Global Sensitivity Analysis as Good Modelling Practices tool for the identification of the most influential process parameters of the primary drying step during freeze-drying

Van Bockstal, Pieter-Jan ; Mortier, Séverine Thérèse F.C.; Corver, Jos; Nopens, Ingmar; Gernaey, Krist V.; De Beer, Thomas

Published in:
European Journal of Pharmaceutics and Biopharmaceutics

Link to article, DOI:

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Global Sensitivity Analysis as Good Modelling Practices tool for the identification of the most influential process parameters of the primary drying step during freeze-drying

Pieter-Jan Van Bockstal\textsuperscript{a,1,2,*}, Séverine Thérèse F.C. Mortier\textsuperscript{a,1}, Jos Corver\textsuperscript{a}, Ingmar Nopens\textsuperscript{b}, Krist V. Gernaey\textsuperscript{c}, Thomas De Beer\textsuperscript{a}

\textsuperscript{a}Laboratory of Pharmaceutical Process Analytical Technology, Department of Pharmaceutical Analysis, Faculty of Pharmaceutical Sciences, Ghent University, O’tempo steenweg 460, 9000 Ghent, Belgium
\textsuperscript{b}BIOMATH, Department of Mathematical Modelling, Statistics and Bioinformatics, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Ghent, Belgium
\textsuperscript{c}Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, 2800 Kgs. Lyngby, Denmark

Abstract

Pharmaceutical batch freeze-drying is commonly used to improve the stability of biological therapeutics. The primary drying step is regulated by the dynamic settings of the adaptable process variables, shelf temperature $T_s$ and chamber pressure $P_c$. Mechanistic modelling of the primary drying step leads to the optimal dynamic combination of these adaptable process variables in function of time. According to Good Modelling Practices, a Global Sensitivity Analysis (GSA) is essential for appropriate model building. In this study, both a regression-based and variance-based GSA were conducted on a validated mechanistic primary drying model to estimate the impact of several model input parameters on two output variables, the product temperature at the sublimation front $T_i$ and the sublimation rate $\dot{m}_{sub}$. $T_s$ was identified as most influential parameter on both $T_i$ and $\dot{m}_{sub}$, followed by $P_c$ and the dried product mass transfer resistance $\alpha_{Rp}$ for $T_i$ and $\dot{m}_{sub}$, respectively. The GSA findings were experimentally validated for $\dot{m}_{sub}$ via a Design of Experiments (DoE) approach. The results indicated that GSA is a very useful tool for the evaluation of the impact of different process variables on the model outcome, leading to essential process knowledge, without the need for time-consuming experiments (e.g., DoE).

Keywords: Freeze-drying, Mathematical modelling, Global Sensitivity Analysis

1. Introduction

Biological drug products like therapeutic proteins and vaccines, gain more and more interest within the pharmaceutical industry [1]. However, the stability of these biopharmaceuticals in aqueous solution is often limited due to water-mediated degradation pathways. Freeze-drying (lyophilization) is a frequently applied drying process to improve the stability of these biopharmaceuticals during storage and distribution, despite the long processing time and high costs [2]. Approximately 50% of the biopharmaceutical drug products approved by the regulatory authorities (> 300) are freeze-dried formulations [3]. Conventional pharmaceutical freeze-drying of unit doses is a batch-wise process during which all vials of

*Corresponding author
\textit{Email addresses: pieterjan.vanbockstal@ugent.be} (Pieter-Jan Van Bockstal), \textit{severine.mortier@ugent.be} (Séverine Thérèse F.C. Mortier), \textit{jos.corver@RheaVita.nl} (Jos Corver), \textit{ingmar.nopens@ugent.be} (Ingmar Nopens), \textit{KVG@kt.dtu.dk} (Krist V. Gernaey), \textit{thomas.debeer@ugent.be} (Thomas De Beer)

\textsuperscript{1}Phone number: +32(0)9 264 83 55
\textsuperscript{2}Fax number: +32(0)9 264 81 96
the same batch are processed through a sequence of consecutive steps (i.e., freezing, primary drying and secondary drying) until the dried end product is obtained [4]. The lyophilization process is initiated by gradually chilling the temperature-controlled shelves, which are packed with glass vials containing the aqueous drug formulation, to circa -45°C, depending on the formulation characteristics. During cooling of the aqueous solution, stochastic ice nucleation takes place leading to the formation of ice crystals [5]. These crystals start growing upon further freezing, leading to a gradual increase in solute concentration in between them. Freeze-concentration continues until at the eutectic temperature $T_e$, a saturated solution is obtained. From this point onward, water and solutes start to crystallize simultaneously, leading to the formation of a fully crystalline matrix. Amorphous products behave differently during freeze-concentration as they form a supersaturated solution, without any solute crystallization. Due to the gradual increase in viscosity associated with this supersaturation, ice crystallization slows down until maximum freeze-concentration is achieved at the glass transition temperature $T_g$. At this characteristic temperature, the viscosity has raised to such an extent that no further freezing occurs [5]. Hence, ice crystallization is interrupted prematurely, explaining the presence of residual unfrozen water in the amorphous glass. At the end of the freezing step, the shelf temperature, and consequently the product temperature, reach a value well below $T_e$ or $T_g$ for crystalline or amorphous products, respectively, resulting in complete solidification of the product. In some cases, the freezing stage is followed by an annealing step to allow full crystallization of crystalline components or to enhance the batch uniformity of the ice crystal size [6–8]. Here, the shelf temperature is temporarily increased to a temperature above $T_g$, but below $T_e$ and maintained for a few hours before re-establishing the final freezing temperature. At the start of the primary drying step, the pressure in the drying chamber is lowered until 5 to 30 Pa, before the shelf temperature is raised to supply energy to promote ice sublimation. The sublimation front, which separates the ice-free product from the ice crystal matrix, gradually moves downward until ice sublimation is finished. During the secondary drying stage, the desired residual moisture level is achieved through desorption of remaining unfrozen water in the amorphous matrix or by the removal of hydrate water from the crystalline product. At the end of the lyophilization process, after stoppering and capping under controlled conditions, the aqueous drug formulation is transformed into a dried cake with an increased shelf life.

