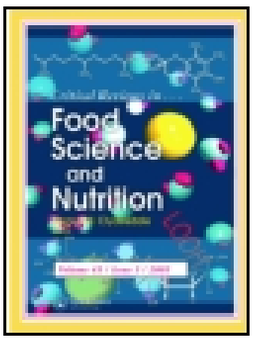

Taylor & Francis
Taylor & Francis Group

Critical Reviews in Food Science and Nutrition

ISSN: 1040-8398 (Print) 1549-7852 (Online) Journal homepage: http://www.tandfonline.com/loi/bfsn20

Towards predictive food process models: A protocol for parameter estimation

Carlos Vilas, Ana Arias-Méndez, Miriam R. Garcia, Antonio A. Alonso & E. Balsa-Canto

To cite this article: Carlos Vilas, Ana Arias-Méndez, Miriam R. Garcia, Antonio A. Alonso & E. Balsa-Canto (2016): Towards predictive food process models: A protocol for parameter estimation, Critical Reviews in Food Science and Nutrition, DOI: 10.1080/10408398.2016.1186591

To link to this article: http://dx.doi.org/10.1080/10408398.2016.1186591

Accepted author version posted online: 31 May 2016.
Published online: 31 May 2016.

Towards predictive food process models: A protocol for parameter estimation

CARLOS VILAS, ANA ARIAS-MÉNDEZ, MÍRIAM R. GARCÍA, ANTONIO A. ALONSO and E. BALSA-CANTO[2]

Bioprocess Engineering Group. IIM-CSIC, Vigo, Spain

Abstract

Mathematical models, in particular, physics-based models, are essential tools to food product and process design, optimization and control.

The success of mathematical models relies on their predictive capabilities. However, describing physical, chemical and biological changes in food processing requires the values of some, typically unknown, parameters. Therefore, parameter estimation from experimental data is critical to achieving desired model predictive properties.

This work takes a new look into the parameter estimation (or identification) problem in food process modeling. First, we examine common pitfalls such as lack of identifiability and multimodality. Second, we present the theoretical background of a parameter identification protocol intended to deal with those challenges. And, to finish, we illustrate the performance of the proposed protocol with an example related to the thermal processing of packaged foods.

Keywords

Model identification; Parameter estimation; Identifiability; Experimental design; Food process engineering

[2]Address correspondence to: E. Balsa-Canto. Bioprocess Engineering Group. IIM-CSIC. c/ Eduardo Cabello,6. 36208 Vigo, Spain. Tel:+34 986 231 930 (Ext. 403). E-mail: ebalsa@iim.csic.es

INTRODUCTION

Consumers demand high quality added value food products. Although price remains a key criterion of most purchasing decisions, pleasure, convenience and health are driving factors of food market evolution. This has motivated researchers and the food industry to explore and to adopt cost-effective technologies that offer better quality and safe products. In this context, there is a growing interest in computer aided food process engineering, i.e. in the use of model-based process design, optimization, and control (e.g. Bruin and Jongen (2003); Banga et al. (2003); Datta (2008); Banga et al. (2008); Perrot et al. (2011); Ho et al. (2013), Galanakis et al., 2015).

Modeling is at the core of food process engineering. The most rigorous models combine two elements. First, conservation laws of energy, mass, and momentum to describe the evolution of thermo-physical variables (temperature, moisture content, etc.). Second, kinetic equations to describe the evolution of quality and safety related attributes (color, vitamins content, microbial content, etc.). Recent works consider the role of biological variability in either the conservation mechanisms or in the evolution of quality attributes. Several works propose the use of stochastic simulation while others propose Monte Carlo sampling methods to assess the uncertainty of model predictions due to variability (Koutsumanis et al., 2002; Hertog et al., 2007; Alonso et al., 2014; García et al., 2015; Ho et al., 2016).

Despite all efforts on developing rigorous models and the necessary numerical simulation techniques, model validation is still a challenge (Datta, 2008; Trystram, 2012). The reasons for that are twofold.

On the one hand, it must be admitted that often experimentation and modeling are fully separate tasks. Usually, experiments are not designed for the purpose of modeling and models are derived without paying especial attention to available experimental data or experimentation capabilities. On the other hand, most models consist of sets of partial and ordinary differential equations (PDEs and ODEs) that depend on, usually unknown, transport and kinetic parameters. When, at some point, the parameter estimation problem is put on the table, modelers use available experimental data to somehow tune the unknown parameters. For example, Maroulis et al. (1995) used least squares plots to compute the heat and mass transfer coefficients in the modeling of air drying of foods; Song et al. (2002) estimated the parameters of a model describing respiration-transpiration in a modified atmosphere packaging using a model linearization based analysis; other works suggest inferring parameters from literature and, in some cases, results are complemented with a parametric sensitivity analysis (Marra et al., 2010; Salvi et al, 2011; Warning et al., 2012; Gulati et al, 2015; to name a few). Unfortunately, these approaches may result in inaccurate or experiment dependent parameter estimates.

Optimization-based approaches can be used for more systematic parameter estimation (Rodriguez et al., 2007; Mariani et al., 2008; Scheerlinck et al., 2008; Dasilva et al., 2009). The idea is to compute the parameter values that minimize the distance between the experimental data and the corresponding model predictions.

However, although the methodology is mathematically well-grounded, in practice, it may suffer from a number of pitfalls. These pitfalls may originate on i) the structure of the model, ii) the available experimental data or iii) the numerical methods used to solve the problem.

In what concerns to model structure and available data, the critical question is whether the parameters are identifiable, i.e. can be given unique values. As such, this issue is often disregarded due to the difficulty to provide an answer *a priori* (Rodriguez et al., 2007). Therefore, in practice, parameters are assumed to have unique solutions and a standard least squares optimizer is used to compute such solutions.

If the model with the optimal parameter values fits the data, the optimal values are considered the "true" parameter values. If, on the contrary, the model does not fit the data, the question is whether the model is not correct or the parameter values are suboptimal. In this concern, the use of global optimization methods was proposed to enable the distinction between a suboptimal fit and a wrong model (Rodriguez et al., 2007).

In many publications related to food process modeling and predictive microbiology, the fitted curves, are reported as "model predictions" and optimal parameters are considered the "true" parameters (Peleg and Corradini, 2011). Results are then given mechanistic interpretations with no consideration for the uncertainty associated neither with the data nor with the parameter estimates.

The uncertainty related to the "true" parameter values will play a major role in the capability to predict a new data set. Therefore, previous works paid special attention to optimal experimental design as a means to obtain accurate parameter estimates and thus improve model validation (Nahor et al., 2001; Balsa-Canto et al., 2007; Scheerlinck et al., 2008; Vanderlinden et al., 2013; Garcia et al., 2015).

Nevertheless, there are still some questions that deserve further attention: identifiability analysis both, *a priori* and *a posteriori*, and model validation.

This work takes a new look at these issues and proposes a data-driven protocol for efficient model parametric identification. As a starting point, we use a number of simple examples to illustrate key pitfalls found in common practice. We continue by presenting the theoretical background of the proposed protocol that includes identifiability analyses, parameter estimation with global optimization methods, optimal experimental design and the analysis of model core predictions. To finish, we show the performance of the protocol by experimentally validating a model that describes the thermal sterilization of solid packaged foods.

COMMON PITFALLS IN PARAMETER ESTIMATION

Illustrative example 1: several parameter values provide exactly the same quality of fit to the data

In some cases different parameter values lead to exactly the same fit. This may become critical if the range of possible solutions for the parameters is high or even infinite.

Let us illustrate the problem with a model that describes the following biochemical reactions:

$$X_1 \xrightarrow{k_1} X_2 \xrightarrow{k_2} X_3 \xrightarrow{k_3} \emptyset$$

where a given reactant X_1 degrades to produce X_2 which in turn degrades to produce X_3 which eventually degrades. Note that the reactions may be reversible but we are considering that reaction rate coefficients in backwards direction are negligible.

This system can be modeled using the law of mass action as follows:

$$\dot{x}_1 = -k_1 x_1$$

$$\dot{x}_2 = k_1 x_1 - k_2 x_2 \qquad (1)$$

$\dot{x}_3 = k_2 x_2 - k_3 x_3$

where $[k_1, k_2, k_3]$ correspond to the unknown kinetic constants and x_i regards the concentration of the reactant X_i. Initial concentration of X_1 is known and set to $x_1(0)$ whereas $x_2(0) = x_3(0) = 0$.

The evolution of the different reactants can be written as follows:

$$x_1(t) = x_1(0)\exp(-k_1 t)$$

$$x_2(t) = x_1(0)\,\frac{k_1}{k_2 - k_1}\left[\exp(-k_1 t) - \exp(-k_2 t)\right] \qquad (2)$$

$$x_3(t) = x_1(0)\,\frac{k_2 k_1}{k_2 - k_1}\left[\frac{(\exp(-k_1 t) - \exp(-k_3 t))}{k_3 - k_1} - \frac{(\exp(-k_2 t) - \exp(-k_3 t))}{k_3 - k_2}\right]$$

provided $k_1 \neq k_2 \neq k_3$.

Let us now assume that we may perform experiments so as to estimate unknown kinetic constants by data fitting. Three different scenarios will be considered: 1) only the concentration of X_1 can be followed through time; 2) the concentrations of X_1 and X_2 can be measured and 3) all three concentrations can be measured. It is also assumed that we may have as many noise free data as desired for the three cases.

1. It is self-evident that continuous noise free data of $x_1 = x_1(0)\exp(-k_1 t)$ will provide information to estimatek_1, but there are infinite possible values for k_2 and k_3 compatible with the experimental data. In other words, if the least squares method is used to compute the values of the kinetic constants, it will not be possible to find a unique solution for k_2 and k_3, no matter the initial concentration of x_1, the duration of the experiment or the number of data.

