\max}$ given by the optimizer, we observe that the precision of parameter γ is not sufficient. This is quite an interesting result: parameter γ was demonstrated to have little influence on the KI (see Appendix A); however, the optimization procedure shows that for the attained precision of the estimates of the most influential parameters, an excessive uncertainty on γ would jeopardize quality of model predictions, which are not good enough to achieve the desired fidelity. Also, it is important to note that the maximum allowable uncertainty (e.g., $\varepsilon_{\gamma,\max}$ for parameter γ) depends on the attained precision for all other parameters, and cannot be set a-priori. To meet the pre-set requirements on *TS* prediction, the iterative cycle of Fig. 2 is repeated; the MBDoE procedure is now performed to maximize the experimental information needed to estimate parameter γ.

Eight iterations are needed to reach the required model fidelity with respect to *TS* prediction, i.e., a total of eight experiments need to be performed for the tablet press unit operation.

Estimated values of model parameters and statistics after the final iteration are collected in Table 5, while the successful assessment of the fidelity towards the *TS* prediction is presented in Table 6, where all (95 %CI_i)/2 are lower than the corresponding maximum uncertainties, $\varepsilon_{i,\,\max}$. Note that higher parameter precision on the most influential parameters (i.e., a_{sf}, b_1) leads to the relaxation of the precision requirements on less influential ones, e.g., parameter γ. Moreover, it can be observed that the precision of parameter γ after the final iteration does not improve (Table 4 vs. Table 6). This may be due to numerical reasons or may suggest that there is some correlation between parameters.

The next unit is the tablet disintegration test, where the KI of interest is the disintegration time. According to the general procedure in Fig. 2, MBDoE is first performed to organize the experimental campaign in order to maximize the information for the most relevant model parameters (i.e., n and S_P) by acting on the operating variables P and K. Disintegration data are, then, used to estimate all the parameters for the current unit (Table 7). Results indicate a poor level of precision for parameters C_2, C_3, $\dot{\epsilon}$, and S_p – their *t*-values are lower than the reference value, and the correspondent 95 % CIs exceed ±50% the parameter nominal value.

Results from the estimation activity are compared with the maximum uncertainties given by the optimizer (Table 8). It can be first read that the computed relative maximum uncertainty for parameters C_2, C_3, and $\dot{\epsilon}$ are equal to the UB, i.e., $\xi_{C_2,\,\max} = \xi_{C_1,\,\max} = \xi_{\dot{\epsilon},\,\max} = 0.500$. The result suggests that in this case we do not need their precise estimation to guarantee the pre-set requirements on the prediction of the tablet disintegration time, and confirms that full model identifiability may be unnecessary for the purpose of achieving high model fidelity, as large uncertainty on those parameters (which have been previously ranked as having low influence) does not produce large uncertainty on the prediction of the KI of interest. Therefore, we can exclude those parameters from the assessment of model fidelity, and consider their uncertainty to be equal to ±50 % of their current value. Conversely, parameters n and S_P require higher precision.

MBDoE is applied to increase the precision of parameters n and S_P by acting on the design variables K and P. Disintegration time data are then

Table 3
Tablet press unit operation. Estimated values of model parameters with their 95 % CIs and t-values: first iteration.

Parameter	Units	Nominal	Estimated	95 % CI	*t*-value
a_1	MPa	11.04	11.93	0.210	56.95
a_2	–	1.091	1.449	9.535×10^{-2}	15.19
a_{sf}	–	0.463	0.413	4.230×10^{-3}	97.61
b_1	–	−8.202	−6.235	0.105	59.66
b_2	–	0.326	0.205	1.758×10^{-2}	11.67
b_{sf}	MPa^{-1}	2.460×10^{-2}	1.890×10^{-2}	3.891×10^{-4}	48.56
γ	dm^{-1}	1.211×10^{-3}	7.365×10^{-3}	2.050×10^{-4}	3.59
					t_{ref}=1.943

Table 4
Tablet press unit operation. First estimation-optimization iteration: comparison between optimization results and current parameter uncertainties. Parameters which do not satisfy the stopping criteria are highlighted in boldface.

Parameter	Units	$\xi_{i,\max}$ [–]	$\varepsilon_{i,\max}$	(95 %CI_i)/2
a_1	MPa	1.358×10^{-2}	0.162	0.105
a_2	–	6.046×10^{-2}	8.761×10^{-2}	4.768×10^{-2}
a_{sf}	–	6.230×10^{-3}	2.573×10^{-3}	2.115×10^{-3}
b_1	–	9.328×10^{-3}	5.816×10^{-2}	5.250×10^{-2}
b_2	–	9.834×10^{-2}	2.016×10^{-2}	8.790×10^{-3}
b_{sf}	MPa^{-1}	1.363×10^{-2}	2.576×10^{-4}	1.946×10^{-4}
γ	dm^{-1}	5.536×10^{-3}	4.077×10^{-5}	1.025×10^{-4}

Table 5
Tablet press unit operation. Estimated values of model parameters with their 95 % CIs and t-values: last (eight) iteration.

Parameter	Units	Nominal	Estimated	95 % CI	*t*-value
a_1	MPa	11.04	11.09	0.179	62.08
a_2	–	1.091	1.088	0.110	9.85
a_{sf}	–	0.463	0.455	4.256×10^{-3}	1.068×10^{2}
b_1	–	−8.202	−7.961	0.103	59.89
b_2	–	0.326	0.321	1.454×10^{-2}	22.11
b_{sf}	MPa^{-1}	2.460×10^{-2}	2.445×10^{-2}	3.772×10^{-4}	51.03
γ	dm^{-1}	1.211×10^{-3}	1.202×10^{-3}	3.118×10^{-4}	3.86
					t_{ref}=1.895

Table 6
Tablet press unit operation. Final estimation-optimization iteration: comparison between optimization results and current parameter uncertainties.