The primary drying step during batch-wise freeze-drying is extensively described through mechanistic modelling [9–11]. These models allow the determination of the optimal combination of the adaptable process variables, i.e., shelf fluid inlet temperature $T_s$ and chamber pressure $P_c$, in function of the primary drying time, to achieve a dried product without any signs of cake collapse in a minimum time frame. The appearance of the dried product is one of the Critical Quality Attributes (CQAs) of freeze-dried products, as structural loss of the cake is undesirable because of aesthetic purposes and the negative impact on the reconstitution time [12]. Cake collapse occurs when the product temperature at the sublimation front $T_i$ exceeds the critical product temperature $T_{i,crit}$ during the primary drying stage [5]. $T_{i,crit}$ depends on the formulation characteristics and is identified as $T_e$ or the collapse temperature $T_c$ for crystalline and amorphous products, respectively. In general, $T_c$ is a few degrees higher than $T_g$ because viscous flow is limited near $T_g$ due to the sufficiently high viscosity of the amorphous solid [5].

As the mechanistic primary drying models are a mathematical approximation of the sublimation process, inherently a few assumptions and simplifications are included (e.g., a planar sublimation front and a steady-state system are assumed). In addition, the model input parameters are often an estimation of the real value, while other input variables (i.e., the dried product mass transfer resistance $R_p$ and the vial heat transfer coefficient $K_v$) are quantified by one global value for the entire batch, despite the inter-vial variability. Because of these reasons, the computational model output can be correlated with a degree of uncertainty, which could lead to a deviation from the actual (experimental) primary drying behaviour. This parameter uncertainty can be included in the model leading to the quantitative estimation of the Risk of Failure (RoF), i.e., the risk of cake collapse, for each combination of the adaptable process parameters, $T_s$ and $P_c$ [13–17].

This quantitative risk assessment is essential for the construction of the Design Space, defined as the multi-dimensional combination and interaction of input variables and process parameters leading to the expected product specifications with a controlled probability [18]. In this way, the optimal combination of $T_s$ and $P_c$ can be determined to maximize the process efficiency for a specific RoF acceptance level, which is defined as the chance of batch rejection due to macroscopic cake collapse in one or more vials [17]. The progression
of primary drying and the corresponding increase in dried layer thickness leads to a continuous change of
dependent process parameters (e.g., $R_p$). Therefore, the optimal combination of $T_s$ and $P_c$ changes in time,
which results in a dynamic Design Space [11, 17].
The determination of the impact of several process parameters and input variables on the primary drying
behaviour would require a lot of time-consuming experiments, even via a DoE approach. The information
gathered via experimental designs leads to the development of empirical models, which provide some, but
no full physical insight in the primary drying dynamics. Mechanistic modelling can offer an alternative as
it allows the identification of critical process parameters. In this respect, the application of a GSA on the
model is recommended for appropriate model building according to the Good Modelling Practices [19]. A
sensitivity analysis is the study of how the uncertainty in the output of a model (numerical or otherwise) can
be apportioned to different sources of uncertainty in the model input (or model structure, parameters) [20].
In combination with the theoretical and physical background used for the development of the mechanistic
model, a lot of process knowledge is obtained in an efficient way. Other incentives to perform a sensitivity
analysis, can be model calibration, model reduction and prediction of uncertainty [21]. In this study, a
regression-based and variance-based GSA were conducted on a, previously validated, mechanistic primary
drying model to estimate the impact of several model input parameters on two output variables, the product
temperature at the sublimation front $T_i$ and the sublimation rate $\dot{m}_{sub}$.

2. Materials and methods

2.1. Freeze-drying model

The detailed description of the (validated) mechanistic model describing the primary drying step of a
batch freeze-drying process is given by Mortier et al. [16]. The nominal values of the model input parameters
are listed in table 1, which are equal to the nominal values previously used in a published study regarding
the computation of the dynamic Design Space for an aqueous formulation containing 30 mg/mL sucrose
(Sigma-Aldrich, Saint-Louis, MO, USA) [17]. The coefficients describing $R_p$ for the model formulation were
obtained from literature [22]. $T_i, crit$, which was assumed to be equal to $T_g'$ of the model formulation, and the
coefficients describing $K_v$ were experimentally determined in the context of the dynamic Design Space study
[17]. The mechanistic primary drying model assumes both a planar sublimation front and a steady-state
system, i.e., the supplied energy was only consumed for ice sublimation.

Table 1: Nominal values of the parameters of the model describing the primary drying step

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inner radius of 10R vial</td>
<td>0.011 m</td>
</tr>
<tr>
<td>Outer radius of 10R vial</td>
<td>0.012 m</td>
</tr>
<tr>
<td>Radius of the vial neck</td>
<td>0.0063 m</td>
</tr>
<tr>
<td>Radius of the duct of the dryer</td>
<td>0.08 m</td>
</tr>
<tr>
<td>$K_v$-coefficient</td>
<td>11.18 J/(m²sK)</td>
</tr>
<tr>
<td>$K_v$-coefficient</td>
<td>1.435 J/(m²sKPa)</td>
</tr>
<tr>
<td>$K_v$-coefficient</td>
<td>0.04008 1/Pa</td>
</tr>
<tr>
<td>$R_p$-coefficient</td>
<td>9.124 10⁶ 1/s</td>
</tr>
<tr>
<td>$R_p$-coefficient</td>
<td>0.1/m</td>
</tr>
<tr>
<td>$T_i, crit$</td>
<td>-32.5°C</td>
</tr>
<tr>
<td>Filling volume</td>
<td>2.5 mL</td>
</tr>
<tr>
<td>Density of ice</td>
<td>919.4 kg/m³</td>
</tr>
<tr>
<td>Volume fraction of ice</td>
<td>0.97</td>
</tr>
<tr>
<td>Ratio of the specific heat for water</td>
<td>1.33</td>
</tr>
<tr>
<td>Molecular weight of water</td>
<td>0.018015 kg/mol</td>
</tr>
</tbody>
</table>
2.2. Global Sensitivity Analysis (GSA)