2. If x_2 can also be measured, from the structure of the model, it can be concluded that, provided we have continuous noise free data, it is possible to find unique values for k_1 and k_2, whereas infinite possibilities remain for k_3.
3. If all concentrations could be measured, all parameters could be given unique values.

In cases 1 and 2, the model is said to be *structurally non-identifiable*, since it is impossible to find unique solutions for all parameters despite the quality and amount of experimental data. Note that, the impossibility of measuring all states in the model may induce structural non-identifiability.

However, full observation is not guarantee of structural identifiability. Let us consider a slab of length L which is heated by its ends with two heat sources at temperatures T_a and T_b, respectively. The distribution of temperature in the slab can be described as follows:

$$T(z,t) = T_a + \frac{T_b - T_a}{L} z + \sum_{n=1}^{\infty} a_n \exp\left(-\frac{k}{\rho\, c_p} \frac{n^2\pi^2}{L^2} t\right) \sin\left(\frac{n\pi}{L} z\right) \quad (3)$$

where $T(z,t)$ represents temperature as a function of space (z) and time (t); a_n are series coefficients depending on the initial and boundary slab temperature (T_0, T_a, T_b); k, ρ and c_p correspond to the material conductivity, density and specific heat coefficient, respectively. In this case, it is easy to realize that, regardless of the possibility of measuring the temperature at any spatial point and at any time, there are infinite combinations of k, ρ and c_p that produce the same result. Only the possibility of performing *ad-hoc* experiments would enable the full characterization of the system. Otherwise one has to restrict to estimate the aggregate $\alpha = k(\rho c_p)^{-1}$.

Most of food processing models are distributed, dynamic and non-linear, thus being impossible to find an analytical solution for the observed quantities. This makes undoubtedly difficult to

anticipate whether a unique value for the unknown parameters can be found attending exclusively to the mathematical structure of the model. However it is important to anticipate such a problem for the sake of model identification.

Illustrative example 2: The model does not fit the data

It is common practice to use standard Newton type, quasi-Newton, Gauss-Newton or Levenberg-Marquardt optimization methods to solve the parameter estimation problem (see, for example, Walter and Pronzato (1997) for a revision of optimization methods). However, in the presence of suboptimal solutions, this type of methods does not guarantee convergence to the best solution (global optimum). Sub-optimal solutions lead to bad fits to the data. In this scenario, it will be rather difficult to distinguish between deficiencies in the model and a suboptimal parameter estimation.

To illustrate this point, let us consider the following classical model of bacterial growth in a batch culture:

$$\dot{c}_b = \frac{\mu_{max}}{K_s + c_s} c_s c_b - K_d c_b$$
$$\dot{c}_s = -\frac{\mu_{max}}{Y(K_s + c_s)} c_s c_b \quad (4)$$

Here, c_b and c_s denote the concentrations of bacteria and nutrients while $\theta = [\mu_{max}, K_s, K_d, Y]$ is the set of unknown parameters representing: the maximum growth rate achieved by c_b; the Michaelis constant; the degradation rate; and the yield coefficient, respectively.

To estimate the parameters we use 10 equidistant measurements obtained from an experiment with initial conditions $c_{b0} = 2$ and $c_{s0} = 30$ and duration 10 hours. We assume that noise in the measurements is Gaussian with a maximum standard deviation of 1 for c_b and 3 for c_s

(approximately a 10 % of the maximum concentration value). The problem is solved with a standard optimizer. Figures 1(a)-1(b) illustrate the complexity of the problem by showing the projection of the least squares over two parameters, μ_{max} and Y. Figures 1(c)-1(d) illustrate the distributions of solutions achieved out of 500 initializations drawn randomly within the parameter bounds [0,10]. The large distributions of solutions in both the least squares of the residuals and the parameters values indicate that results highly depend on the initial guess selected to start the optimization.

The consequences in terms of model predictions are diverse (see Figure 2). Grey curves present model predictions with the different optima obtained. Some solutions correspond to a very fast bacterial decay as compared to the data; others do not predict bacterial decay at all and others result in negative values for the nutrient concentration (those in which the optimizer computed $K_s = 0$), which is physically unreasonable. Only after running a multi-start of the optimizer it is possible to detect that model is correct (at least in what concerns to this data set) and "true" parameter values are $\theta^* = [0.4, 5.0, 0.05, 0.5]$. Bad fits correspond to sub-optimal parameter estimates.

Finding the global optima despite of the initial guess or bounds for the parameters is therefore of critical importance.

Illustrative example 3: A good fit but unsuccessful validation

As mentioned before, often experiments are not designed for the purpose of modeling. Typically, factorial experimental designs are used to test the system responses to different levels of the factors influencing the process (e.g., temperature levels, processing times, etc.). Although quite intuitive

and simple to implement, this technique presents a number of limitations. Especially when using the data to estimate parameters for dynamic models.

Let us recall the bacterial growth model described above -Eqs. (4)-. Eight experiments with different durations and initial concentrations for c_b and c_s were performed according to the design in Table 1. In each experiment we measure both states (c_b, c_s) at final time. Thus, a total amount of 16 data is available to estimate 4 model parameters. We used a global optimization method to compute the parameter values and obtained $\theta^* = [3.47, 7.64, 0.00, 0.55]$. Figure 3 illustrates the difficulty of the optimization problem by representing a projection of the least squares function over μ_{max} and Y. Mean differences between model predictions and experimental data, using this set of parameters, are within the experimental error. However we can already realize that the model will fail to predict a possible decay in c_b since $K_d = 0$.

To illustrate this point, let us consider several validation experiments with $c_{b0} = 1$, $c_{s0} = 60$ and different experimental durations $t_f = [5, 10, 15, 20, 25, 30]$. Validation results are shown in Figure 4. Continuous lines represent the model evolution obtained using the estimated θ whereas marks correspond with the experimental measurements at final time. As expected, the model with the estimated parameters fails to represent the decay of c_b once the nutrients are consumed. It is, therefore, of critical importance to rely on an experimental design intended to parameter estimation.

MODEL IDENTIFICATION PROTOCOL

We propose an identification protocol suited to diagnose for possible sources of difficulties in parameter identification and to iteratively improve model predictive capabilities with the minimum experimental cost. The main steps of the protocol are depicted in Figure 5.

Model formulation

Thermo-physical variables as well as quality and safety attributes (**w** and **z**) usually vary with both time and position within the food product. In addition, they depend on process conditions (**u**) and such dependency is different from product to product and from package to package. As a result, most models consist of sets of partial and ordinary differential equations (PDEs and ODEs) that depend on, usually unknown, transport and kinetic parameters (**p**).

For the sake of mathematical and numerical treatment, it is quite convenient to transform PDEs into ODEs. Typically, classical spatial discretization techniques are used for that purpose. However, their application results in large scale computationally demanding ODE sets. This poses serious difficulties for tasks requiring symbolic manipulation. In addition, the computational cost of those tasks that involve the use of optimizers is extremely large.

In order to overcome this limitation, this work proposes the use of reduced order models by means of global spatial functions defined in the domain of interest.

Reduced order modeling

The Proper Orthogonal Decomposition (POD) approach (Sirovich, 1987) is proposed since its application results in the minimum number of spatial functions required to describe the system behavior. The POD technique is based on the concept of "snapshot". Snapshots are intended to

capture the representative behavior of the system under different operation conditions. The Finite Element Method (FEM) can be used to obtain snapshots and to efficiently obtain the POD reduced order model, see details in Garcia et al. (2007).

Remark that, taking into account the feasible range of parameter values implies a modification in the snapshots acquisition procedure. In this regard, several operating conditions and parameter values are required to generate the snapshots database.

As a result of the POD technique, a set of ordinary differential equations:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p}, t) \tag{5}$$

is obtained on a new set of dynamic variables **x** from which relevant process variables (**z** and **w**) can be recovered.

Structural identifiability analysis

The aim of the structural identifiability analysis is to check whether the unknown parameters could be uniquely determined in the scenario of perfect, noise-free continuous, data. A model is said to be globally structurally identifiable if all parameters can be given unique values in the feasible parameter domain. If more than one, but finite, solutions are possible, then the model is said to be locally structurally identifiable. In the scenarios of structural identifiability, we can proceed with the identification protocol. Otherwise, the model may need some reformulation.

The analysis takes into account model structure, i.e. the set of ordinary differential equations (5) and the observation function which relates the measured quantities with the states in the model, $\mathbf{y} \equiv \mathbf{g}(\mathbf{x}, \mathbf{p})$.

There are only a few methods to test structural identifiability for non-linear dynamic models (Chapman et al., 2003). The most popular are those based on differential algebra (Ljung and Glad, 1994; Bellu et al., 2007) or on power series expansions (Pohjanpalo, 1978; Walter and Lecourtier, 1982). All methods require symbolic manipulation of the model. Therefore, the size of the model, the number of parameters and the number of available observables are critical for the performance of the method. In this respect, power series approaches seem to offer the best compromise between robustness and efficiency (Chis et al., 2011b).

Let's consider the idea that it is possible to find an explicit algebraic dependency of the observables with respect to model states, parameters and controls. If this description, which explicitly depends on the parameters, is unique in the parameter space, then the model is structurally identifiable. Unfortunately, except for very simple examples, it is impossible to find such a description. Power series expansions may be used instead. Similarly, if the power series expansion of the observables with respect to the time and control is unique, then the model is structurally identifiable. Remarkably, power series coefficients can be obtained using symbolic manipulation tools, and their uniqueness can be assessed by solving a system of non-linear equations using the so called identifiability *tableaus* – for details see Balsa-Canto et al. (2010).

Taylor series (Pohjanpalo, 1978) and generating series (Walter and Lecourtier, 1982) are good candidates to handle models related to food processing. However the generating series approach presents a number of advantages. The most important being that the system of equations on the parameters is typically simpler which translates into less computational requirements.