Parameter	Units	$\xi_{i,\max}$ [–]	$\varepsilon_{i,\max}$	(95 %CI$_i$)/2
a_1	MPa	9.468×10^{-3}	0.105	8.950×10^{-2}
a_2	–	5.076×10^{-2}	5.523×10^{-2}	5.500×10^{-2}
a_{sf}	–	4.678×10^{-3}	2.129×10^{-3}	2.128×10^{-3}
b_1	–	8.353×10^{-3}	6.650×10^{-2}	5.150×10^{-2}
b_2	–	2.738×10^{-2}	8.791×10^{-3}	7.270×10^{-3}
b_{sf}	MPa^{-1}	9.796×10^{-3}	2.395×10^{-4}	1.886×10^{-4}
γ	dm^{-1}	0.135	1.623×10^{-4}	1.559×10^{-4}

Table 7
Tablet disintegration test unit. Estimated values of model parameters with their 95 % CIs and t-values: first iteration. † = 95 % CI larger than ±50 % the parameter nominal value. * = precision is not statistically satisfactory.

Parameter	Units	Nominal	Estimated	95 % CI	*t*-value
C_2	MPa	1.000×10^2	52.23	7.475×10^5 †	6.987×10^{-5} *
C_3	MPa	1.000×10^2	99.93	4.195×10^5†	6.060×10^{-4} *
$\dot{e}$	m s^{-1}	1.000×10^{-3}	1.480×10^{-3}	0.373†	3.969×10^{-3} *
n	–	0.900	0.905	0.184	4.92
S_P	–	0.524	0.535	0.547†	0.978 * t_{ref}=1.647

Table 8
Tablet disintegration test unit. First estimation-optimization iteration: comparison between optimization results and current parameter uncertainties. ** = maximum relative uncertainty equal to the upper bound (UB) of the preset range of variation. Parameters which do not satisfy the stopping criteria are highlighted in boldface.

Parameter	Units	$\xi_{i,\max}$ [–]	$\varepsilon_{i,\max}$	(95 %CI$_i$)/2
C_2 **	MPa	0.500	26.12	3.738×10^5
C_3 **	MPa	0.500	49.97	2.098×10^5
$\dot{e}$ **	m s^{-1}	0.500	7.402×10^{-4}	0.187
n	–	1.855×10^{-2}	1.679×10^{-2}	9.200×10^{-2}
S_P	–	0.106	5.686×10^{-2}	0.274

used to estimate all the parameters of the current unit – although C_2, C_3, and $\dot{e}$ are not impacting model fidelity. Five iterations, i.e., five experiments for the current unit, are necessary to reach the required model fidelity with respect to the KI prediction. Results after the final iteration are collected in Table 9, from which it is verified that we do not require satisfactory statistical precision of parameters C_2, C_3, and $\dot{e}$ – the correspondent *t*-values are smaller than the reference, and CIs still exceed ±50 % of the values of parameter estimates.

The successful assessment of the fidelity towards the prediction of the tablet disintegration time is reported in Table 10, where (95 %CIs)/2

Table 9
Tablet disintegration test unit. Estimated values of model parameters with their 95 % CIs and t-values: final iteration. † = 95 % CI larger than ±50 % the parameter nominal value. * = precision is not statistically satisfactory.

Parameter	Units	Nominal	Estimated	95 % CI	*t*-value
C_2	MPa	1.000×10^2	88.72	3.895×10^5 †	2.278×10^{-3} *
C_3	MPa	1.000×10^2	99.17	2.174×10^3 †	4.561×10^{-2} *
$\dot{e}$	m s^{-1}	1.000×10^{-3}	9.835×10^{-4}	6.964×10^{-2} †	1.412×10^{-2} *
n	–	0.900	0.903	3.491×10^{-2}	35.86
S_P	–	0.524	0.533	0.110	4.853 t_{ref}=1.652

Table 10
Tablet disintegration test unit. Final estimation-optimization iteration: comparison between optimization results and current parameter uncertainties. ** = maximum relative uncertainty equal to the upper bound (UB) of the preset range of variation.

Parameter	Units	$\xi_{i,\max}$ [–]	$\varepsilon_{i,\max}$	(95 %CI$_i$)/2
C_2 **	MPa	0.500	44.36	1.948×10^5
C_3 **	MPa	0.500	49.59	1.087×10^3
$\dot{e}$ **	m s^{-1}	0.500	4.918×10^{-4}	3.482×10^{-2}
n	–	2.111×10^{-2}	1.906×10^{-2}	1.746×10^{-2}
S_P	–	0.109	5.848×10^{-2}	5.500×10^{-2}

Table 11
In vitro dissolution test unit. Estimated values of model parameters with their 95 % CIs and t-values: first iteration.

Parameter	Units	Nominal	Estimated	95 % CI	*t*-value
k_{API}	$(m^3 \times kg^{-1})^{(n_{API})}$ s^{-1}	2.300×10^{-12}	2.226×10^{-12}	9.432×10^{-16}	2.360×10^3
n_{API}	–	1.00	1.10	1.521×10^{-3}	9.538×10^2 t_{ref}=1.652

Table 12
In vitro dissolution test unit. Final estimation-optimization iteration: comparison between optimization results and current parameter uncertainties.

Parameter	Units	$\xi_{i,\max}$ [–]	$\varepsilon_{i,\max}$	(95 %CI$_i$)/2
k_{API}	$(m^3 \times kg^{-1})^{n_{API}}$ s^{-1}	1.870×10^{-3}	4.163×10^{-15}	4.716×10^{-16}
n_{API}	–	4.951×10^{-3}	5.446×10^{-3}	7.605×10^{-4}

of both n and S_P are lower than the corresponding maximum uncertainties. Again, parameters C_2, C_3, and $\dot{e}$ are non-influential.