A GSA was conducted on two model outputs, i.e., the product temperature at the sublimation front $T_i$ and the sublimation rate $\dot{m}_{\text{sub}}$. The sensitivity of the parameters was calculated after different primary drying times, i.e., after 1, 2, 3 and 4 h. Initially, the model is run for different combinations of input variables based on their uncertainty range, i.e., Monte Carlo simulations. The sample matrix required for the calculation of the output vector, is generated using the Sobol sampling technique. The number of Monte Carlo simulation runs has an influence on the sensitivity indices. For this reason, the optimal number of samples was determined in a first step, before the regression-based and variance-based GSA methods were conducted.

The parameters involved in the GSA, their uncertainty level and the reason for their inclusion are listed in Table 2. The estimation of the uncertainty level for each parameter and the reason for their inclusion is previously described [17]. For most factors, the uncertainty was not implemented as a percentage of the nominal value, as the range is highly dependent on the magnitude of the factor. For these factors (i.e., $T_s$, $P_c$, inner vial radius $r_{v,i}$ and outer vial radius $r_{v,o}$), the uncertainty was defined by an absolute range, based on experimental data or data provided by the supplier. In this study, the uncertainty on $L_{\text{dried}}$ was assumed to be constant. However, it has to be remarked that the uncertainty on $L_{\text{dried}}$ can also be calculated based on the other (uncertain) model input variables (i.e., error propagation), as previously demonstrated for the uncertainty analysis [17]. $K_v$ is described in function of $P_c$ via [16]:

$$K_v = \alpha + \beta P_c \frac{1}{1 + \gamma P_c}$$

(1)

with $\alpha$ (J/(m²sK)), $\beta$ (J/(m²sKPa)) and $\gamma$ (1/Pa) constants describing $K_v$ quantitatively in function of $P_c$ (Table 1). $R_p$ is described in function of $L_{\text{dried}}$ via [16]:

$$R_p = R_{p,0} + \frac{A_{R_p} L_{\text{dried}}}{1 + B_{R_p} L_{\text{dried}}}$$

(2)

with $R_{p,0}$ (m/s), $A_{R_p}$ (1/s) and $B_{R_p}$ (1/m) constants describing $R_p$ quantitatively in function of $L_{\text{dried}}$ (Table 1). The uncertainty on these coefficients is lumped by introducing an uncertainty on the entire parameter, i.e., $\alpha_{R_p}$ and $\alpha_{K_v}$. The uncertainty on $r_{v,i}$ and $r_{v,o}$ is based on data provided by the supplier of the 10R glass vials (Schott, Müllheim, Germany) used during the experimental validation (Section 2.3). The filling volume $V$ was included in the analysis because of the inherent error that is made during the filling stage. Depending on the location of the vial on the shelf, $T_s$ can be higher or lower than the set value. This error has been experimentally determined for both a constant value of $T_s$ and during a shelf temperature ramp [17]. The error on $P_c$ is based on the maximum allowed deviation of the feedback pressure control system of the freeze-dryer.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Uncertainty level</th>
<th>Reason of inclusion in GSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_{\text{dried}}$</td>
<td>5%</td>
<td>Error propagation</td>
</tr>
<tr>
<td>$R_p$</td>
<td>20%</td>
<td>Based on experimental data</td>
</tr>
<tr>
<td>$K_v$</td>
<td>20%</td>
<td>Based on experimental data</td>
</tr>
<tr>
<td>$r_{v,i}$</td>
<td>$10^{-4}$ m</td>
<td>Process variable with inherent error</td>
</tr>
<tr>
<td>$r_{v,o}$</td>
<td>$10^{-4}$ m</td>
<td>Process variable with inherent error</td>
</tr>
<tr>
<td>$V$</td>
<td>10%</td>
<td>Process variable with inherent error</td>
</tr>
<tr>
<td>$T_s$</td>
<td>2.5 °C</td>
<td>Based on experimental data</td>
</tr>
<tr>
<td>$P_c$</td>
<td>1 Pa</td>
<td>Process variable with inherent error</td>
</tr>
</tbody>
</table>
2.2.1. Regression-based sensitivity analysis

The use of the regression-based GSA technique allows the sensitivity ranking to be determined based on the relative magnitude of the Standardized Regression Coefficients (SRCs). Therefore, the ranking is based on a linear regression model on the output of the Monte Carlo simulations via least square computation, not on the original model [23, 24]. For this method, it is assumed that no correlation exists between the different input factors. The linear regression is evaluated via the model coefficient of determination $R^2_Y$, which is equal to the fraction of variance of the original data, i.e., the results of the Monte Carlo simulation, explained by the regression model [23]. Ideally, for all regression-based methods, $R^2_Y$ is close to 1, meaning that most of the model output variance is explained by the regression coefficients [24]. A low value for $R^2_Y$ indicates non-linear model behaviour where the SRCs fail to provide an adequate ranking [25]. The minimum criterion for the linear regression is a value of 0.7 for $R^2_Y$ [26, 27]. The disadvantage of the regression-based method is that it is computationally expensive, because many samples are required. However, after simulation of the model, the analysis is fast and all measures can be calculated using the same sample and model. The obtained SRCs can only be used to assess the importance of a given parameter. Moreover, it is difficult to determine the number of samples to obtain valuable results. The details regarding the calculation of the SRCs are described by Mortier et al. [21].