The generating series method can be applied to models which are linear on the control variables, as follows:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = f_0(\mathbf{x}, \mathbf{p}, t) + \sum_{i=1}^{n_c} u_i(t) f_i(\mathbf{x}, \mathbf{p}, t) \quad (6)$$

Note that this is quite often the case in food processing. f_i in Eq. (6) correspond with the i^{th} component of **f**.

The observables are expanded in series with respect to time and the control variables in such a way that the coefficients of these series correspond to $\mathbf{g}(\mathbf{x_0}, \boldsymbol{\theta})$, with $\boldsymbol{\theta} \subset \mathbf{p}$ being the set of unknown model parameters. The number of coefficients required is unknown *a priori*, it certainly depends on the number of observables and the number of parameters, but also on the nature of function **g**.

GenSSI toolbox (Chis et al., 2011a) enables non-expert users to carry out the structural identifiability analysis of non-linear dynamic models**.** The toolbox is implemented using the MATLAB computing environment. It makes use of symbolic manipulation tools to automatize the computation of the power series coefficients and to inform about the global or local structural identifiability of a given model, not requiring high-level programming or advanced mathematical skills from users. In the case of non-identifiable models, GenSSI may detect the subsets of identifiable parameters, therefore guiding in the reformulation of the model when needed.

Parameter estimation with global optimizers

The maximum-likelihood principle yields an appropriate measure of the distance among experimental data and model predictions (Walter and Pronzato, 1997):

$$J_{mc}(\boldsymbol{\theta}) = \sum_{k=1}^{n_{\exp}} \sum_{j=1}^{n_{obs}} \sum_{i=1}^{n_{st}} \left(\frac{y_{k,j,i}(\boldsymbol{\theta}) - y_{k,j,i}^{m}}{\sigma_{k,j,i}} \right) \quad (7)$$

where n_{exp}, n_{obs} and n_{st} are, respectively, the number of experiments, observables, and sampling times while $\sigma_{k,j,i}$ represents the standard deviation of the measurements.

Parameter estimation is then formulated as a nonlinear optimization problem where the decision variables are the unknown parameters ($\boldsymbol{\theta}$). The objective is to minimize $J_{mc}(\boldsymbol{\theta})$ subject to the system dynamics in Eqs. (5) and possible algebraic constraints on the parameters.

As for the case of the illustrative example, in food process models it is quite often the case that this problem is multimodal, i.e. it has several sub-optimal solutions. In this respect, the use of global optimizers has been suggested (Rodriguez et al., 2007). In particular, we propose the use of the *Enhanced Scatter Search* (eSS) (Egea et al., 2009), to find the global solution in reasonable time. eSS combines the scatter search approach, to explore the feasible parameter domain and scape from local optima, and local optimization methods, to enhance convergence to the global solution.

Practical identifiability analysis

Practical identifiability analysis is intended to evaluate the uncertainty of the parameter estimates, in terms of confidence intervals, given a set of experimental data. Confidence intervals may be obtained through the covariance matrix (**C**) as follows:

$$\pm t_{\alpha/2}^{\gamma} \sqrt{\mathbf{C}_{i,i}} \quad (8)$$

where $t_{\alpha/2}^{\gamma}$ is given by Students t-distribution, $\gamma = N_d - n_\theta$ corresponds to the number of degrees of freedom (amount of data - number of unknown parameters) and α is the $(1-\alpha)$ 100% confidence interval selected, typically 95% is used.

Since for non-linear dynamic models there is no exact way to obtain **C**, the Cramér-Rao inequality is typically used (Walter and Pronzato, 1997; Ljung, 1999). Under some assumptions the covariance matrix **C** is bounded by the Fisher information matrix (*F*):

$$\mathbf{C} \geq \boldsymbol{F}^{-1}; \qquad \boldsymbol{F} = E\left\{\left[\frac{\partial J_{mc}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]\left[\frac{\partial J_{mc}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]^T\right\} \tag{9}$$

Note, however, that conditions to use Cramér-Rao bound are not usually satisfied in typical food engineering models. As an alternative, the uncertainty on the parameters can be robustly quantified using a Monte Carlo-based sampling method (Joshi et al., 2006; Balsa-Canto et al., 2010). The underlying idea is to solve the parameter estimation problem for hundreds of experimental replicates. Since in practice it is not realistic to perform such a large number of experiments, we use simulated realizations following the assumed measurements statistics. As a result, a cloud of solutions, which is assumed to be fully contained into a hyper-ellipsoid, are obtained. Correlation and accuracy of parameters are related to the confidence hyper-ellipsoid eccentricity and pseudo-volume, respectively. In this regard, when the parameters are highly correlated the associated hyper-ellipsoids will be highly eccentric. Also, the smaller the volume of the hyper-ellipsoid, the more accurate the parameter values. Confidence intervals can be computed using the distributions of parameter solutions. See Balsa-Canto et al. (2010) for further details.

Model based optimal experimental design

In order to obtain the experimental schemes resulting on the most informative data, a measure of the information quality is required. Data information may be quantified using the Fisher Information Matrix (***F***) defined in Eq (9). Note that *F* is related to the sensitivity of $J_{mc}(\boldsymbol{\theta})$ to variations on the unknown parameters.

The OED problem is then defined as the maximization of a scalar measure of *F* subject to the system dynamics -Eqs. (5)- and to other algebraic constraints associated with experimental limitations. The decision variables of the problem are those related to the experimental scheme, namely, time-varying control profile, sampling times, experiments duration and (possibly) initial conditions.

As a measure of the experimental information we propose the use of the following objective function (Garcia, 2008):

$$J_{OED} = \lambda_{max-1}\lambda_{max-2} \dots \lambda_{min}^2 \quad (10)$$

where λ_{max} and λ_{min} are, respectively, the maximum and minimum eigenvalues of ***F***. Note that the maximization of J_{OED} offers a good compromise between ellipsoid volume and eccentricity.

The OED problem is then solved using the control vector parametrization (CVP) approach (Balsa-Canto et al., 2007) summarized in the following steps:

- Divide the duration of the experiment into a number of intervals
- Approximate the control variables in these intervals using low order polynomials. Usually, constant (zeroth order) or linear (first order) polynomials are considered to represent the controls

- Compute the polynomial coefficients that maximize/minimize J_{OED}

Model validation

The agreement between model and data supports the parameter identification approach. But this agreement does not confirm the validity of the model assumptions. A true validation of the model comes from its ability to predict correctly experimental results not used in parameter estimation.

In this regard, once parameters are computed, it is necessary to test predictive capabilities of the model for a new experimental scenario. We propose to use the concept of core predictions (Garcia et al., 2015). The idea is to analyze how the uncertainty associated to parameter values propagates to model predictions. A Monte Carlo sampling method can be used for this purpose. A sufficiently large number of parameter vectors are selected within the confidence hyper-ellipsoid and model is simulated for each vector under the validation experiment conditions. The core predictions correspond with the whole distribution of solutions.

All the steps of the identification protocol, except for the structural identifiability analysis, are implemented in the AMIGO toolbox (Balsa-Canto and Banga, 2011). The toolbox works under the popular MATLAB environment and interfaces to C or FORTRAN codes to enhance computational performance. It is intended to deal with general non-linear dynamic models. In this regard it offers access to a suite of state-of-the-art simulation and optimization methods. Users can link AMIGO to their own simulators (for example, COMSOL) or optimizers, as well. It allows for simulation, local and global sensitivity analysis and rank of parameters, several practical identifiability tests, parameter estimation using local and global optimization methods and optimal dynamic experimental design.

CASE STUDY: THERMAL STERILIZATION OF PACKAGED FOODS

The proposed protocol is now applied to the identification of a model that describes the thermal sterilization of packaged solid foods. The food container is made of glass with a metallic cover -see Figure 6(a)-. The wall bypass channel located at the center of the cover allows the introduction of thermocouples to measure temperature inside the food product. Thermocouples are then located at $r = 0$ and at any desired height z. The package was filled with a 9% solution of Bentonite, since it has similar thermal properties to several food products (Uno and Hayakawa, 1980).

Mathematical model formulation

We assume that, during the sterilization process, the temperature of the medium surrounding the package, i.e. the retort temperature T_r, is homogeneous along the vessel (Alonso et al., 1997). Therefore, axial symmetry allows us to reduce the original 3D spatial domain to a 2D version presented, together with the spatial discretization scheme, in Figure 6(b).

The dynamics and spatial distribution of the product temperature -$T(r, z, t)$- is described by Fourier's heat equation:

$$\rho c_p \frac{\partial T}{\partial t} = \kappa \left(\frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial z^2} \right) T \tag{11}$$

Heat flux boundary conditions are considered:

$$\kappa \mathbf{n} \cdot \boldsymbol{\nabla} T_g = h_g (T_r - T_g) \tag{12}$$

$$\kappa \mathbf{n} \cdot \boldsymbol{\nabla} T_m = h_g (T_r - T_m) \tag{13}$$

with T_g and T_m being, respectively, the temperature in the glass and metal boundaries. ∇ represents the gradient operator. T_r corresponds with the steam retort temperature which can be measured and controlled. Finally, **n** is a unit vector pointing outward the boundary.

The solution of PDE system (11)-(13) can be numerically found using the finite element method (FEM). In this case, *N*=462 discretization points -as shown in Figure 6(b)- have been chosen since finer grids did not alter the solution significantly.

Remark that the FEM description is used to numerically generate snapshots for the purpose of reduced order modeling (ROM) and to validate ROM predictions.

Reduced order modeling

As mentioned before, we suggest using the POD technique to obtain a reduced order model of the system. The details about the derivation of the POD model are given in Appendix A. Here we show its final form:

$$\rho c_p \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} = \left(\kappa \mathbf{A}_\kappa + h_g \mathbf{A}_g + h_m \mathbf{A}_m\right)\mathbf{m} + \left(h_g \mathbf{B}_g + h_m \mathbf{B}_m\right)T_r \qquad (14)$$

where the vector of states is now $\mathbf{m} = [m_1,\ m_2, \ldots, m_n]^T$. The original state variable $T(r,z,t)$ can be recovered from **m** as indicated in Appendix A. In Eq. (14), $\mathbf{A}_\kappa$, $\mathbf{A}_g$, $\mathbf{A}_m$, $\mathbf{B}_g$ and $\mathbf{B}_m$ are given matrices.