We finally move to the *in vitro* dissolution test unit. MBDoE is applied to increase the precision of both parameters k_{API} and n_{API}. Data on API dissolution profile are used for the estimation activity; results are reported in Table 11. One single iteration (i.e., one experiment for the current unit) is necessary to reach the required model fidelity with respect to the KI prediction, as shown in Table 12: (95 %CIs)/2 of both n_{API} and k_{API} are lower than the corresponding maximum uncertainties.

5.1.2. Global approach

In the global approach, all KIs are targeted simultaneously, and the parameters of all sub-models are considered at the same time. MBDoE is first applied to maximize the information of the most influential model parameters towards the prediction of all model KIs, i.e., a_{sf}, b_{sf}, b_1 (tablet press sub-model), n, S_P (tablet disintegration sub-model), k_{API}, and n_{API} (*in vitro* dissolution sub-model). Design variables K and P are again used in the MBDoE problem. Measured data of *TS*, disintegration time, and API dissolution profile are used to estimate all the parameters of the systems model (Table 13).

To assess whether the current precision of model parameter estimates is sufficient to guarantee the fidelity requirements on all model KIs, we compare results of the optimized $\varepsilon_{i,\max}$ with the (95 % CIs)/2. Results in Table 14 show that the required tolerance is not met due to unsatisfactory precision in parameters γ (sub-model for the tablet press unit operation) and n (sub-model for the tablet disintegration test unit). Therefore, the iterative cycle of Fig. 2 is repeated; MBDoE procedure is performed to maximize the information of parameters γ and n. Design

Table 13
Global approach focusing on all KIs simultaneously. Estimated values of model parameters with their 95 % CIs and t-values: first iteration. † = 95 % CI larger than ±50 % the parameter nominal value. * = precision is not statistically satisfactory.

Parameter	Units	Nominal	Estimated	95 % CI	*t*-value
a_1	MPa	11.04	11.94	0.214	38.95
a_2	–	1.091	1.120	0.119	9.45
a_{sf}	–	0.463	0.453	3.524×10^{-3}	1.285×10^{2}
b_1	–	−8.202	−7.862	0.105	56.87
b_2	–	0.326	0.314	1.596×10^{-2}	19.66
b_{sf}	MPa^{-1}	2.460×10^{-2}	2.416×10^{-2}	3.878×10^{-4}	62.30
γ	dm^{-1}	1.211×10^{-3}	7.649×10^{-4}	3.124×10^{-4}	3.57
					t_{ref}=1.943
C_2	MPa	1.000×10^{2}	66.06	2.462×10^{5} †	2.683×10^{-4} *
C_3	MPa	1.000×10^{2}	1.071×10^{2}	5.712×10^{4} †	1.874×10^{-3} *
$\dot{\epsilon}$	$m\ s^{-1}$	1.000×10^{-3}	1.451×10^{-3}	4.548×10^{-2} †	3.189×10^{-2} *
n	–	0.900	0.901	6.649×10^{-2}	13.54
S_P	–	0.524	0.538	0.102	5.29
					t_{ref}=1.686
k_{API}	$(m^3 \times kg^{-1})^{n_{API}}\ s^{-1}$	2.300×10^{-12}	2.247×10^{-12}	9.512×10^{-16}	2.362×10^{3}
n_{API}	–	1.00	1.07	1.152×10^{-3}	9.281×10^{2}
					t_{ref}=1.652

Table 14
Global approach focusing on all KIs simultaneously. First estimation-optimization iteration: comparison between optimization results and current parameter uncertainties. ** = maximum relative uncertainty equal to the upper bound (UB) of the preset range of variation. Parameters which do not satisfy the stopping criteria are highlighted in boldface.

Parameter	Units	$\xi_{i,max}$ [–]	$\varepsilon_{i,max}$	(95 %CI_i)/2
a_1	MPa	1.189×10^{-2}	0.142	0.107
a_2	–	7.601×10^{-2}	8.513×10^{-2}	5.950×10^{-2}
a_{sf}	–	5.709×10^{-3}	2.586×10^{-3}	1.762×10^{-3}
b_1	–	9.866×10^{-3}	7.757×10^{-2}	5.250×10^{-2}
b_2	–	4.516×10^{-2}	1.418×10^{-2}	7.980×10^{-3}
b_{sf}	MPa^{-1}	8.647×10^{-2}	2.089×10^{-4}	1.939×10^{-4}
γ	dm^{-1}	7.968×10^{-2}	6.095×10^{-5}	1.562×10^{-4}
C_2**	MPa	0.500	33.03	1.231×10^{5}
C_3**	MPa	0.500	53.54	2.856×10^{4}
$\dot{\epsilon}$**	$m\ s^{-1}$	0.500	7.253×10^{-4}	2.274×10^{-2}
n	–	2.102×10^{-2}	1.894×10^{-2}	3.325×10^{-2}
S_P	–	0.118	6.334×10^{-2}	5.100×10^{-2}
k_{API}	$(m^3 \times kg^{-1})^{nAPI}\ s^{-1}$	1.853×10^{-3}	4.163×10^{-15}	4.756×10^{-16}
n_{API}	–	4.982×10^{-3}	5.331×10^{-3}	5.760×10^{-4}

variables *K* and *P* are used in the MBDoE problem. For each iteration, three different experiments for the three units need to be simultaneously performed. Measured output variables *TS*, disintegration data, and API dissolution profile are used to re-estimate all model parameters.

Eight iterations are needed to reach the required fidelity of the KIs, i.e., eight experimental runs would need to be performed for each unit simultaneously, i.e., 24 experiments altogether. Estimated values of model parameters after the final iteration are reported in Table 15,

Table 15
Global approach focusing on all KIs simultaneously. Estimated values of model parameters with their 95 % CIs and t-values: last iteration. † = 95 % CI larger than ±50 % the parameter nominal value. * = precision is not statistically satisfactory.

