

Optimal Experimental Design for Estimating the Kinetic Parameters of Processes Described by the Weibull Probability Distribution Function

Luís M. Cunha^a, Fernanda A. R. Oliveira^{b*} & Jorge C. Oliveira^b

^aEscola Superior de Biotecnologia, Universidade Católica Portuguesa, Rua Dr. António Bernardino de Almeida, 4200 Porto, Portugal

^bInstituto Inter-Universitário de Macau, Palácio da Penha, Macau, Portugal

(Received 1 December 1997; accepted 26 May 1998)

ABSTRACT

The optimum experimental design for determining the kinetic parameters of the model resulting from the Weibull probability density function was studied, by defining the sampling conditions that lead to a minimum confidence region of the estimates, for a number of observations equal to the number of parameters. It was found that for one single isothermal experiment the optimum sampling times corresponded always to fractional concentrations that are irrational numbers (approximately 0.70 and 0.19) whose product is exactly $1/e^2$. The experimental determination of the equilibrium conversion (for growth kinetics) is very important, but in some situations this is not possible, e.g. due to product degradation over the length of time required. Sampling times leading to a maximum precision were determined as a function of the maximum conversion (or yield) attainable. For studies of kinetic parameters over a range of temperatures, performed with a minimum of three isothermal experiments, it was proved that the optimum design consists of two experiments at one limit temperature with two sampling times (those corresponding to fractional concentrations of approximately 0.70 and 0.19) and another at the other limit temperature for a sampling time such that the fractional concentration is $1/e$. Case studies are included for clarification of the concepts and procedures.

© 1998 Elsevier Science Limited. All rights reserved

NOMENCLATURE

C_i	Number or concentration of a component or quality factor at time t_i
C_{ini}	Initial concentration of a component or quality factor
C_∞	Equilibrium concentration of a component or quality factor
E_a	Activation energy (kJ mol^{-1})

*Author to whom correspondence should be addressed.

E.D.	Efficiency of the design (%)
L/D	Ratio between bed length and diameter
F	Matrix of the derivatives of the response function
n	Number of experimental points
p	Number of parameters
R	Universal gas constant ($\text{kJ mol}^{-1} \text{K}^{-1}$)
T	Temperature ($^{\circ}\text{C}$)
t_i	Sampling time of experiment i (min)
u_f	Superficial velocity (m/s)

Greek letters

α	Scale parameter (min)
α_0	Pre-exponential scale parameter (min)
β	Shape parameter
Δ	Modulus of the determinant of F (min^{-1} or $\text{min}^{-1} ^{\circ}\text{C}^{-1}$)
Δ_{\max}	Modulus of the determinant of F for optimal conditions (min^{-1} or $\text{min}^{-1} ^{\circ}\text{C}^{-1}$)
Δt	Time lag between t_1 and t_2 (min)
η_i	Fractional concentration for the experiment i
θ	Vector of p parameters

INTRODUCTION

The Weibull distribution is one of many probability distribution functions used to describe the behaviour of systems or events that have some degree of variability. It was originally developed in 1939 by W. Weibull to statistically analyse the ultimate strengths of materials (Simon & Woeste, 1980). The probability density function of the Weibull distribution may be described as (Hahn & Shapiro, 1967):

$$f(t) = \begin{cases} \frac{\beta}{\alpha} \left(\frac{t}{\alpha} \right)^{\beta-1} e^{-\left(\frac{t}{\alpha} \right)^{\beta}}, & t > 0 \\ 0, & \text{elsewhere} \end{cases} \quad (1)$$

with $\alpha > 0$ and $\beta > 0$. The corresponding cumulative distribution is:

$$F(t) = \int_0^t f(t) dt = 1 - e^{-\left(\frac{t}{\alpha} \right)^{\beta}} \quad (2)$$

This model has an interesting potential for describing microbial, enzymatic and chemical degradation kinetics (failure of the system after a given time subjected to stress conditions), considering the scale parameter (α) as a reaction rate constant and the shape parameter (β) as a behaviour index. The model reduces to a first order decay/growth kinetics for $\beta = 1$ (see Fig. 1). As discussed by Hahn and Shapiro (1967), Nelson (1969) and Gacula and Kubala (1975), the failure rate for the Weibull model is an increasing function of time for $\beta > 1$ and a decreasing function for $\beta < 1$. When $\beta = 1$, the failure rate is constant. These authors also

emphasise that α is a characteristic time to failure, as it corresponds to the $100 \cdot (1 - 1/e) = 63.2$ percentile of the distribution, regardless of the value of β (see Fig. 1(b)).