ROM is now compared to FEM model in a number of numerical experiments with time varying processing temperatures and different values for the parameters. Results reveal that the ROM is able to represent the system dynamics with satisfactory accuracy. Some examples are presented in Figure 7. Three different points inside the package (l_1, l_2, l_3) are used in the comparison. Parameter values used in the figures are:

- Figure 7(a): $\rho = 1348\ \text{kg m}^{-3}$, $c_p = 3542\ \text{J kg}^{-1}\ \text{K}^{-1}$, $\kappa = 0.73\ \text{W K}^{-1}\ \text{m}^{-1}$, $h_g = 200\ \text{W K}^{-1}\ \text{m}^{-2}$, $h_m = 500\ \text{W K}^{-1}\ \text{m}^{-2}$
- Figure 7(b): $\rho = 2500\ \text{kg m}^{-3}$, $c_p = 2140\ \text{J kg}^{-1}\ \text{K}^{-1}$, $\kappa = 0.14\ \text{W K}^{-1}\ \text{m}^{-1}$, $h_g = 9\ \text{W K}^{-1}\ \text{m}^{-2}$, $h_m = 20\ \text{W K}^{-1}\ \text{m}^{-2}$

Note that ROM and FEM solutions are equivalent despite the fact that both combinations of parameters result into rather different dynamic behaviors (dynamics in Figure 7(a) are much faster).

The objective now is to estimate, from experimental data, the foodstuff and food package unknown parameters, namely, foodstuff density (ρ), specific heat (c_p) and thermal conductivity (κ) as well as the steam-glass and steam-metal heat transfer coefficients (h_g, h_m).

Structural identifiability analysis

As mentioned above, the parameters considered in the identifiability study are $\theta = [\rho, c_p, \kappa, h_g, h_m]$. The use of the generating series approach is suggested for such study. Note, however, that similarly to what we observed for illustrative example 1:

- Parameters ρ and c_p cannot be given unique values because infinite combinations of ρ and c_p lead to the same value of $p = \rho c_p$. For instance, using $\rho = 1500$, $c_p = 2000$ results into identical, and therefore indistinguishable, solutions as using $\rho = 1000$, $c_p = 3000$. At most it will be only possible to estimate the product $p = \rho c_p$.
- Still, $p = \rho c_p$, κ, h_g and h_m cannot be given unique values. The reason is that the same model solutions are obtained for different values of the parameters. For instance,

multiplying all parameters by an arbitrary quantity $q \in \mathbb{R}$ i.e., $(pq, \kappa q, h_g q, h_m q)$ will result into identical model solutions independently of the value of q.

It must be remarked that these issues are not related to the quantity or quality of the experimental measurements but only depend on the model structure. Hence, let us reformulate the model and divide equation (14) by ρc_p:

$$\rho c_p \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} = \left(p_\kappa A_\kappa + p_g \mathbf{A}_g + p_m \mathbf{A}_m\right)\mathbf{m} + \left(p_g \mathbf{B}_g + p_m \mathbf{B}_m\right)T_r \tag{15}$$

where $p_\kappa = \kappa\left(\rho c_p\right)^{-1}$, $p_g = h_g\left(\rho c_p\right)^{-1}$ and $p_m = h_m\left(\rho c_p\right)^{-1}$. The objective will be now to identify the vector of unknown parameters $\boldsymbol{\theta} = [p_\kappa, p_g, p_m]$ from experimental data.

In this case, simple model inspection is not enough to draw conclusions on the identifiability of parameters p_κ, p_g and p_m. Thus the generating series approach is employed. The GenSSI toolbox was used to analyze structural identifiability. The following main conclusions were drawn:

- The use of one thermocouple (one observable), for instance the measurement at the center of the package, results into two solutions for the parameter values. Parameters are, therefore, locally structurally identifiable.
- The use of two thermocouples (two observables) results into a unique parameter solution. Thus, parameters are structurally globally identifiable.

This result leads us to design experiments with two thermocouples.

Parameter estimation and identifiability analysis°

Data for parameter estimation were acquired in a pilot plant setup that enables monitoring and controlling the thermal process. The estimation experiment was performed with following conditions:

- The product initial temperature is $T(r, z, 0) = 124$ ℃
- Processing temperature was set to $T_r(t) = 119$ ℃.
- Temperature evolution was measured at two different locations (n_{obs}), $z_1 = 0.032$ and $z_2 = 0.052$. Such locations are, respectively, the center of the product and a point close to the product top.
- Retort temperature (boundary condition of the model) was also measured.
- The duration of the experiment was 70 .
- Three replicates were performed so as to estimate experimental error.

The AMIGO toolbox, with the hybrid optimization algorithm eSS, was used to compute the unknown parameter values that minimize J_{mc}. Figure 8 shows best fit, i.e. a comparison between model predictions and experimental data for the optimal parameter values.

The optimal solution was $p_\kappa = 4.12 \times 10^{-7}$ m^2 s^{-1} , $p_m = 1.5 \times 10^{-3}$ m s^{-1} and $p_g = 0.82 \times 10^{-5}$ m s^{-1}. Note that p_m is two-to-three orders of magnitude larger than p_g. This was expected as heat transfer is faster on metallic surfaces as compared to glass surfaces. Such a large value for p_m is equivalent to having a Dirichlet boundary condition in Eq (13) and increasing this

value does not affect the quality of the fit. Therefore p_m was fixed and, from now on, we focus on parameters p_κ and p_g.

Confidence intervals for these parameters were robustly computed using the Monte Carlo sampling method. The resulting mean values and confidence intervals were $p_\kappa = (4.13 \pm 0.73) \times 10^{-7}$ m^2 s^{-1} and $p_g = (0.83 \pm 0.39) \times 10^{-5}$ m s^{-1}. Parameter uncertainty is therefore rather large, 17.6% and 46.7% of the values of p_κ and p_g, respectively.

In order to illustrate the model predictive capabilities and how the uncertainty on parameters propagates to model predictions we performed a second experiment, this time for model validation.

Model validation

Experimental conditions for the validation are different from those used in the estimation experiment:

- The product is initially at room temperature (around 25 °C).
- A dynamic retort temperature profile was implemented. First the retort temperature is stabilized at 105 °C and kept constant during $t = 50$ min, subsequently, step-up to 111 °C was introduced.

Figure 9 presents model core predictions for this validation experiment. Model predictions for the mean parameter values (continuous lines) are rather far from the experimental behavior (marks). In fact the absolute mean error is 3.9 °C for the center of the product and 2.9 °C for the point close to the product top. Notably, at high temperatures, the mean errors between model and experimental data are 4.8 °C for the center of the product and 1.8 °C for the point close to the product top. These

differences will compromise reliability on the predictions of quality and safety related attributes, thus impacting process design results.

In order to improve predictive capabilities and to reduce the parameter confidence intervals, with a minimum experimental effort, a new experiment was optimally designed.

Model based optimal experimental design

The objective was to compute the food initial temperature and the retort temperature profile, that maximizes J_{OED} -see Eq (10)-. To parametrize the retort temperature profile -$T_r(t)$- we considered step-wise and piece-wise linear approximations. A maximum of three different stages was considered. It must be noted that fast changes on T_r result into steep pressure changes which may damage the package, particularly when temperature drops. To avoid package damage we imposed the following constraints in the formulation of the experimental design problem:

- Step-wise processing profiles: $-5℃ < T_{r,i} - T_{r,i+1} < 6℃$.
- Piece-wise linear processing profiles: $\frac{T_{r,i+1}-T_{r,i}}{t_{i+1}-t_i} \in [-1,1]$ ℃ min^{-1}

In addition the experiment duration was fixed to $t_f = 30$ min.

The best objective function was obtained using piece-wise linear processing temperature profile. Such profile is depicted in Figure 10(a) with a black dashed line. Optimal initial product temperature corresponds to $T(r,z,0) = 105$ ℃. At the beginning of the experiment, the retort temperature increases at the maximum rate allowed, i.e. 1 ℃. From minute 6.2, changes in the temperature are smoother.

This experiment was implemented in the pilot plant and the recorded data were added to the previous experimental data.

The Monte Carlo sampling method was then used to robustly compute the confidence intervals associated to the parameters. In this case we obtained $p_\kappa = (4.57 \pm 0.28) \times 10^{-7}$ m^2 s^{-1} and $p_g = (1.40 \pm 0.17) \times 10^{-5}$ m s^{-1}. Confidence intervals are now a 6,3% and 12,2%, respectively.

Figures 10 summarize results of the OED and the Monte Carlo sampling method. Figure 10(a) presents the best fit for the optimally designed experiment showing a good agreement between model predictions and data. Figure 10(b) compares the confidence ellipses obtained by the Monte Carlo sampling method before (dashed line) and after (continuous line) experimental design. It can be observed that parameters are rather correlated. However, the information provided by the optimally designed experiment contributed to reduce the area of the ellipse in such a way that correlation is no longer an issue. Figures 10(c) and 10(d) present the parameter distributions obtained in the Monte Carlo method. Figures highlight two important results obtained by the use of optimal experimental design:

- confidence on parameters has improved (in around 3 times)
- the mean parameter values ($p_\kappa = 4.57 \times 10^{-7}$ m^2 s^{-1} and $p_g = 1.40 \times 10^{-5}$ m s^{-1} have changed, particularly in the case of p_g.

We now recall the validation experiment to check the model predictive capabilities with these new parameter values and their associated uncertainties.