Parameter	Units	Nominal	Estimated	95 % CI	*t*-value
a_1	MPa	11.04	11.05	0.161	82.74
a_2	–	1.091	1.078	8.811×10^{-2}	12.23
a_{sf}	–	0.463	0.457	3.332×10^{-3}	1.370×10^{2}
b_1	–	−8.202	−8.097	0.108	74.68
b_2	–	0.326	0.326	1.023×10^{-2}	31.82
b_{sf}	MPa^{-1}	2.460×10^{-2}	2.494×10^{-2}	3.772×10^{-4}	66.11
γ	dm^{-1}	1.211×10^{-3}	1.210×10^{-3}	1.172×10^{-4}	10.32
					t_{ref}=1.812
C_2	MPa	1.000×10^{2}	88.13	4.542×10^{4} †	1.940×10^{-3} *
C_3	MPa	1.000×10^{2}	1.406×10^{2}	6.918×10^{3} †	2.032×10^{-2} *
$\dot{\epsilon}$	$m\ s^{-1}$	1.000×10^{-3}	9.075×10^{-4}	3.190×10^{-2} †	2.884×10^{-2} *
n	–	0.900	0.900	1.849×10^{-2}	48.67
S_P	–	0.524	0.538	5.852×10^{-2}	9.20
					t_{ref}=1.664
k_{API}	$(m^3 \times kg^{-1})^{nAPI}\ s^{-1}$	2.300×10^{-12}	2.278×10^{-12}	8.941×10^{-16}	3.923×10^{3}
n_{API}	–	1.00	1.07	1.003×10^{-3}	9.993×10^{2}
					t_{ref}=1.521

Table 16
Global approach focusing on all KIs simultaneously. Final estimation-optimization iteration: comparison between optimization results and current parameter uncertainties. ** = maximum relative uncertainty equal to the upper bound (UB) of the preset range of variation.

Parameter	Units	$\xi_{i,\max}$ [–]	$\varepsilon_{i,\max}$	(95 %CI$_i$)/2
a_1	MPa	1.357×10^{-2}	0.150	8.050×10^{-2}
a_2	–	8.114×10^{-2}	8.747×10^{-2}	4.406×10^{-2}
a_{sf}	–	6.217×10^{-3}	2.841×10^{-3}	1.666×10^{-3}
b_1	–	1.059×10^{-2}	8.581×10^{-2}	5.400×10^{-2}
b_2	–	4.184×10^{-2}	1.364×10^{-2}	5.115×10^{-3}
b_{sf}	MPa^{-1}	1.376×10^{-2}	3.432×10^{-4}	1.886×10^{-4}
γ	dm^{-1}	5.721×10^{-2}	6.408×10^{-5}	5.860×10^{-5}
C_2**	MPa	0.500	44.07	2.271×10^{4}
C_3**	MPa	0.500	70.32	3.459×10^{3}
$\dot{\epsilon}$**	$m\ s^{-1}$	0.500	4.538×10^{-4}	1.595×10^{-2}
n	–	2.101×10^{-2}	1.891×10^{-2}	9.245×10^{-3}
S_P	–	0.118	6.345×10^{-2}	2.926×10^{-3}
k_{API}	$(m^3 \times kg^{-1})^{n_A PI}\ s^{-1}$	1.894×10^{-3}	4.314×10^{-15}	4.471×10^{-16}
n_{API}	–	4.887×10^{-3}	5.229×10^{-3}	5.015×10^{-4}

together with their 95 % CIs and *t*-values. Observations regarding estimated values of model parameters and their precision are similar to the case discussed in Section 5.1.1. The final iteration is presented in Table 16, where all (95 % CI$_i$)/2 are lower than the corresponding maximum uncertainties, $\varepsilon_{i,\max}$.

5.2. Discussion

The case study demonstrated the effectiveness of the proposed approach for systematic evaluation of pharmaceutical process systems models. Here are some additional observations:

- The optimization method presented in this study relies on the original mathematical (systems) model without any linearization of the relationship between parameters and KIs. The procedure is only based on optimization results, and can be implemented easily.
- The number of experimental runs is equal to that obtained by Geremia et al. (2023) for both modular and global approaches; this shows that the linearization introduced in the methodology proposed by Geremia et al. (2023) did not affect the results for the case study being investigated. The modular approach is more efficient than the global one as the total number of experimental runs that are needed to attain the required fidelity on model KIs is less (14 vs. 24) – however, no clear rule could be postulated, and the selection of the best strategy should be done on a case by case basis. Also note that the modular approach relies on the assumption that the parametric precision attained on a previous unit does not need to be improved in order to satisfy the KI fidelity targets that are required in subsequent units; this may not be true and therefore the global approach represents a more general procedure.
- In case a large number of parameters need considering, the optimization problem discussed in this work may lead to numerical issues, long computational times, and results based on local rather than global optimality. The choice of the most effective optimization approach, which is beyond the scope of this work, may be crucial to find a solution.
- The presented methodology relies on the assumption that the only mismatch between model and process depends on the parameter values (parametric mismatch); no structural mismatch has been considered. Moreover, we did not account for measurement noise or bias, which are typically encountered in a real-industrial environment. Accounting for these aspects may impact the capability of reducing parameter uncertainties up to the point that the preset fidelity may not be possible to attain. How structural mismatch and process noise/wrong measurements can be handled effectively is subject of further investigation.
- Uniform parameter distribution is assumed for the given optimization problem. Furthermore, the effect of parameter correlation on uncertainty representation is not accounted for. However, if sufficient information is available more rigorous methods should be considered to characterize the regions of parameter uncertainty (e.g., Schenkendorf et al. 2018) and a different approach to optimization may be needed (for instance, a stochastic optimization formulation could be more effective to solve the problem).