2.2.2. Variance-based sensitivity analysis

Variance-based methods are very popular nowadays, and the obtained indices are much more informative compared to the SRCs. These indices are derived from the decomposition of the total output variance into the contribution of the input factors [20, 24, 28]. Two sensitivity measures are calculated by the variance-based GSA, i.e., the first order effect $S_i$ and the total order effect $S_{Ti}$. $S_i$ indicates the actual fraction of variance accounted for by each factor, whereas $S_{Ti}$ represents the total effect of factor $i$, i.e., the sum of the first order effects, higher order effects and all interactions with other factors [25]. Different designs are known to calculate these sensitivity indices [29, 30]. The design proposed by Saltelli et al. was applied in this study [31]. The mathematical details regarding the computation of $S_i$ and $S_{Ti}$ are described by Mortier et al. [21].

2.2.3. Determination of sample size

The Central Limit Theorem (CLT) was used to determine the optimal number of samples $N$ required for both the regression-based and variance-based GSA, i.e., the number of samples required to achieve a normal distribution of the composite [32]. Therefore, for an increasing number of base samples (different values for $N$), $R$ replicas were taken and the corresponding sensitivity measures were calculated, i.e., the SRCs for the regression-based technique and $S_i$ and $S_{Ti}$ for the variance-based technique. For each $N$ value, this leads to $R$ values for each sensitivity measure. The mean and the coefficient of variation $CV$ was calculated. According to the CLT, $R$ should be at least 30, however, in this contribution 100 replicas were used [32].

2.3. Experimental design methodology

The results of the GSA were experimentally validated via a DoE approach for one specific response, i.e., the sublimation rate $\bar{m}_{sub}$. The parameters used for the GSA were included as factors in the experimental design (Table 3), except for the vial dimensions $r_{v,i}$ and $r_{v,o}$, which were kept constant as we did not deem these variables to provide more scientific insight in this stage. The center points for each factor of the DoE were derived from the nominal values of the model parameters used for the computation of the dynamic Design Space for an aqueous 30 mg/mL sucrose formulation (Table 1), except for $T_s$ and $P_e$, which were based on the output of this model [17]. The lowest and highest level for the quantitative factors $T_s$, $P_e$ and $V$ were based on their corresponding uncertainty level in the GSA (Table 2). Both $R_p$ and $K_e$ were treated as two level qualitative factors which approach the relative uncertainty of 20% on the nominal value, as included in the GSA (Table 2). Specifically, the uncertainty on $R_p$ was incorporated in the design by including an aqueous 30 mg/mL lactose formulation (Fagron, Waregem, Belgium), for which the $R_p$ profile in function of $L_{dried}$ was approximately 20% higher compared to the nominal value for the 30 mg/mL sucrose model formulation [22]. The uncertainty on $K_e$ was incorporated by comparing the sublimation rate for the edge
vials and center vials. A fractional factorial screening design consisting of 16 experiments, expanded with two double center points, was constructed using MODDE Pro (Version 11.0.0, Umetrics, Umeå, Sweden). An overview of all individual experiments is given in Table 4.

The response $\dot{n}_{\text{sub}}$ was gravimetrically determined. The experiments were conducted in a laboratory scale freeze-dryer Lyobeta 25 (Telstar, Terrassa, Spain) equipped with four shelves, each with a heatable area of 0.16 m$^2$, in the vacuum-tight drying chamber with a volume of 0.2 m$^3$. The system contains a condenser with 35 kg of ice capacity and a vacuum pump to evacuate the drying chamber. Chamber pressure control was performed via a Baratron Type 626A capacitance manometer (MKS Instruments, Andover, MA, USA). Comparative pressure measurements were conducted using a Pirani Type PSG-502-S thermal conductivity gauge (Inficon, Bad Ragaz, Switzerland). Temperature control of shelf and condenser was based on Pt100 thermal sensors (WIKA Instruments, Klingenberg, Germany) measurements. For each individual experiment of the DoE, 10 vials were filled with a certain volume of the specified formulation, as defined by the DoE (Table 4). The vials were placed on a precooled shelf (3°C) in the drying chamber, randomly distributed at the edge or center of the batch. The batch was expanded with dummy vials containing 2.5 mL of water to a total of 100 vials. The size and arrangement of the batch was equal to the experimental determination of $K_v$ [17]. The freezing step for each validation run was equal to the freezing procedure used to determine $R_p$.

### Table 3: Overview factors DoE with lowest and highest level

<table>
<thead>
<tr>
<th>Factors</th>
<th>Level</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_s$</td>
<td>-18.5</td>
<td>°C</td>
</tr>
<tr>
<td>$P_c$</td>
<td>9</td>
<td>Pa</td>
</tr>
<tr>
<td>$V$</td>
<td>2.25</td>
<td>mL</td>
</tr>
<tr>
<td>Qualitative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_p$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_v$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4: Overview experiments DoE, the experiments in bold were performed twice (center points)