Model validation after optimal experimental design

Figure 11 presents model core predictions against experimental data for the validation experiment. Table 2 presents the evolution of model predictive capabilities, in terms of prediction error, throughout the identification protocol. The application of the proposed parametric identification

protocol converged to a set of parameter values and confidence intervals that lead to reasonably accurate predictions. Maximum prediction errors are found, as expected, at the venting stage when product temperature is below 100 °C. Note that, at venting stage, there is air in the retort and the assumption of homogeneous temperature may not hold. However, errors in the heating phase, in which quality and safety attributes are affected, are lower than the experimental error. Thus rendering a model which is now suitable for process design.

CONCLUSIONS

In this work we presented and tested a data-based protocol for model identification in food engineering. The protocol is intended to diagnose for possible sources of difficulties in parameter identification and to iteratively improve model predictive capabilities with the minimum experimental cost.

First we discussed, using some illustrative examples, the typical difficulties found in the parametric identification of nonlinear models, namely lack of structural or practical identifiability and multimodality.

Then we presented the theoretical background of the parametric identification procedure. Such procedure incorporates the following steps: model reduction; structural identifiability analysis; parameter estimation using global optimizers; practical identifiability analysis; optimal experimental design and model validation.

Finally, we illustrated the performance of the protocol with the experimental validation of a model describing the thermal processing of packaged foods. Results revealed that the use of the proposed identification protocol enabled the iterative improvement of model predictive capabilities with minimum experimental effort.

The protocol may be applied to general distributed and dynamic models in food engineering. It has been designed to aid modelers in the diagnose for possible sources of difficulties in parameter identification and to iteratively improve model predictive capabilities with the minimum experimental cost.

Acknowledgements

The authors acknowledge financial support from the EU (Project CAFE FP7-KBBE-2007-1(212754)), Spanish Ministry of Science and Innovation (Project ISFORQUALITY AGL2012-39951-C02-01) and CSIC (Project CONTROLA PIE 201270E075). Authors acknowledge the collaboration of J. I. Molina in the experimental validation.

References

Alonso, A.A., Banga, J.R., and Pérez-Martín, R.I. (1997). A complete dynamic model for the thermal processing of bioproducts in batch units and its application to controller design. *Chemical Engineering Science*, **52**: 1307–1322.

Alonso, A.A., Molina, I. and Theodoropoulos, C. (2014). Modeling bacterial population growth from stochastic single-cell dynamics. *Applied and Environmental Microbiology*, **80**(17), pp. 5241-5253.

Balsa-Canto, E., Alonso, A.A., and Banga, J.R. (2010). An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Systems Biology*, **4**: 1–18.

Balsa-Canto, E. and Banga, J.R. (2011). AMIGO, a toolbox for advanced model identification in systems biology using global optimization. *Bioinformatics*, **27**: 2311–2313.

Balsa-Canto, E., Rodríguez-Fernández, M., and Banga, J.R. (2007). Optimal design of dynamic experiments for improved estimation of kinetic parameters of thermal degradation. *Journal of Food Engineering*, **82**: 178–188.

Banga, J.R., Balsa-Canto, E., and Alonso, A.A. (2008). Quality and safety models and optimization as part of computer-integrated manufacturing. *Comprehensive Reviews in Food Science and Food Safety*, **7**: 168–174.

Banga, J.R., Balsa-Canto, E., Moles, C., and Alonso, A.A. (2003). Improving food processing using modern optimization methods. *Trends in Food Science and Technology*, **14**: 131–144.

Bellu, G., Saccomani, M.P., Audoly, S., and D'Angio, L. (2007). DAISY: A new software tool to test global identifiability of biological and physiological systems. *Computer Methods and Programs in Biomedicine*, **88**: 52–61.

Bruin, S. and Jongen, T.R.G. (2003). Food process engineering: The last 25 years and challenges ahead. *Comprehensive Reviews in Food Science and Food Safety*, **2**: 42–81.

Chapman, M.J., Godfrey, K.R., Chappell, M.J., and Evans, N.D. (2003). Structural identifiability for a class of non-linear compartmental systems using linear/non-linear splitting and symbolic computation. *Mathematical Biosciences*, **183**: 215–215.

Chis, O., Banga, J.R., and Balsa-Canto, E. (2011a). GenSSI: a software toolbox for structural identifiability analysis of biological models. *Bioinformatics*, **27**: 2610–2611.

Chis, O., Banga, J.R., and Balsa-Canto, E. (2011b). Structural identifiability of systems biology models: A critical comparison of methods. *Plos One*, **6**: e27755.

da Silva, C., da Silva, Z., and Mariani, V. (2009). Determination of the diffusion coefficient of dry mushrooms using the inverse method. *Journal of Food Engineering*, **95**: 1–10.

Datta, A.K. (2008). Status of physics-based models in the design of food products, processes, and equipment. *Comprehensive Reviews in Food Science and Food Safety*, **7**: 121–129.

Egea, J.A., Vazquez, E., Banga, J.R., and Marti, R. (2009). Improved scatter search for the global optimization of computationally expensive dynamic models. *Journal of Global Optimization*, **43**: 175–190.

Galanakis, C.M., Patsioura, A. and Gekas, V. (2015) Enzyme Kinetics Modeling as a Tool to Optimize Food Industry: A Pragmatic Approach Based on Amylolytic Enzymes. *Critical Reviews in Food Science and Nutrition*, **55(12)**:1758-1770.

García, M.R., Vilas, C., Herrera, J.R., Bernárdez, M., Balsa-Canto, E., and Alonso, A.A. (2015). Quality and shelf-life prediction for retail fresh hake (merluccius merluccius). *International Journal of Food Microbiology*, **208** :65–74.

García, M.R. (2008). *Identification and Real Time Optimisation in the Food Processing and Biotechnology Industries.* PhD thesis, University of Vigo, Spain.

García, M.R., Vilas, C., Banga, J.R., and Alonso, A.A. (2007). Optimal field reconstruction of distributed process systems from partial measurements. *Industrial & Engineering Chemistry Research*, **46**: 530–539.

Gulati, T., Datta, A.K., Doona, C.J., Ruan, R.R., Feeherry, F.E. (2015). Modeling moisture migration in a multi-domain food system: Application to storage of a sandwich system, *Food Research International*, **76**(3): 427-438.

Ho, Q., Carmeliet, J., Datta, A., Defraeye, T., Delele, M., Herremans, E., Opara, L., Ramon, H., Tijskens, E., van der Sman, R., Liedekerke, P.V., Verboven, P., and Nicolai, B. (2013). Multiscale modeling in food engineering. *Journal of Food Engineering*, **114**: 279 – 291.

Ho, Q.T., Rogge, S., Verboven, P., Verlinden, B.E., Nicolaï, B.M. Stochastic modelling for virtual engineering of controlled atmosphere storage of fruit, *Journal of Food Engineering*, **176**: 77-87, 2016.

Joshi, M., Seidel-Morgenstern, A., and Kremling, A. (2006). Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. *Metabolic Engineering*, **8**: 447–455.

Koutsoumanis, K., Giannakourou, M., Taoukis, P., and Nychas, G. (2002). Application of shelf life decision system (slds) to marine cultured fish quality. *International Journal of Food Microbiology*, **73**: 375-382.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall PTR.

Ljung, L. and Glad, T. (1994). On global identifiability for arbitrary model parametrizations. *Automatica*, **30**: 265–176.

Mariani, V., de Lima A.G.B., and Coelho, L. (2008). Apparent thermal diffusivity estimation of the banana during drying using inverse method. *Journal of Food Engineering*, **85**: 569–579.

Maroulis, Z.B., Kiroanoudis, C.T., Marinos-Kouris, D. (1995) Heat and mass transfer modeling in air drying of foods. *Journal of Food Engineering*, **26**:113-130.

Marra, F., De Bonis, M.V., Ruocco, G. (2010) Combined microwaves and convection heating: A conjugate approach. *Journal of Food Engineering*, **97**:31-39.

Nahor, H.B., Scheerlinck, N., Verniest, R., De Baerdemaeker, J., and Nicolai, B.M. (2001). Optimal experimental design for the parameter estimation of conduction heated foods. *Journal of Food Engineering*, **48**: 109–119.

Peleg, M. and Corradini, M.G. (2011) Microbial Growth Curves: What the Models Tell Us and What They Cannot, *Critical Reviews in Food Science and Nutrition*, **51(10):** 917-945.

Perrot, N., Trelea, I.C., Baudrit, C., Trystram, G., and Bourgine, P. (2011). Modelling and analysis of complex food systems: State of the art and new trends. *Trends in Food Science and Technology*, **22**: 304–314.

Pohjanpalo, H. (1978). System identifiability based on power-series expansion of the solution. *Mathematical Biosciences*, **41**: 21–33.

Rodríguez-Fernández, M., Balsa-Canto, E., Egea, J.A., and Banga, J.R. (2007). Identifiability and robust parameter estimation in food process modeling: Application to a drying model. *Journal of Food Engineering*, **83**: 374–383.

Salvi, D., Boldor, D., Aita, G.M., Sabliov, C.M. (2011) COMSOL Multiphysics model for continuous flow microwave heating of liquids. *Journal of Food Engineering*, **104**:422-429.

Scheerlinck, N., Berhane, N., Moles, C., Banga, J., and Nicolaï, B. (2008). Optimal dynamic heat generation profiles for simultaneous estimation of thermal food properties using a hotwire probe: computation, implementation and validation. *J. Food Eng.*, **84**: 297–306.

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures. Part I: Coherent structures. *Quaterly of Applied Mathematics*, **45**: 561–571.

Song, Y., Vorsa, N., Yam, K.L. (2002) Modeling respiration-transpiration in a modified atmosphere packaging system containing blueberry. *Journal of Food Engineering*, **53**:103-109.

Trystram, G. (2012). Modelling of food and food processes. *Journal of Food Engineering*, **110**:269–277.

Uno, J.I. and Hayakawa, K.I. (1980). A method for estimating thermal diffusivity of heat conduction food in a cylindrical can. *Journal of Food Science*, **45**: 692–695.