6. Conclusions

Our study supports the use of standardized approaches for model evaluation, and aims at enhancing the systematic use of quantitative models for pharmaceutical process development, optimization, and decision-making. The proposed methodology is based on an optimization framework, and allows the assessment of fidelity in model predictions by directly tackling uncertainty in model parameters. It can be exploited to ensure pre-set requirements on parameters towards model KIs in an explicit way.

Results demonstrate the effectiveness of the method and its consistency when compared to the evaluation framework presented by Geremia et al. (2023). One clear advantage is that no simplification (i.e., linearization of the relationship between model parameters and predicted outputs) is introduced in the model structure.

Future work will aim at testing the procedure experimentally, and at investigating model-process structural mismatch thoroughly. The effect of higher uncertainty in initial values of model parameters should also be further analyzed.

CRediT authorship contribution statement

Margherita Geremia: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Giulio Cisco:** Methodology, Software, Formal analysis. **Samir Diab:** Conceptualization, Methodology, Writing – review & editing. **Gabriele Bano:** Conceptualization, Writing – review & editing. **Fabrizio Bezzo:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

S.D., and G.B. are employees of the GSK group of companies. G.B. reports ownership of GSK shares and/or restricted GSK shares.

Data availability

No data was used for the research described in the article.

Acknowledgment

This study was funded by a Digital Design capability project at GSK.

Appendix A

In this work, we rely on variance-based Sobol's global sensitivity analysis (GSA) (Sobol, 1993) to study how the variance of one model output of interest depends on the input parameters that are affected by uncertainty.

For a generic KI of sub-model M_j, $K_{M_j} = h_{M_j}(\boldsymbol{\theta})$, the finite variance decomposition is computed as:

$$K_{M_j}(\boldsymbol{\theta}) = K_{M_j,\,0} + \sum_{i=1}^{N_\theta} K_{M_j,\,i}(\theta_i) + \sum_{1<i<t<N_\theta} K_{M_j,\,it}(\theta_i, \theta_j) + \ldots + K_{M_j,\,12}(\theta_1,\ \ldots,\ \theta_{N_\theta}), \tag{A.1}$$

where N_θ indicates the number of model parameters, and the partial functions are:

$$K_{M_j,\,0} = E(K_{M_j}), \tag{A.2}$$

$$K_{M_j,\,\boldsymbol{p}}(\boldsymbol{\theta_p}) = E_{\boldsymbol{\theta}_{\sim \boldsymbol{p}}}(K_{M_j}|\boldsymbol{\theta_p}) - \sum_{s\ \subset\ \boldsymbol{p}} K_{M_j,\,s} - K_{M_j,\,0}, \tag{A.3}$$

with $E(\cdot)$ denoting the expectation operation, while $\sim \boldsymbol{p}$ indicating the complementary subset of $\boldsymbol{p}$.

The orthogonal property of the function decomposition (i.e., $\int K_{M_j,\,\boldsymbol{p}}(\boldsymbol{\theta_p}) K_{M_j,\,\boldsymbol{v}}(\boldsymbol{\theta_v}) K_{M_j,\,\boldsymbol{\theta}}(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} = 0$ with $\boldsymbol{p} \neq \boldsymbol{v}$, $\boldsymbol{p}$ and $\boldsymbol{v}$ different subsets) allows to express the variance of the KI of interest $V(K_{M_j})$ as:

$$V(K_{M_j}) = \sum_{i\ =\ 1}^{n} V_i + \sum_{1<i<t<N_\theta} V_{it} + \ \ldots\ + V_{12..N_\theta}, \tag{A.4}$$

with

$$V_{\boldsymbol{p}} = \boldsymbol{V}_{\boldsymbol{\theta_p}}\ \left(E_{\boldsymbol{\theta}_{\sim \boldsymbol{p}}}(K_{M_j}|\boldsymbol{\theta_p})\right)\ - \sum_{s\ \subset\ \boldsymbol{p}} V_s\ . \tag{A.5}$$

V_i, V_{it}, …, $V_{123\ldots\ N_\theta}$ are partial variances which describe the parameters effect on the variance of the response.

According to Homma and Saltelli (1996), results from Sobol's GSA can be collected in two different metrics, i.e., the first-order effect index S_i, and the total effect index $S_{i,\ TOT}$. The first-order effect index, S_i, represents the direct effect contribution of each parameter i to the variance of the output: the higher S_i value, the higher the influence of the i^{th} parameter on the output. The total effect index, $S_{i,\ TOT}$, accounts for the total contribution to the output variance of the i^{th} parameter including both its individual contribution and all higher-order effects due to interactions with other factors.

With respect to the KI of interest, K_{M_j}, they are defined as:

$$S_i = \frac{V_{\theta_i}\left(E_{\boldsymbol{\theta}_{\sim i}}(K_{M_j}|\theta_i)\right)}{V(K_{M_j})}, \tag{A.6}$$

$$S_{i,TOT} = 1 - \frac{V_{\boldsymbol{\theta}_{\sim i}}\left(E_{\theta_i}(K_{M_j}|\boldsymbol{\theta}_{\sim i})\right)}{V(K_{M_j})}\ . \tag{A.7}$$

Sobol's GSA is applied to assess the impact of model parameters on the prediction of model KIs, i.e., TS, disintegration time, and $\%LC$ attained in 25 min. Uniform distributions for the model parameters are assumed. Bounds are chosen assuming initial parameter uncertainties equal to ±50 % of their nominal values that are reported in Table 1. Control variables $\boldsymbol{u}$ (i.e., the tablet press model inputs P and K, which are time-invariant operating conditions) are fixed so that the predicted KIs based on initial parameter values are equal to the targets (i.e., $P = 200$ MPa and $K = 990$ dm). Results are reported in Table A.1 (tablet press unit), Table A.2 (tablet disintegration test unit), and Table A.3 (*in vitro* dissolution test unit).