<table>
<thead>
<tr>
<th>Experiment number</th>
<th>$T_s$ (°C)</th>
<th>$P_c$ (Pa)</th>
<th>$V$ (mL)</th>
<th>$K_v$</th>
<th>$R_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-18.5</td>
<td>9</td>
<td>2.75</td>
<td>Formulation 1</td>
<td>Edge vial</td>
</tr>
<tr>
<td>2</td>
<td>-13.5</td>
<td>9</td>
<td>2.25</td>
<td>Formulation 1</td>
<td>Edge vial</td>
</tr>
<tr>
<td>3</td>
<td>-18.5</td>
<td>11</td>
<td>2.75</td>
<td>Formulation 1</td>
<td>Edge vial</td>
</tr>
<tr>
<td>4</td>
<td>-13.5</td>
<td>11</td>
<td>2.25</td>
<td>Formulation 1</td>
<td>Edge vial</td>
</tr>
<tr>
<td>5</td>
<td>-18.5</td>
<td>9</td>
<td>2.25</td>
<td>Formulation 2</td>
<td>Edge vial</td>
</tr>
<tr>
<td>6</td>
<td>-13.5</td>
<td>9</td>
<td>2.75</td>
<td>Formulation 2</td>
<td>Edge vial</td>
</tr>
<tr>
<td>7</td>
<td>-18.5</td>
<td>11</td>
<td>2.75</td>
<td>Formulation 2</td>
<td>Edge vial</td>
</tr>
<tr>
<td>8</td>
<td>-13.5</td>
<td>11</td>
<td>2.25</td>
<td>Formulation 2</td>
<td>Edge vial</td>
</tr>
<tr>
<td>9</td>
<td>-18.5</td>
<td>9</td>
<td>2.25</td>
<td>Formulation 1</td>
<td>Center vial</td>
</tr>
<tr>
<td>10</td>
<td>-13.5</td>
<td>9</td>
<td>2.75</td>
<td>Formulation 1</td>
<td>Center vial</td>
</tr>
<tr>
<td>11</td>
<td>-18.5</td>
<td>11</td>
<td>2.75</td>
<td>Formulation 1</td>
<td>Center vial</td>
</tr>
<tr>
<td>12</td>
<td>-13.5</td>
<td>11</td>
<td>2.25</td>
<td>Formulation 1</td>
<td>Center vial</td>
</tr>
<tr>
<td>13</td>
<td>-18.5</td>
<td>9</td>
<td>2.75</td>
<td>Formulation 2</td>
<td>Center vial</td>
</tr>
<tr>
<td>14</td>
<td>-13.5</td>
<td>9</td>
<td>2.25</td>
<td>Formulation 2</td>
<td>Center vial</td>
</tr>
<tr>
<td>15</td>
<td>-18.5</td>
<td>11</td>
<td>2.25</td>
<td>Formulation 2</td>
<td>Center vial</td>
</tr>
<tr>
<td>16</td>
<td>-13.5</td>
<td>11</td>
<td>2.75</td>
<td>Formulation 2</td>
<td>Center vial</td>
</tr>
<tr>
<td>17</td>
<td>-16</td>
<td>10</td>
<td>2.5</td>
<td>Formulation 1</td>
<td>Edge vial</td>
</tr>
<tr>
<td>18</td>
<td>-16</td>
<td>10</td>
<td>2.5</td>
<td>Formulation 2</td>
<td>Edge vial</td>
</tr>
</tbody>
</table>
of the model formulation [22]. The shelves were chilled at a rate of 1 °C/min until -45 °C and maintained at this temperature for 1 hour. An annealing step was included in the freezing procedure, during which $T_s$ was increased in 25 minutes until -20 °C and kept at this value for 2 hours. After re-establishing the final freezing temperature of -45 °C at a rate of 1 °C/min, $T_s$ was kept for another hour at this value, before the condenser was cooled and the vacuum pump was activated. The settings of $T_s$ and $P_c$ during primary drying were characteristic for each individual DoE experiment. Primary drying was interrupted after approximately 7 to 8h, to avoid ice sublimation was completely finished. The vials were weighed before and after freeze-drying, allowing the calculation of the mean value of $\dot{m}_{sub}$.

MODDE Pro was used for the analysis of the experimental design. A regression model for the response was calculated via Multiple Linear Regression (MLR). All factors were scaled and centered, which makes the regression coefficients for the different factors comparable. Each quantitative factor was set to +1 and −1 for their highest and lowest value, respectively. Qualitative factors require an alternative approach via a mathematical re-expression (regression coding). A k-level qualitative parameter will be expanded into k-1 artificial categorical variables. Each expanded term results in one regression coefficient, which in our case led to one coefficient for level 2 (regular mode). In addition, the level 1 coefficient is equal to the negative coefficient of the expanded term of level 2 (extended mode). The regression coefficient of a specific factor represents the quantitative change in response value when this factor is increased from its average to its high level, keeping all other factors at their average value. The effect of a specific factor is defined as the quantitative change in response value when this factor is increased from its low to its high level, keeping all other factors at their average value. Based on these definitions, the effect of a factor is twice the corresponding coefficient. The constant term is related to the response value at the design center point, with all factors at their nominal value. For all regression coefficients the 95% confidence interval was calculated. Factors or interactions between factors were considered significant if the 95% confidence interval of the corresponding regression coefficient did not contain zero.

3. Results

3.1. Determination of the required number of simulations for the GSA

To determine the number of samples required for the GSA, the convergence of the sensitivity measures was verified for a base scenario, i.e., the output after 1,800 s. This could be done on a graphical basis, by visually checking the convergence. However, a more objective evaluation was conducted using a technique based on CLT [32].

For the regression-based GSA, the coefficient of determination $R^2_Y$ should be higher than 0.7. Only when $R^2_Y$ is higher than 0.7, the SRCs can be used to rank the parameters of the mechanistic primary drying model. It was possible to perform the analysis for both output variables, i.e., $T_i$ and $\dot{m}_{sub}$, as $R^2_Y$ was above 0.7, without the need for a rank transformation. On the left side of figure 1, the mean of the quotient of two successive SRC values is presented for an increasing number of simulations $N$ for $T_i$ as output variable. For each parameter, the quotient of two successive SRC values should converge to 1 for an increasing value of $N$, therefore, the mean over all parameters should also converge towards one. However, the question arises for which $N$ the value is close enough to 1. Therefore, a more in-depth method has been used to objectively determine the optimal number of samples, i.e., CLT. Besides the mean quotient of two successive SRC values, the $R^2_Y$-value also converges for an increasing number of Monte Carlo simulations (Right part of figure 1).