The Weibull distribution was successfully applied to describe shelf-life failure (Gacula & Kubala, 1975; Schmidt & Bouma, 1992). Page (1949) developed an equation to describe thin-layer drying of shelled corn, similar to the Weibull proba-

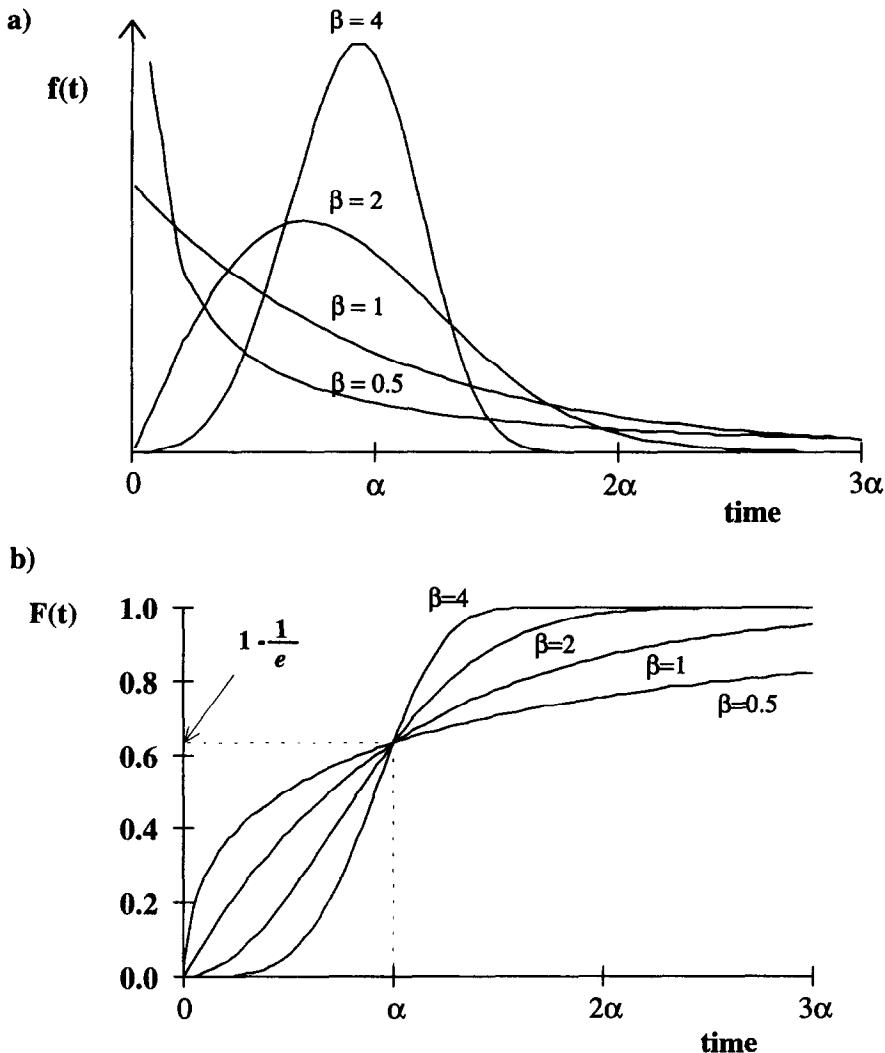


Fig. 1. Effect of the shape parameter β on (a) the Weibull probability density function, $f(t)$, (b) the Weibull distribution function, $F(t)$.

bilistic model that was later used to model thin-layer drying of yellow corn (Misra & Brooker, 1980; Li & Morey, 1984), of soybeans (White *et al.*, 1981), of pigeon pea (Shepherd & Bhardwaj, 1988) and of adzuki beans (Tagawa *et al.*, 1996). This model was also used to describe the drying kinetics of rough rice (Agrawal & Singh, 1977; Wang & Singh, 1978; Lu *et al.*, 1994). Chhinnan (1984) evaluated four mathematical models, i.e., the exponential model, the diffusion model, an approximation of the diffusion model and the Page (Weibull) model, for describing thin-layer drying of in-shell pecans, recommending the Page (Weibull) equation for future modelling studies. Moreover, the Weibull distribution has been used to model failure of apple tissue under cyclic loading (McLaughlin & Pitt, 1984), grape pomace elutriation (Peraza *et al.*, 1986) and water uptake and soluble solids losses during rehydration of dried apple pieces (Ilincanu *et al.*, 1995). Machado *et al.* (1997) adequately described moisture uptake and soluble solid losses by breakfast cereals immersed in water or milk using the Weibull model, after unsuccessful attempts to use the Peleg (1988), the diffusion and the first order models. Recently, the Weibull model has also been used to describe microbial death kinetics under high-pressure conditions (Heinz & Knorr, 1996).