Ranking of model parameters is coherent with results obtained adopting the procedure proposed by Geremia et al. (2023), where a methodology based on principal component analysis was implemented.

Table A.1
Sobol's sensitivity indices for the parameters of the direct compression systems model with respect to the tensile strength. The most influential model parameters are in boldface.

Parameter	Units	S_i	$S_{i,\ TOT}$
a_1	MPa	3.309×10^{-2}	4.471×10^{-2}
a_2	–	9.319×10^{-3}	1.004×10^{-2}
$\boldsymbol{a_{sf}}$	–	0.589	0.630
$\boldsymbol{b_1}$	–	0.156	0.177
b_2	–	7.126×10^{-2}	9.631×10^{-2}
b_{sf}	MPa^{-1}	8.124×10^{-2}	9.297×10^{-2}
γ	dm^{-1}	8.805×10^{-3}	9.628×10^{-3}

Table A.2
Sobol's sensitivity indices for the parameters of the direct compression systems model with respect to the disintegration time. The most influential model parameters are in boldface.

Parameter	Units	S_i	$S_{i,\ TOT}$
a_1	MPa	4.767×10^{-4}	1.925×10^{-2}
a_2	–	3.762×10^{-4}	5.478×10^{-3}
$\mathbf{a_{sf}}$	–	$\mathbf{8.425 \times 10^{-2}}$	**0.286**
$\mathbf{b_1}$	–	$\mathbf{1.562 \times 10^{-3}}$	$\mathbf{5.431 \times 10^{-2}}$
b_2	–	1.273×10^{-4}	1.634×10^{-2}
$\mathbf{b_{sf}}$	MPa^{-1}	$\mathbf{2.245 \times 10^{-2}}$	**0.139**
γ	dm^{-1}	2.332×10^{-4}	2.695×10^{-4}
C_2	MPa	2.815×10^{-4}	3.992×10^{-3}
C_3	MPa	2.209×10^{-4}	3.731×10^{-3}
$\dot{\epsilon}$	$m\ s^{-1}$	3.603×10^{-4}	5.042×10^{-4}
$\mathbf{n}$	–	**0.641**	**0.881**
$\mathbf{S_p}$	–	$\mathbf{1.312 \times 10^{-2}}$	**0.101**

Table A.3
Sobol's sensitivity indices for the parameters of the direct compression systems model with respect to the %LC attained in 25 min. The most influential model parameters are in boldface.

Parameter	Units	S_i	$S_{i,\ TOT}$
a_1	MPa	7.248×10^{-3}	9.525×10^{-3}
a_2	–	1.492×10^{-3}	1.893×10^{-3}
$\mathbf{a_{sf}}$	–	$\mathbf{3.253 \times 10^{-2}}$	$\mathbf{4.558 \times 10^{-2}}$
$\mathbf{b_1}$	–	$\mathbf{1.127 \times 10^{-2}}$	$\mathbf{1.687 \times 10^{-2}}$
b_2	–	5.504×10^{-3}	5.594×10^{-4}
$\mathbf{b_{sf}}$	MPa^{-1}	$\mathbf{2.042 \times 10^{-2}}$	$\mathbf{3.774 \times 10^{-2}}$
γ	dm^{-1}	4.320×10^{-3}	6.111×10^{-3}
C_2	MPa	4.201×10^{-4}	4.462×10^{-4}
C_3	MPa	2.677×10^{-4}	3.302×10^{-4}
$\dot{\epsilon}$	$m\ s^{-1}$	1.785×10^{-4}	3.797×10^{-4}
$\mathbf{n}$	–	**0.529**	**0.678**
$\mathbf{S_p}$	–	$\mathbf{2.220 \times 10^{-2}}$	$\mathbf{8.327 \times 10^{-2}}$
$\mathbf{k_{API}}$	$(m^3 \times kg^{-1})^{n_{API}}\ s^{-1}$	**0.275**	**0.353**
n_{API}	–	1.165×10^{-2}	1.333×10^{-2}

Appendix B

Here we report the full set of equations of the direct compression (DC) systems model presented in Geremia et al. (2023).

B.1. Model for the tablet press unit operation

The model equations of the sub-model for the tablet press unit operation are listed in the following (Nassar et al., 2021). Variables appearing in the model equations for the tablet press unit operation are reported in Table B.1.

$$sf = \frac{a_{sf}\left(1 + b_{sf}P\right)}{1 + a_{sf}b_{sf}P} \tag{B.1}$$

$$TS = TS_0((1-\beta) + \beta\exp(-\gamma K)), \tag{B.2}$$

$$TS_0 = a_1\exp(b_1(1-sf)), \tag{B.3}$$

$$\beta = a_2(1-sf) + b_2 \tag{B.4}$$

Seven model parameters associated with the tablet press unit need to be estimated: a_{sf} [–], b_{sf} [MPa^{-1}], γ [dm^{-1}], a_1 [MPa], b_1 [–], a_2[–], b_2 [–].

Table B.1
List of process operating variables, model variables, and model parameters to be estimated into the sub-model for the tablet press unit operation.