Van Derlinden, E., Mertens, L., and Van Impe, J. (2013). The impact of experiment design on the parameter estimation of cardinal parameter models in predictive microbiology. *Food Control*, **29**: 300–308.

Walter, E. and Lecourtier, Y. (1982). Global approaches to identifiability testing for linear and non-linear state-space models. *Mathematics and Computers in Simulation*, **24**: 472–482.

Walter, E. and Pronzato, L. (1997). *Identification of parametric models from experimental data*. Springer.

Warning, A., Dhall, A., Mitrea, D., Datta, A.K. (2012) Porous media based model for deep-fat vacuum frying potato chips. *Journal of Food Engineering*, **110** (3), 428–440.

APPENDIX A: DERIVATION OF REDUCED ORDER MODELS USING THE PROPER ORTHOGONAL DECOMPOSITION

Let us first rewrite, for the sake of clarity, model equations (11)-(13) describing the product temperature of the food package being sterilized:

$$\rho c_p \frac{\partial T}{\partial t} = \kappa \Delta T; \qquad \Delta = \left(\frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial z^2} \right) \tag{16}$$

Boundary conditions are of the form:

$$\kappa \mathbf{n} \cdot \boldsymbol{\nabla} T_g = h_g \left(T_r - T_g \right) \tag{17}$$

$$\kappa \mathbf{n} \cdot \boldsymbol{\nabla} T_m = h_g (T_r - T_m) \tag{18}$$

There exist a wide range of possibilities to solve systems of PDEs of the form of (16)-(18). Probably the most well-known belong to the family of weighted residuals in which the state variable, i.e. the temperature inside the food product, is written as a linear combination of time dependent coefficients (m_i) and spatial dependent functions -basis functions -$\phi_i(r,z)$-:

$$T(r,z,t) = \sum_{i=1}^{\infty} \phi_i(r,z)\, m_i(t) \approx \sum_{i=1}^{n} \phi_i(r,z)\, m_i(t) \tag{19}$$

where n is the number of elements required to obtain a sufficiently accurate approximation. Expression (19) is substituted into the PDE equation which, at the same time is projected onto given weighting functions -$w_i(r,z)$-. The projection step will be explained below.

Depending on the selection of the basis and weighting functions, different methods arise. In the Galerkin method the basis functions are the same as the weighting functions, i.e. $\phi_i(r,z) = w_i(r,z)$. The POD belongs to the family of Galerkin methods and the basis/weighting functions

are computed from experimental (*in silico* or *in vivo*) data. The main idea is to take measurements of the state variable at different spatial points and at given times (snapshots). Then the snapshots are used to construct the following two point correlation kernel:

$$\mathcal{K} = \frac{1}{p}\sum_{i=1}^{p} T_i\, T_i'$$

where $T_i \in \mathbb{R}^N$ is the vector of temperature values at different *N* spatial points and at time t_i. The prime symbol indicates the transpose.

It is important to mention here that the selection of the set of snapshots is a key step in the derivation of the reduced order model as the quality of the approximation will be highly affected by it. In this regard, the snapshots must be representative of the behavior of the system. i.e., the snapshots should be taken within the range of the plant operational conditions. Also, when system dynamics are fast, the time interval between two snapshots should be short. Finally, it should be also mentioned that, when the model is used for parameter calibration and the snapshots are obtained from *in silico* experiments then the range of possible values for the parameters should be used in such experiments.

For instance, in the case considered in this work, five parameters $(\rho, c_p, \kappa, h_m, h_g)$ are unknown so the following sets of values were chosen for the parameters: $p_\kappa = \kappa(\rho c_p)^{-1} = [1,\ 0.2,\ 0.1,\ 0.04] \times 10^{-6}\ \mathrm{m^2\ s^{-1}}$, $h_m = [10, 10^2, 10^3, 10^4]\ \mathrm{W\ m^2\ K^{-1}}$ and $h_g = [10,\ 10^2, 10^3, 10^4]\ \mathrm{W\ m^2\ K^{-1}}$. All the possible parameter combinations were used to obtain the snapshots.

Once the kernel is obtained, the basis functions are computed by solving the following eigenvalue problem:

$$\lambda_i \phi_i = \int_V \mathcal{K} \phi_i \mathrm{d}V$$

where V represents the volume of the package.

The next step is to obtain the time dependent functions of the series expansion (19), i.e. the coefficients $m_i(t)$. To that purpose, the PDE is projected over the basis functions. In practice, projection is performed by multiplying the PDE by the basis functions and integrating over the spatial domain:

$$\rho c_p \int_V \phi_j \frac{\partial \mathrm{T}}{\partial \mathrm{t}} \mathrm{d}V = \kappa \int_V \phi_j \Delta \mathrm{T} \mathrm{d}V\,; \qquad j = 1, \dots, n$$

Substituting T by the series expansion (19) and using the boundary conditions (17)-(18), the following set of ODEs is obtained:

$$\rho c_p \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} = \left(\kappa \mathbf{A}_\kappa + h_g \mathbf{A}_g + h_m \mathbf{A}_m\right)\mathbf{m} + \left(h_g \mathbf{B}_g + h_m \mathbf{B}_m\right) T_r \qquad (20)$$

where $\mathbf{m} = [m_1,\ m_2, \dots, m_n]^T$ is the vector of time dependent functions in the series (19) and each element (i,j) of $\mathbf{A}_\kappa$, $\mathbf{A}_g$, $\mathbf{A}_m$, $\mathbf{B}_g$ and $\mathbf{B}_m$ are given matrices.

$$\mathbf{A}_\kappa(i,j) = \int_V \nabla \phi_i \nabla \phi_j \mathrm{d}V\,; \quad \mathbf{A}_\ell(i,j) = \int_{\mathfrak{B}_\ell} \phi_i \phi_j \mathrm{d}\mathfrak{B}_\ell\,; \quad \mathbf{B}_\ell(i) = \int_{\mathfrak{B}_\ell} \phi_i \mathrm{T_r} \mathrm{d}\mathfrak{B}_\ell\,; \quad \ell = g, m$$

with $\mathfrak{B}_g$ and $\mathfrak{B}_m$ being, respectively, the glass and metal boundaries. In the derivation of Eq. (20) we took advantage of the orthonormality of the basis functions, i.e.

$$\int_V \phi_i \phi_j \mathrm{d}V = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

The computation of the spatial integrals can be performed using the finite element structure as shown in Garcia et al. (2007).

List of Tables

Table 1. Factorial experiment design to estimate bacterial growth model parameters.

	Experiment							
Parameter	E1	E2	E3	E4	E5	E6	E7	E8
c_{s0}	45	45	45	45	15	15	15	15
c_{b0}	3	3	1	1	3	3	1	1
t_f	3	6	3	6	3	6	3	6

Table 2. Mean errors (in ℃) between model predictions and experimental data in the validation experiment at both measurement points. "Heating" refers to that part of the process where product temperatures are close to 100 ℃ or above.

	Center		Top	
	Whole process	Heating	Whole process	Heating
Before OED	3.9	4.8	2.9	1.8
After OED	2.2	0.6	1.0	0.3

List of Figures

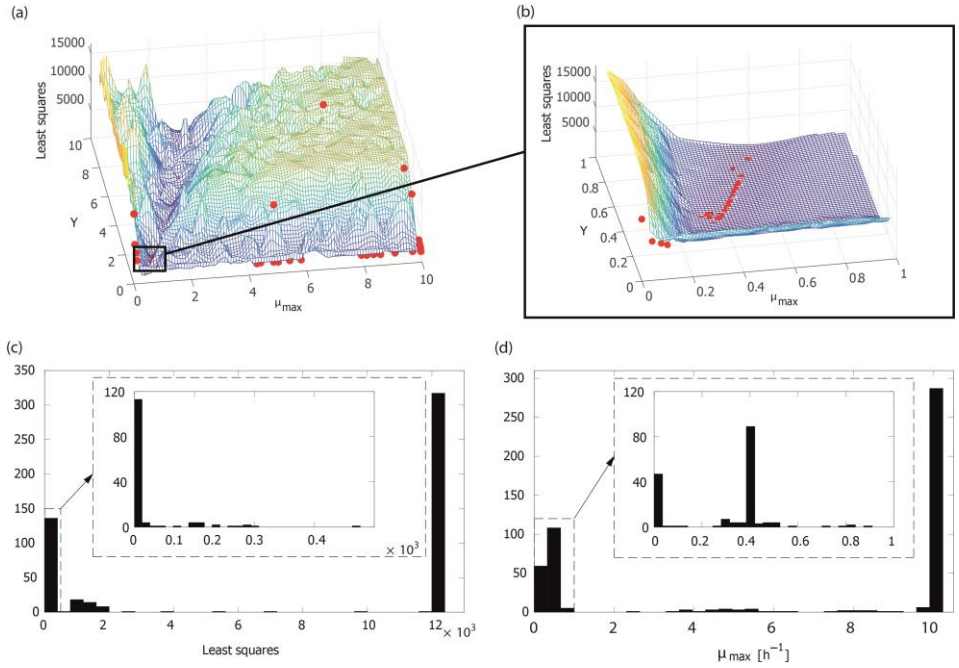

Figure 1. Illustrative representation of the optimization problem and its solution with a multi-start of a standard optimizer. (a) Projection of the objective function over μ_{max} and Y in the search space together with the multi-start solutions; (b) zoom-in of the objective function over μ_{max} and Y in the vicinity of the global solution; (c) distribution of objective function values; (d) Distribution of μ_{max} values.