As in any other model, the usefulness of the Weibull model for predictive or descriptive purposes greatly depends on the precision and accuracy of the model parameter estimates. Rigorous parameter estimation requires an adequate statistical analysis of experimental data, but also a careful experimental design. As stressed by Bates and Watts (1988), even very careful data analysis is unable to recover information that is not present in the experimental data.

Box and Lucas (1959) proposed an optimum design criterion for non-linear models that allows the selection of sampling conditions that lead to a minimum confidence region, also known as *D*-optimal design (Bates & Watts, 1988). The rationale is that as the 'real' parameters of a system are probably contained in the error region of the estimates, the smaller this region the higher the precision. This criterion has been applied in different situations, such as, first order kinetics under isothermal conditions (Box & Lucas, 1959), diffusional processes under non-isothermal conditions (Oliveira *et al.*, 1995) and the Thermal Death Time model under both isothermal and non-isothermal conditions (Cunha *et al.*, 1997).

The main objective of this work was to define the best experimental conditions according to the *D*-optimal design for systems that may be described by the Weibull probabilistic model.

MATHEMATICAL METHODS

For any choice of the design variable (i.e., the independent variable, t) the size of the parameters joint confidence region is proportional to the Jacobian $|F^T F|^{-1/2}$ of the derivative matrix F (where $F \equiv [f_{i,j}]$, with $f_{i,j} = \partial x_i / \partial \theta_j$ evaluated at $t = t_i$ with i ranging from 1 to n ; x represents the system response and θ is a kinetic parameter). Thus, a logical choice of the design criterion is to choose sampling points so that the size of this joint confidence region is minimised, that is, the determinant $D \equiv |F^T F|$ should be maximised. This can be done for a number of sampling times equal to the number of parameters to be estimated, which is the minimum possible. Atkinson and Hunter (1968) showed that for a number of observations greater than the number of parameters, the optimal design often corresponds to r replications of the

optimal p sampling times ($r = n/p$). According to Box and Lucas (1959), if a sequence of n observations is to be designed for a p -parameter model, in the case where $n = p$, the D -optimal design can be simplified from the maximisation of $D \equiv |F^T F|$ to the maximisation of $\Delta \equiv \text{mod}(|F|)$ (Δ denotes the modulus of the determinant of the matrix F). Due to the complexity of the mathematical expressions of Δ derived for the Weibull model, there was no analytical solution, and therefore a numerical procedure was used. Numerical optimisation was performed using *Mathematica*® for Windows—2.2 Enhanced Version (Wolfram, 1993).

If one considers that: (i) η_i corresponds to $(C_i - C_\infty)/(C_{\text{ini}} - C_\infty)$, the fractional amount of a given component C , changing from an initial value (C_{ini}) to a final equilibrium value (C_∞), at a time t , and that (ii) the time required to reach a certain value of η is represented by the continuous random variable τ , with probability density function $f(t)$, where $f(t)$ is the Weibull distribution function, then $\eta(t)$ can be defined as the probability of having a certain fractional amount of C for, at least, a specified time t , under specified experimental conditions. Therefore:

$$\eta(t) = P(\tau > t) = \int_0^\infty f(t) dt = 1 - F(t) \quad (3)$$

Experiments conducted at a single temperature

Combining eqns (2) and (3), the response function estimated at time t_i , can be expressed as:

$$C_i = C_\infty + (C_{\text{ini}} - C_\infty)\eta_i = C_\infty + (C_{\text{ini}} - C_\infty)e^{-\left(\frac{t_i}{\alpha}\right)^\beta} \quad (4)$$