	Symbol	Units
Process operating variables		
Compaction pressure	P	MPa
Lubrication extent	K	dm
Model variables		
Tensile strength at zero porosity	TS_0	MPa
Total fraction of tensile strength that can be lost due to lubrication	β	–

(*continued on next page*)

Table B.1 (*continued*)

	Symbol	Units
Model parameters to be estimated		
Extended Kushner parameter	a_1	MPa
Extended Kushner parameter	a_2	–
Kawakita model parameter	a_{sf}	–
Extended Kushner parameter	b_1	–
Extended Kushner parameter	b_2	–
Kawakita model parameter	b_{sf}	MPa^{-1}
Lubrication rate constant	γ	dm^{-1}

B.2. Model for the tablet disintegration test unit

The model equations of the sub-model for the tablet disintegration test unit are listed in the following (Geremia et al., 2023). Variables appearing in the model equations for the tablet disintegration test unit are reported in Table B.2.

$$V_c = (H_{coat} - \dot{\epsilon}t)A_t, \tag{B.5}$$

$$\frac{dP_d}{dt} = \left(\frac{P}{F_L/A_t}\right)^{n\left(T_{t/2}-P_d\right)/T_{t/2}} \left[\frac{d_h^2 \varepsilon}{S_p \tau_{or}^2 \mu P_d}\right] p_c, \tag{B.6}$$

$$\tau = -TS + C_2 w_l + C_3\sqrt{w_l}\ . \tag{B.7}$$

$$\tau = \frac{G_0 \exp\left(-\frac{E\varepsilon}{1-\varepsilon}\right)\lambda t}{T_{t/2}} \tag{B.8}$$

Five model parameters associated with the disintegration test unit need to be estimated: C_2 [MPa], C_3 [MPa], $\dot{\epsilon}$ [m s^{-1}], n [–], S_p [–].

Table B.2
List of process operating variables, model variables, and model parameters to be estimated into the sub-model for the tablet disintegration test unit.

	Symbol	Units
Process operating variables		
Compaction pressure	P	MPa
Model variables		
Coating volume	V_c	m^3
Thickness of the coating layer	H_{coat}	m
Time	t	s
Tablet surface area	A_t	m^2
Water penetration depth	P_d	m
Tablet hydraulic diameter	d_h	m
Average tablet tortuosity	τ_{or}	–
Half tablet thickness	$T_{t/2}$	m
Average porosity of the swollen tablet	ε	–
Total stress	τ	MPa
Liquid content in the tablet	w_l	–
Swelling rate	λ	s^{-1}
Elastic constant	G_0	MPa
Elastic constant	E	–
Liquid viscosity	μ	Pa s
Capillary pressure	p_c	Pa
Model parameters to be estimated		
Peppas and Colombo parameter	C_2	MPa
Peppas and Colombo parameter	C_3	MPa
Erosion rate	$\dot{\epsilon}$	m s^{-1}
Swelling parameter	n	–
Pore shape factor	S_P	–

B.3. Model for the in vitro dissolution test unit

The model equations of the sub-model for the *in vitro* dissolution test unit are listed in the following (Bano et al., 2022; Geremia et al., 2023). Variables in the model equations for the *in vitro* dissolution test unit are reported in Table B.3.

$$\frac{\partial N_{API}}{\partial t} = B_{API}\delta\left(l - l_{0,\,API}\right) + R_{API,l}\frac{\partial N_{API}}{\partial l}, \tag{B.9}$$

$$B_{API} = \frac{1}{\rho_p}\left(\frac{x_{API}}{\phi l_{0,API}^3}\right)\frac{dM_t}{dt}, \tag{B.10}$$

$$R_{API,l} = k_{API}\left(c_{sat} - c_{API}\right)^{n_{API}}, \tag{B.11}$$

$$\%LC = 100\,\frac{c_{API}V_m}{x_{API}M_{t,0}}\,. \tag{B.12}$$

Two model parameters need to be estimated for the *in vitro* tablet dissolution test unit: k_{API} [$(m^3 \times kg^{-1})^{n_{API}}\ s^{-1}$], and n_{API} [–].

Table B.3
List of model variables and model parameters to be estimated into the sub-model for the *in vitro* dissolution test unit.

	Symbol	Units
Model variables		
Rate of release of API	B_{API}	s^{-1}
Particle size	l	m
Particle dissolution coefficient	$R_{API,l}$	$m\ s^{-1}$
Dirac delta function	δ	–
Mass fraction of API	x_{API}	–
Tablet mass	M_t	kg
API bulk concentration	c_{API}	$kg\ m^{-3}$
Percentage of label content	$\%LC$	–
Initial mass of the tablet	$M_{t,0}$	kg
Liquid volume in the test vessel	V_m	m^3
Density of particles	ρ_p	$kg\ m^{-3}$
Shape of particles	ϕ	–
Particle size at the beginning	$l_{0,\ API}$	m
API saturation concentration	c_{sat}	$kg\ m^{-3}$
Model parameters to be estimated		
Mass transfer coefficient of API	k_{API}	$(m^3 \times kg^{-1})^{nAPI}\ s^{-1}$
Order of dissolution	n_{API}	–

References

Avraam, M.P., Shah, N., Pantelides, C.C., 1998. Modelling and optimisation of general hybrid systems in the continuous time domain. Comput. Chem. Eng. 22, S221–S228.

Bai, J.P.F., Earp, J.C., Pillai, V.C., 2019. Translational quantitative systems pharmacology in drug development: from current landscape to good practices. AAPS J. 21, 72.

Bano, G., Dhenge, R.M., Diab, S., Goodwin, D.J., Gorringe, L., Ahmed, M., Elkes, R., Zomer, S., 2022. Streamlining the development of an industrial dry granulation process for an immediate release tablet with systems modelling. Chem. Eng. Res. Des. 178, 421–437.

Bard, Y., 1974. Nonlinear Parameter Estimation. Academic Press, New York, NY.

Boukouvala, F., Niotis, V., Ramachandran, R., Muzzio, F.J., Ierapetritou, M.G., 2012. An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process. Comput. Chem. Eng. 42, 30–47.

Boggs, P.T., Tolle, J.W., 1995. Sequential quadratic programming. Acta Numer. 4, 1–51.

Braakman, S., Pathamanathan, P., Moore, H., 2022. Evaluation framework for systems models. CPT Pharmacomet. Syst. Pharmacol. 11, 264–289.