In the left part of figure 2, the coefficient of variation $CV$ is plotted in function of the number of base samples $N$ for $T_i$ as output variable. From the graph it can be concluded that this coefficient decreases for a higher number of base samples. After 10,000 samples, the solution seems to be converged, but the optimal number of samples is determined graphically and is not exact. There exists no quantitative method in order to determine the optimal number. From this figure it is apparent that $CV$ is remarkable higher for $V$ and $L_{dried}$ compared to the other parameters. $CV$ in function of $N$ for the sublimation rate $\dot{m}_{sub}$ as output variable is presented on the right side of figure 2. $CV$ is much higher for this output variable, and moreover, the behaviour for $L_{dried}$ is different compared to the other parameters. Based on these results, and given
Figure 1: Mean of the quotient of two successive SRC values (Left) and the $R^2$-value (Right) for increasing number of Monte Carlo simulations performed on the primary drying model for $T_i$ as output variable of the regression-based GSA

the fact that it has no impact on the mean of the SRCs (Figure 3), it was decided to use 10,000 samples for the regression-based GSA.

The same approach was applied for the variance-based technique. The results for the total sensitivity

Figure 2: Coefficient of variation $CV$ for the SRC for an increasing number of base samples $N$ performed on the primary drying model for $T_i$ (Left) and $\dot{m}_{sub}$ (Right) as output variable of the regression-based GSA

Figure 3: Mean of the SRC in function of the number of base samples $N$ performed on the primary drying model for $T_i$ (Left) and $\dot{m}_{sub}$ (Right) as output variable of the regression-based GSA
indices $S_{Ti}$ are presented in figure 4 for the temperature at the sublimation front $T_i$ and the sublimation rate $\dot{m}_{sub}$. It should be noted that the actual number of model runs is presented on the x-axis. Dependent on the formulas used to calculate $S_i$ and $S_{Ti}$, the actual number of model runs varies according to the set sample size. Here, the scheme proposed by Saltelli et al. was used [31]. This means that the actual number of models runs equals $(n_{par} + 2)N_{base}$ with $n_{par}$ the number of parameters included in the sensitivity analysis and $N_{base}$ the base sample size.

The variance-based sensitivity technique does not require a linear model as input, therefore, the method has a broader applicability. To have a converged result for the temperature at the sublimation front $T_i$ (Left part of figure 4), a sample size of 5,000 seems to be needed. The same is valid for the sublimation rate $\dot{m}_{sub}$ (Right part of figure 4). Therefore, a 500 base samples, i.e., 5,000 divided by $n_{par} + 2$, were chosen to perform the further analysis.

![Figure 4: Coefficient of variation CV for the $S_{Ti}$ in function of the number of base samples $N$ performed on the primary drying model for $T_i$ (Left) and $\dot{m}_{sub}$ (Right) as output variable of the variance-based GSA](image)

### 3.2. Gathering process knowledge using graphical tools

Graphical sensitivity analysis tools, which are qualitative techniques (i.e., scatter plots or Contribution to Sample Mean (CSM)/Contribution to Sample Variance (CSV) plots), are interesting to detect the relationship between uncertain model inputs and the resulting output [33]. The scatter plots are presented for two parameters with a different behaviour, after 1 h of simulated process time (Figure 5). From the scatter plot of $P_c$, it can be concluded that the influence of $P_c$ on $T_i$ is rather limited, as the same value of $T_i$ can be obtained for each value of the chamber pressure. The effect of $\alpha_{Rp}$ on the sublimation rate (Right part of figure 5) is more pronounced as for lower values of $\alpha_{Rp}$ the sublimation rate is higher due to the reduced product resistance to vapour flow.

### 3.3. GSA using a regression- and variance-based technique

To rank the parameters it is suggested to perform a regression- or variance-based GSA technique, where quantitative sensitivity measures are obtained. The GSA was performed for different time steps during the primary drying process. In figure 6, $S_{Ti}$ values for the different parameters are presented for different time steps. Focusing at the results after 1 h for $T_i$ as output variable (Left side of 6), it can be concluded that $T_s$ is the most influential parameter, followed by $P_c$. The significance of $T_s$ is important as this is one of the adaptable process parameters of the freeze-drying process. The other process variable, i.e., $P_c$, has a rather limited influence. Looking at the results for the whole drying process, it is obvious that $T_s$ has the most influence on $T_i$. The influence of $P_c$ is somewhat decreasing over time. The influence of $\alpha_{Rp}$ and $\alpha_{Kv}$ is clearly increasing over time.

On the right side of figure 6, the results for $S_{Ti}$ with $\dot{m}_{sub}$ as output variable are presented. Again, $T_s$ is the most influential parameter, followed by $\alpha_{Rp}$. The influence of $T_s$ is decreasing and the influence for $P_c$ is
increasing, but for the other parameters the value for $S_{Ti}$ is similar over the whole time range. The results using the regression-based analysis with the SRCs are similar and not shown.

3.4. Analysis of experimental design

The mean sublimation rate $\dot{m}_{sub}$ for the individual experiments of the experimental design (Table 4) ranged from approximately 0.15 g/h up to 0.27 g/h. For the filling volume of 2.5 mL, assuming an ice mass of 2.5 g, this would correspond to an estimated primary drying time ranging from approximately 16.67 h to 9.25 h. Evaluation of the raw data indicated that there were no peculiarities in the results: the variability in $\dot{m}_{sub}$ between the repetitions (i.e., center points) was less than the overall variability for the individual experiments. The response data were normally distributed, hence, data transformation before regression analysis was not needed. The regression coefficients, computed via MLR fitting, are displayed for each factor and factor interaction, including their 95% confidence interval, in the extended coefficient plot in Figure 7.