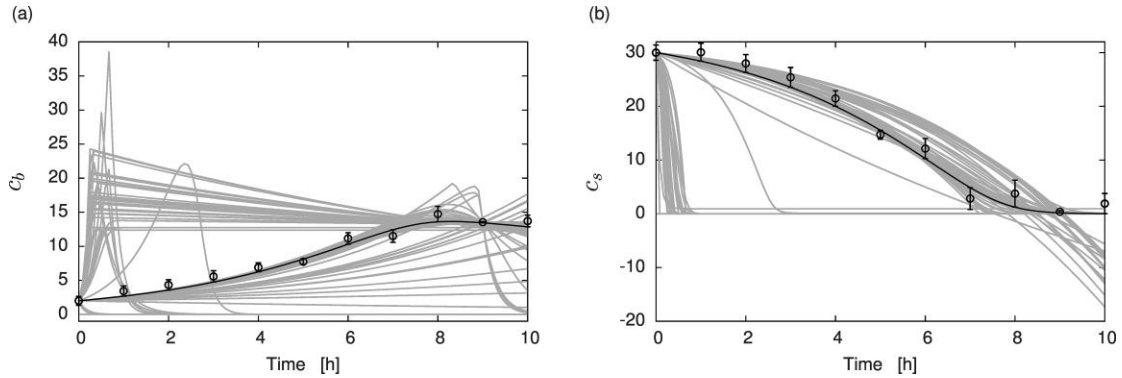

Figure 2. Model predictions out of a multi-start of a standard optimizer. Dots and error bars represent experimental data and the experimental noise. Black continuous line represents model prediction with the best parameter values. Grey lines represent model predictions with local optima.

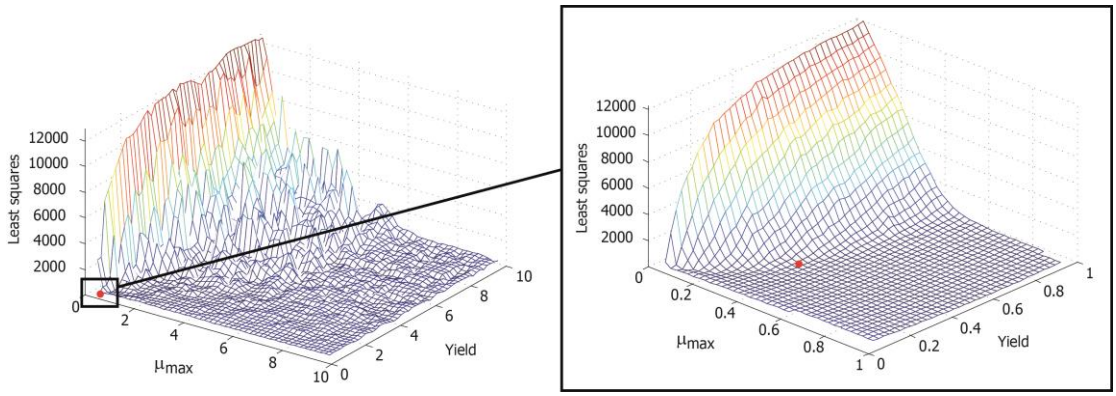

Figure 3. Illustrative representation of the optimization problem for the case of performing a factorial experimental design. (a) Projection of the objective function over μ_{max} and Y in the search space; (b) zoom-in of the objective function over μ_{max} and Y in the vicinity of the global solution

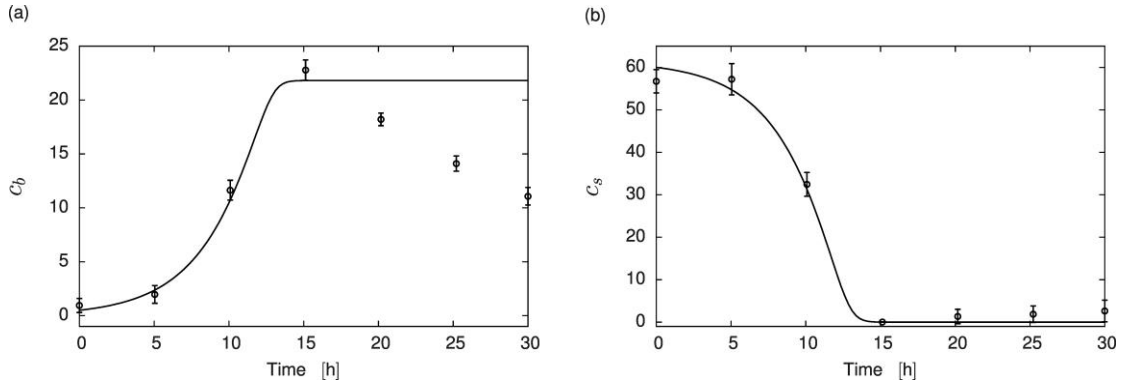

Figure 4. Experimental data vs. solutions of the bacterial growth model with optimal parameter values found using the factorial experimental design in Table 1. Figure (a) shows that the model is not able to predict bacterial decay once the nutrients are consumed.

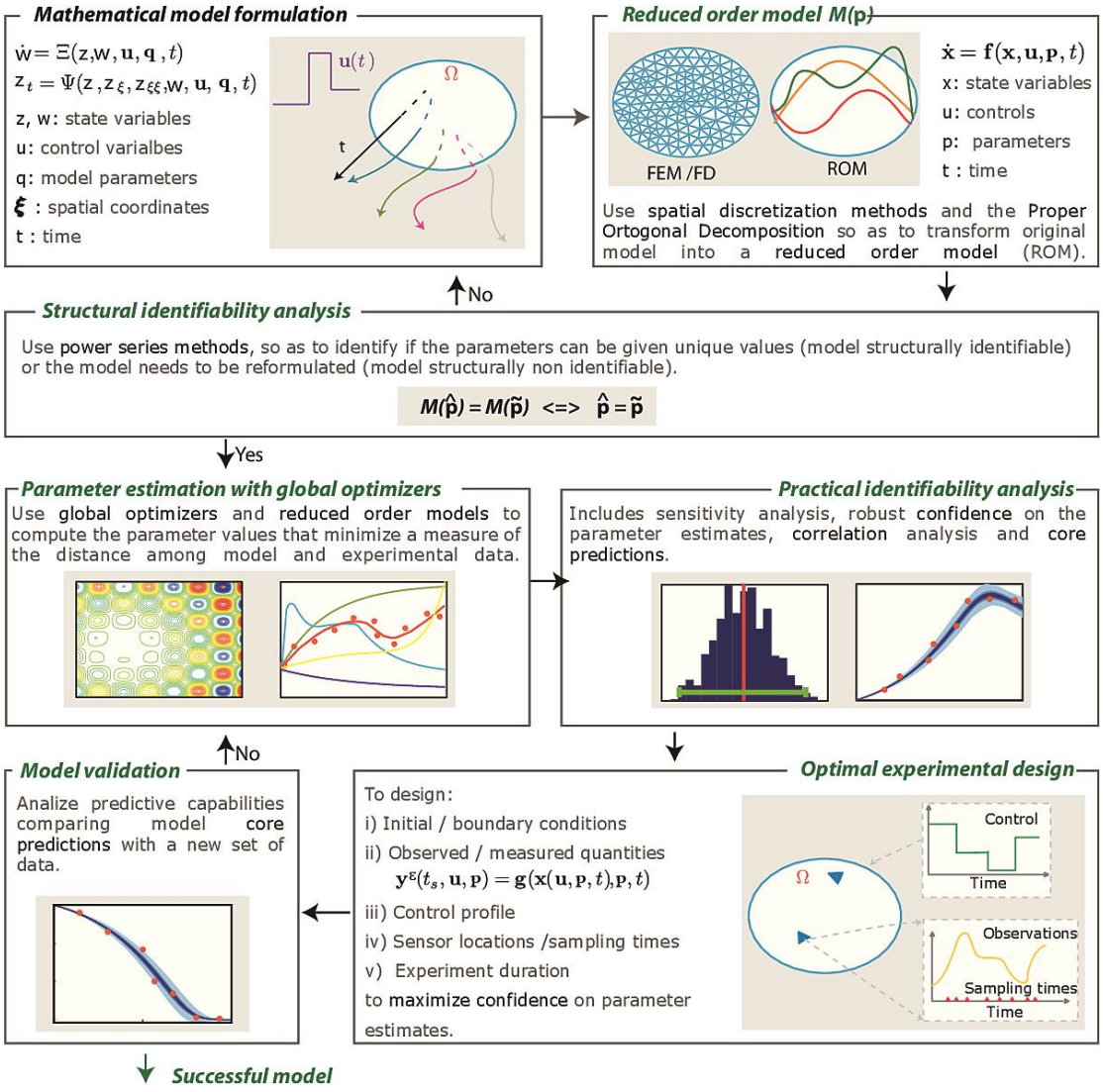

Figure 5. General scheme of the proposed identification protocol.

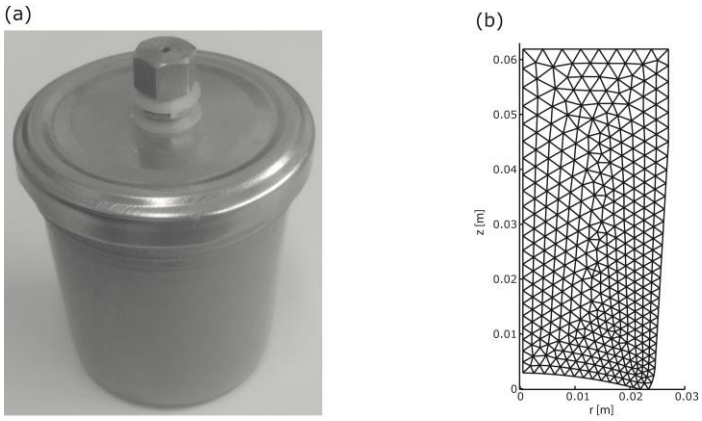

Figure 6. (a) Picture of the glass package with a metallic cover used in the experiments. The wall bypass channel, located at the center of the cover, allows to introduce a thermocouple in the foodstuff. (b) Half section of the package geometry and finite element mesh used to numerically solve the model equations.