Briskot, T., Stückler, F., Wittkopp, F., Williams, C., Yang, J., Konrad, S., Doninger, K., Griesbach, J., Bennecke, M., Hepbildikler, S., Hubbuch, J., 2019. Prediction uncertainty assessment of chromatography models using Bayesian inference. J. Chromatogr. A 1587, 101–110.

Cenci, F., Pankajakshan, A., Facco, P., Galvanin, F., 2023. An exploratory model-based design of experiments approach to aid parameters identification and reduce model prediction uncertainty. Comput. Chem. Eng. 177, 108353.

Dasgupta, S., Mukhopadhyay, S., Keith, J., 2021. G-optimal grid designs for kriging models. arXiv, 2111.06632.

Daryaee, F., Tonge, P.J., 2019. Pharmacokinetic–pharmacodynamic models that incorporate drug–target binding kinetics. Curr. Opin. Chem. Biol. 50, 120–127.

Dekking, F.M., Kraaikamp, C., Lopuhaä, H.P., Meester, L.E., 2005. A Modern Introduction to Probability and Statistics: Understanding Why and How. Springer, London.

Demetriades, M., Zivanovic, M., Hadjicharalambous, M., Ioannou, E., Ljujic, B., Vucicevic, K., Ivosevic, Z., Dagovic, A., Milivojevic, N., Kokkinos, O., Bauer, R., Vavourakis, V., 2022. Interrogating and quantifying *in vitro* cancer drug pharmacodynamics via agent-based and Bayesian Monte Carlo modelling. Pharmaceutics 14, 749.

Destro, F., Barolo, M., 2022. A review on the modernization of pharmaceutical development and manufacturing – trends, perspectives, and the role of mathematical modeling. Int. J. Pharm. 620, 121715.

Diab, S., Bano, G., Christodoulou, C., Hodnett, N., Benedetti, A., Andersson, M., Zomer, S., 2022a. Application of a system model for continuous manufacturing of an active pharmaceutical ingredient in an industrial environment. J. Pharm. Innov. 17, 1333–1346.

Diab, S., Christodoulou, C., Taylor, G., Rushworth, P., 2022b. Mathematical modeling and optimization to inform impurity control in an industrial active pharmaceutical ingredient manufacturing process. Org. Process Res. Dev. 26, 2864–2881.

Fishman, G.S., 1995. Monte Carlo: Concepts, Algorithms, and Applications. Springer-Verlag, N.Y., USA.

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameters precision: state of the art. Chem. Eng. Sci. 63, 4864–4872.

Geremia, M., Diab, S., Christodoulou, C., Bano, G., Barolo, M., Bezzo, F., 2023. A general procedure for the evaluation of the prediction fidelity of pharmaceutical systems models. Chem. Eng. Sci. 289, 118972.

Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. Reliab. Eng. Syst. Saf. 52, 1–17.

Kucherenko, S., Albrecht, D., Saltelli, A., 2015. Exploring multi-dimensional spaces: a comparison of latin hypercube and Quasi Monte Carlo sampling techniques. *ArXiv150502350 Stat.*

Markl, D., Yassin, S., Wilson, D.I., Goodwin, D.J., Anderson, A., Zeitler, J.A., 2017. Mathematical modelling of liquid transport in swelling pharmaceutical immediate release tablets. Int. J. Pharm. 526, 1–10.

McLean, K.A., McAuley, K.B., 2012. Mathematical modelling of chemical processes–obtaining the best model predictions and parameter estimates using identifiability and estimability procedures. Can. J. Chem. Eng. 90, 351–366.

Monaco, D., Reynolds, G.K., Tajarobi, P., Litster, J.D., Salman, A.D., 2023. Modelling the effect of L/S ratio and granule moisture content on the compaction properties in continuous manufacturing. Int. J. Pharm. 633, 122624.

Moreno-Benito, M., Lee, K.T., Kaydanov, D., Verrier, H.M., Blackwood, D.O., Doshi, P., 2022. Digital twin of a continuous direct compression line for drug product and process design using a hybrid flowsheet modelling approach. Int. J. Pharm. 628, 122336.

Nassar, J., Williams, B., Davies, C., Lief, K., Elkes, R., 2021. Lubrication empirical model to predict tensile strength of directly compressed powder blends. Int. J. Pharm. 592, 119980.

Peppas, N.A., Colombo, P., 1989. Development of disintegration forces during water penetration in porous pharmaceutical systems. J. Control. Release 10, 245–250.

Quaglio, M., Fraga, E.S., Cao, E., Gavriilidis, A., Galvanin, F., 2018. A model-based data mining approach for determining the domain of validity of approximated models. Chemom. Intell. Lab. Syst. 172, 58–67.

Schenkendorf, R., Xie, X., Rehbein, M., Scholl, S., Krewer, U., 2018. The impact of global sensitivities and design measures in model-based optimal experimental design. Processes 6, 27.
Sin, G., Gernaey, K.V., Lantz, A.E., 2009. Good modeling practice for PAT applications: propagation of input uncertainty and sensitivity analysis. Biotechnol. Prog. 25, 1043–1053.
Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comput. Exp. 1, 407–414.
USP <701>Disintegration. The United States pharmacopeial convention. 2011.
White, L.R., Molloy, M., Shaw, R.J., Reynolds, G.K., 2022. System model driven selection of robust tablet manufacturing processes based on drug loading and formulation physical attributes. Eur. J. Pharm. Sci. 172, 106140.
Wang, Z., Escotet-Espinoza, M.S., Ierapetritou, M., 2017. Process analysis and optimization of continuous pharmaceutical manufacturing using flowsheet models. Comput. Chem. Eng. 107, 77–91.
Zineh, I., 2019. Quantitative systems pharmacology: a regulatory perspective on translation. CPT Pharmacomet. Syst. Pharmacol. 8, 336–339.