The position of the vial in the batch, i.e., $K_v$(Edge vial) and $K_v$(Center vial), was identified as the factor with the highest influence on $\dot{m}_{sub}$, with a regression coefficient of 0.0317 g/h (Figure 7). For a vial situated at the edge of the shelf, $\dot{m}_{sub}$ increases with 0.0635 g/h compared to a center vial, keeping all other factors at their average level, hence, decreasing the primary drying time with approximately 27% from 14.53 h to 10.61 h for a 2.5 mL filling volume. The shelf fluid inlet temperature $T_s$ had a positive impact on the response
Figure 7: Extended coefficient plot: Regression coefficients displayed including 95% confidence interval, with shelf fluid inlet temperature (Tₛ), chamber pressure (Pₑ), filling volume (V), the expanded model terms related to the qualitative factor Rₛ (Rₛ(For1) and Rₛ(For2)) and the expanded model terms related to the qualitative factor Kᵥ (Kᵥ(Edge vial) and Kᵥ(Center vial)).

variable (regression coefficient 0.0183 g/h). An increase in Tₛ from -16°C to -13.5°C, keeping all other factors at their average level, is associated with a decrease in drying time of approximately 1 h from 12.26 h to 11.25 h. The difference in dried product mass transfer resistance R_p, represented by the two different formulations, i.e., R_p(For1) and R_p(For2), was small with a regression coefficient of 0.00524 g/h. Primary drying of the lactose formulation would take 5% or 38 min longer compared to the sucrose formulation with a primary drying time of 12.59 h and 11.96 h, respectively, keeping all other factors at their average level. With a regression coefficient of 0.00382 g/h, the chamber pressure Pₑ had a very small impact on \( \dot{m}_{sub} \). An increase of Pₑ from 10 Pa to 11 Pa, keeping all other factors at their average level, is associated with a decrease in drying time of approximately 13 min from 12.26 h to 12.04 h. The filling volume V had no significant impact on the response \( \dot{m}_{sub} \) (regression coefficient 0.001003 g/h). In addition, none of the factor interactions was significant.

4. Discussion

The results of the GSA on the mechanistic primary drying model, i.e., the significance ranking \( S_{T_i} \) for both T_i and \( \dot{m}_{sub} \), should be interpreted considering the physical basis of the sublimation process. As one of the adaptable process variables, Tₛ had the largest impact on both T_i and \( \dot{m}_{sub} \) (Figure 6). The shelves provide the energy transfer to the glass vials which impacts both T_i and \( \dot{m}_{sub} \). Therefore, Tₛ is the process parameter with the highest influence on these output variables. The experimental results confirm the significant impact of Tₛ on \( \dot{m}_{sub} \) (Figure 7). The relative impact of Tₛ decreases when the primary drying process progresses (Figure 6) due to the dynamic nature of the optimized cycle as R_p increases in function of time, requiring a gradual decrease in energy input to avoid cake collapse [17]. The other adaptable parameter of the primary drying process, Pₑ, was also identified in the GSA as a factor with a significant impact, mainly on T_i. Pₑ directly impacts \( \dot{m}_{sub} \), as the sublimation rate is based on the difference in vapour pressure at the sublimation front and in the drying chamber. Because of the high accuracy of the pressure control system of the freeze-dryer, the uncertainty level for Pₑ is very narrow, i.e., only 1 Pa, explaining why the impact of Pₑ on \( \dot{m}_{sub} \) is rather limited. Due to the narrow uncertainty range, Pₑ was identified as a factor with a very limited impact on the experimental \( \dot{m}_{sub} \). In addition, it should be noted that Pₑ influences the conductive energy transfer. An increase in Pₑ is related to a higher density of gas molecules in the drying chamber, which enhances the conductive energy transfer from shelf to vial due to
the suboptimal contact between vial and shelf \[34\]. The impact of \( P_t \) could be more relevant in situations where \( T_f \) becomes closer to \( T_i \), for instance in case of a higher value for \( V \) or \( R_p \) or in case of a formulation with a low \( T_w \) which requires a limitation of the energy transfer.

The GSA predicted that the impact of \( \alpha_{R_p} \) on \( \dot{m}_{sub} \) was less pronounced compared to \( T_f \) (Figure 6). However the influence of \( \alpha_{K_v} \) is still significant, especially towards \( \dot{m}_{sub} \). \( R_p \) directly impacts \( \dot{m}_{sub} \), as the sublimation rate is inversely proportional to the resistance of the dried product layer. This is confirmed by the results of the experimental design, where \( R_p \) has a significant, but limited impact. For equal process conditions, the primary drying step lasts 5% longer for the 30 mg/mL lactose formulation compared to the 30 mg/mL sucrose formulation (Section 3.4). \( R_p \) is highly dependent on the formulation characteristics and the pore size of the dried product layer, which, in turn, depends on the ice crystal size, related to the degree of supercooling before ice nucleation takes place \[22, 35\]. Therefore, cycle optimization can only be performed for a specific freezing regime. The stochastic nature of ice nucleation and the resulting interval variability in pore size distribution is countered by including a sufficiently high uncertainty level for \( R_p \).

\( \alpha_{K_v} \) is the only parameter of which the effect on \( \dot{m}_{sub} \) is underestimated by the GSA, compared to the experimental validation. Based on the experimental design, the position of the vial, at the edge of the shelf or in the center of the vial pack, is the parameter with the highest influence on \( \dot{m}_{sub} \). These experimental findings stroke with the difference in \( K_v \) between the edge and center vials, due to the additional radiative heat transfer coming from the surrounding walls and the door \[17, 34, 36\]. This uneven energy transfer between both positions is inherent to the design of the pharmaceutical batch freeze-dryer. Therefore, an appropriate freeze-drying cycle operating strategy is required as a compensation. This parameter was experimentally determined for separate groups. One group consisted of vials completely surrounded by other vials, i.e., center vials, while the other group contained vials situated at the edge of the shelf, with at least one side directed towards the door or the walls of the drying chamber, i.e., edge vials \[17\]. The dynamic Design Space was computed based on these edge vials, as these vials are the limiting factor in the prevention of cake collapse due to the additional radiative energy transfer. As compensation for the center vials with a lower energy input, the optimal primary drying trajectory should be prolonged before the start of the secondary drying step \[17\]. It should be noted that the ratio of edge to center vials decreases considerably during upscaling to industrial scale freeze-driers, however, the edge vials will remain the limiting factor and the same strategy should be applied.