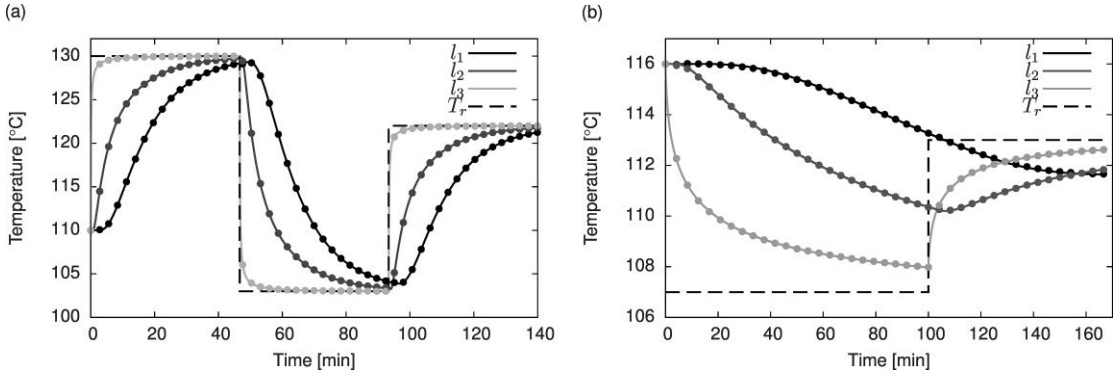

Figure 7. Temperature as predicted by the FEM-based model (continuous lines) and the POD based reduced model (marks). Figures (a) and (b) present four curves: the retort temperature profile, plus the evolution of temperature at three locations inside the food product l1 = (0, 0.031) m; l2 = (0.0177, 0.0517) m and l3 = (0.0265, 0.062) m.

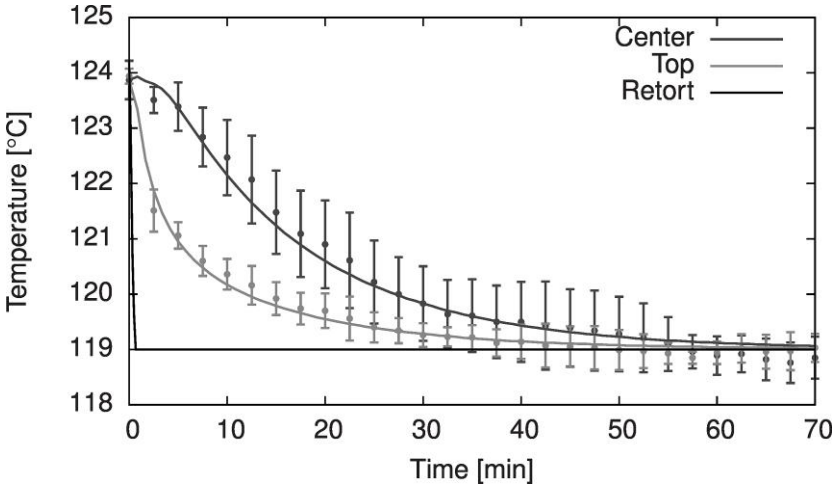

Figure 8. Best fit: model predictions (continuous lines) versus experimental data (marks) for the two measurement points. Black line represents the measured retort temperature.

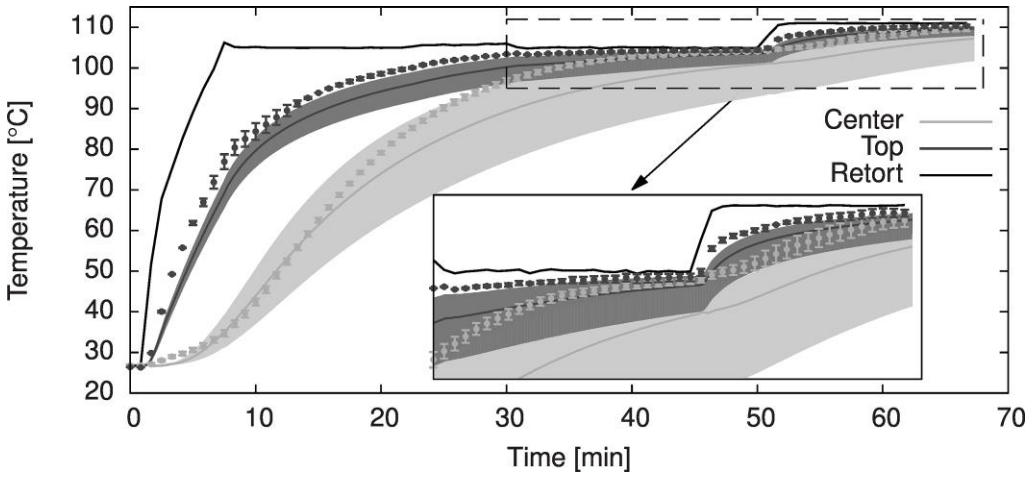

Figure 9. Simulation results vs. experimental data for the validation experiment. Both measurement points are represented. Gray bands represent the core predictions. Continuous gray lines were obtained with the mean value of the parameters computed from the Monte Carlo sampling method.

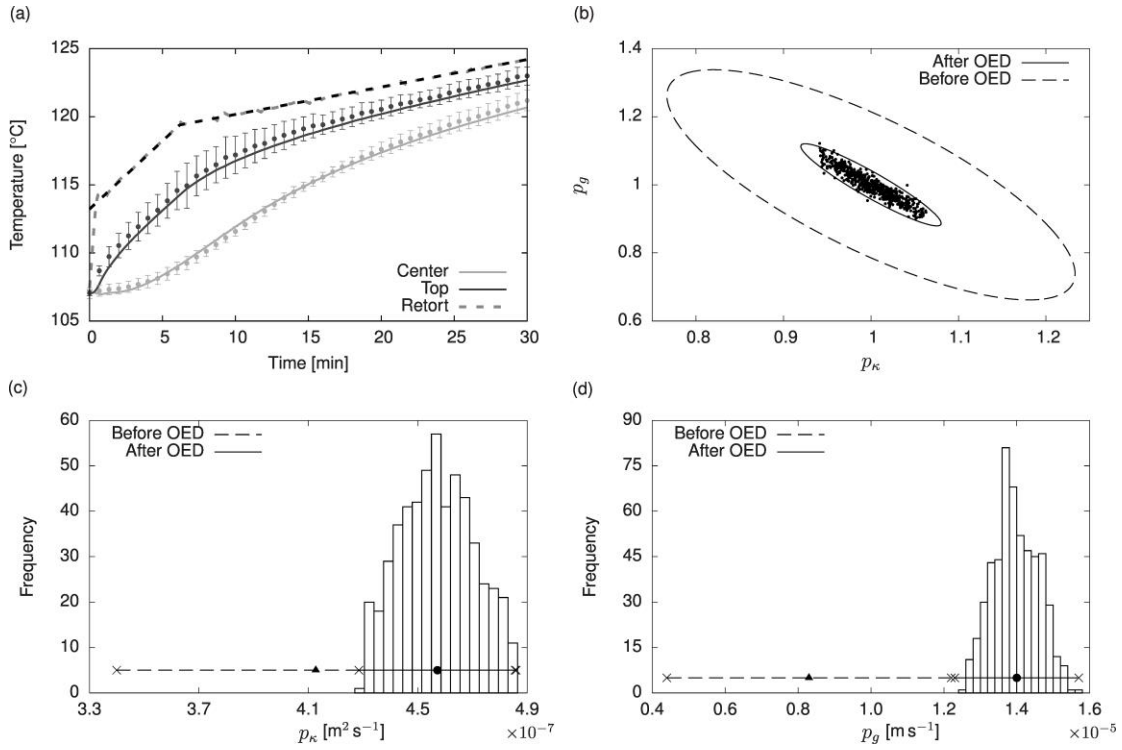

Figure 10. (a) Result of the OED and its implementation on the pilot plant. (b) Ellipses and cloud of points obtained from the robust identification procedure. (c) and (d) histograms of solutions from the Monte Carlo sampling method for parameters p_κ and p_g, respectively. Horizontal lines represent the parameter confidence intervals whereas triangle and circle marks are, respectively, the mean value of the Monte Carlo method before and after including the OED.

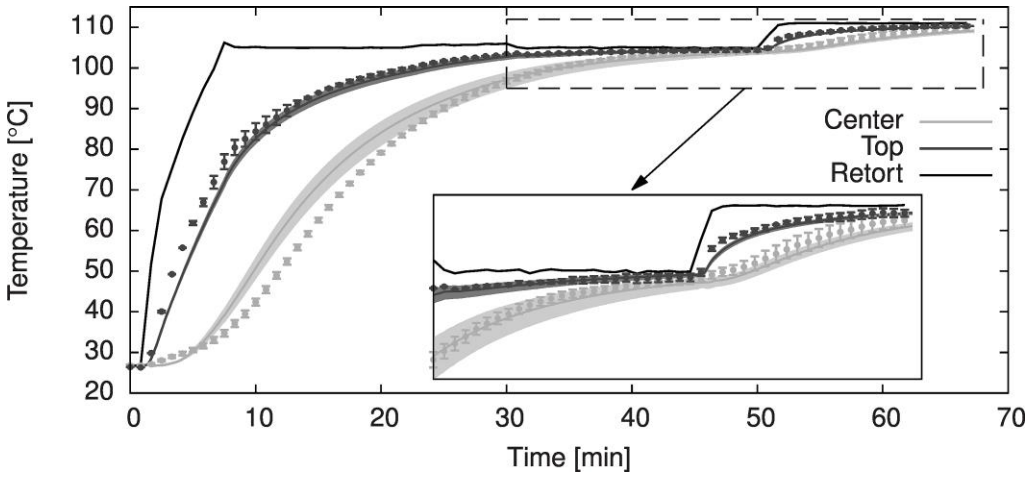

Figure 11. Model simulation using core predictions vs. experimental results in the validation experiment. Marks indicate the experimental data including error bars. Bands represent the core predictions. Continuous gray lines were obtained with the mean value for the parameters. Black line represents the retort temperature.