The computed impact of \( r_{v,i}, r_{v,o}, \) and \( V \) was very limited. Based on the data provided by the supplier, the uncertainty margin on \( r_{v,i}, r_{v,o} \) is very small and the GSA indicates that their impact is negligible. The experimental design confirmed the low impact of \( V \) on \( \dot{m}_{sub} \).

In general, the experimental validation seems to confirm the results of the GSA, with an exception for \( \alpha_{K_v} \). It is important to note that the results of the GSA are highly dependent on the used uncertainty level. Therefore, the uncertainty ranges for each parameter should be carefully chosen. In addition, the uncertainty range should also be considered when interpreting the results of the GSA. For experimental validation of the GSA, it is equally important to choose an appropriate level for each factor of the design, correctly reflecting the uncertainty level for each process variable included in the GSA. For quantitative factors, this can be easily done by adapting the input settings, e.g., \( P_t \) can be set at a specific value via the pressure control system. However, \( K_v \) and \( R_p \) could only be included in the design as qualitative factors, which makes it more difficult to approach the quantitative uncertainty range used in the GSA. This might be the reason why the impact of \( \alpha_{K_v} \) is underestimated in the GSA compared to the experimental validation. As the GSA was conducted on a validated model, essential process knowledge can be extracted. The method allows to determine which uncertain input variables only have a limited influence on the model outcome, even to such extent their uncertainty could be irrelevant. Also, GSA can identify the parameter for which a reduction or elimination of the uncertainty would reduce most of the variance on the output and, hence, lower the risk of cake collapse during primary drying. The experimental DoE was conducted to validate the GSA on the previously validated primary drying model. Once validated, it is no longer required to repeat the post DoE in case different nominal values or uncertainty levels would be used for the model input parameters, e.g., during upscaling or during the optimization of the primary drying cycle for another formulation. This way, similar information can be obtained in a much shorter time frame compared to the labor intensive and more expensive DoE approach. However, freeze-drying cycles designed and optimized
via the mechanistic primary drying model, do require a final experimental validation [17].

5. General conclusion

A GSA has been applied on a, previously validated, mechanistic primary drying model in order to gain process knowledge. First, the optimal number of samples has been determined by using CLT. For the regression-based GSA, 10,000 samples was sufficient to obtain converged SRCs. The variance-based GSA required 500 base samples, corresponding to 5,000 samples \((n_{\text{par}} + 2)N_{\text{base}}\). For both methods, the ranking of the parameters was similar for both output variables of the model, i.e., the temperature at the sublimation front \(T_i\) and the sublimation rate \(\dot{m}_{\text{sub}}\). The shelf temperature \(T_s\) was identified as the factor with the highest impact on both output variables. The chamber pressure \(P_c\) has a large influence on \(T_i\), however, the impact on \(\dot{m}_{\text{sub}}\) is rather limited. The influence of \(\alpha_{K_v}\) on \(\dot{m}_{\text{sub}}\) is significant.

The results of the GSA were largely confirmed by the experimental validation via a DoE approach. Only, the impact of \(\alpha_{K_v}\) was underestimated by the GSA, as this parameter was included as a qualitative factor in the experimental design. The level of uncertainty of each process parameter was identified as a major influence on the significance ranking. Therefore, the uncertainty level of each process parameter should be determined with reason and not be chosen randomly. In summary, the results indicated that GSA is a very useful tool for the evaluation of the impact of different process variables on the model outcome, leading to essential process knowledge, without performing time-consuming experiments.

Acknowledgment

Financial support for this research from the Fund for Scientific Research Flanders (FWO Flanders - Ph.D. fellowship Séverine Thérèse F.C. Mortier) is gratefully acknowledged.

Nomenclature

\(\alpha\) Constant describing \(K_v\) in function of \(P_c\) \((J/(m^2sK))\)

\(\beta\) Constant describing \(K_v\) in function of \(P_c\) \((J/(m^2sKPa))\)

\(\dot{m}_{\text{sub}}\) Sublimation rate \((kg/s)\)

\(\epsilon\) Volume fraction of ice (-)

\(\gamma\) Constant describing \(K_v\) in function of \(P_c\) \((1/Pa)\)

\(\rho_{\text{ice}}\) Density of ice \((kg/m^3)\)

\(A_{R_p}\) Constant describing \(R_p\) in function of \(l\) \((1/s)\)

\(B_{R_p}\) Constant describing \(R_p\) in function of \(l\) \((1/m)\)

\(CV\) Coefficient of variation (-)

\(k\) Ratio of specific heats (-)

\(K_v\) Vial heat transfer coefficient \((J/(m^2sK))\)

\(L_{\text{dried}}\) Dried layer thickness \((m)\)

\(M\) Molecular weight \((kg/mol)\)

\(N\) Number of samples (-)

\(n_{\text{par}}\) Number of parameters
$P_c$ Chamber pressure (Pa)

$R_Y^2$ Model coefficient of determination (-)

$R_p$ Dried product mass transfer resistance (m/s)

$r_d$ Equivalent radius corresponding to duct opening between chamber and condenser (m)

$R_{p,0}$ Constant describing $R_p$ in function of $l$ (m/s)

$r_v,i$ Inner radius of the glass vial (m)

$r_v,n$ Equivalent radius corresponding to vial neck opening (m)

$r_v,o$ Outer radius of the glass vial (m)

$S_i$ First order effect (-)

$S_{T,i}$ Total order effect (-)

$T_c$ Collapse temperature (K)

$T_e$ Eutectic temperature (K)

$T_g'$ Glass transition temperature of the maximum freeze-concentrated formulation (K)

$T_i$ Product temperature at sublimation front (K)

$T_s$ Shelf fluid inlet temperature (K)

$T_{i,crit}$ Critical product temperature (K)

$V$ Filling volume (mL)

References


Primary drying model

\[ T_p, \dot{m}_{sub} = f(T_s, P_c, R_p, K_v, V \ldots) \]

Inherent uncertainty on model input parameters

Global Sensitivity Analysis

Process parameter sensitivity ranking

Primary drying

Uncertain model output