Following the procedure presented by Box and Lucas (1959), the determinant, Δ , of the matrix of derivatives of the response function in order to the model parameters, was built for four time levels t_0 , t_1 , t_2 and t_3 . For the sake of simplification, Δ is represented in terms of η :

$$\Delta = \text{mod} \begin{pmatrix} \frac{\partial C_0}{\partial \alpha} & \frac{\partial C_0}{\partial \beta} & \frac{\partial C_0}{\partial C_{\text{ini}}} & \frac{\partial C_0}{\partial C_\infty} \\ \frac{\partial C_1}{\partial \alpha} & \frac{\partial C_1}{\partial \beta} & \frac{\partial C_1}{\partial C_{\text{ini}}} & \frac{\partial C_1}{\partial C_\infty} \\ \frac{\partial C_2}{\partial \alpha} & \frac{\partial C_2}{\partial \beta} & \frac{\partial C_2}{\partial C_{\text{ini}}} & \frac{\partial C_2}{\partial C_\infty} \\ \frac{\partial C_3}{\partial \alpha} & \frac{\partial C_3}{\partial \beta} & \frac{\partial C_3}{\partial C_{\text{ini}}} & \frac{\partial C_3}{\partial C_\infty} \end{pmatrix}$$

$$= \text{mod} \left(\frac{(C_{\text{ini}} - C_\infty)^2}{\alpha} \left\{ \eta_0 \ln(\eta_0) \eta_1 \ln(\eta_1) \ln \left[\frac{\ln(\eta_0)}{\ln(\eta_1)} \right] (\eta_2 - \eta_3) + \eta_0 \ln(\eta_0) \eta_2 \ln(\eta_2) \right. \right.$$

$$\begin{aligned}
& \times \ln \left[\frac{\ln(\eta_2)}{\ln(\eta_0)} \right] (\eta_1 - \eta_3) + \eta_0 \ln(\eta_0) \eta_3 \ln(\eta_3) \ln \left[\frac{\ln(\eta_0)}{\ln(\eta_3)} \right] (\eta_1 - \eta_2) + \eta_1 \ln(\eta_1) \eta_2 \ln(\eta_2) \\
& \times \ln \left[\frac{\ln(\eta_1)}{\ln(\eta_2)} \right] (\eta_0 - \eta_3) + \eta_1 \ln(\eta_1) \eta_3 \ln(\eta_3) \ln \left[\frac{\ln(\eta_3)}{\ln(\eta_1)} \right] (\eta_0 - \eta_2) \\
& + \eta_2 \ln(\eta_2) \eta_3 \ln(\eta_3) \ln \left[\frac{\ln(\eta_2)}{\ln(\eta_3)} \right] (\eta_0 - \eta_1) \Bigg\} \quad (5)
\end{aligned}$$

As stressed by Cunha *et al.* (1997), when the initial concentration, C_{ini} , is considered to be a model parameter one of the optimal sampling times would be equal to zero (the time at which the initial concentration is determined), leading to $\eta_0 = 1$. Thus:

$$\begin{aligned}
\Delta = \text{mod} \left[\frac{(C_{ini} - C_\infty)^2}{\alpha} \left\{ \eta_1 \ln(\eta_1) \eta_2 \ln(\eta_2) \ln \left[\frac{\ln(\eta_1)}{\ln(\eta_2)} \right] (1 - \eta_3) - \eta_1 \ln(\eta_1) \eta_3 \right. \right. \\
\left. \left. \times \ln(\eta_3) \ln \left[\frac{\ln(\eta_1)}{\ln(\eta_3)} \right] (1 - \eta_2) + \eta_2 \ln(\eta_2) \eta_3 \ln(\eta_3) \ln \left[\frac{\ln(\eta_2)}{\ln(\eta_3)} \right] (1 - \eta_1) \right\} \right] \quad (6)
\end{aligned}$$

Similarly, when the equilibrium value is considered to be a model parameter, an optimal sampling time near infinity would be obtained, that is, $\eta_3 = 0$. Combining both results, the determinant Δ may be simplified as follows:

$$\Delta = \text{mod} \left(\frac{(C_{ini} - C_\infty)^2}{\alpha} \eta_1 \ln(\eta_1) \eta_2 \ln(\eta_2) \ln \left[\frac{\ln(\eta_2)}{\ln(\eta_1)} \right] \right) \quad (7)$$

The optimal experimental design leading to a minimum confidence region was found by determining numerically the conversion values η_1 and η_2 that maximise the modulus of Δ . Substituting these values back into eqn (4), the optimal sampling times, t_1 and t_2 , were obtained.

There are some practical cases where it is impossible or unfeasible to reach the equilibrium value because the time required might be too long and/or the food may degrade. In these cases sub-optimal designs should be developed, taking into consideration the maximum measurable value of ' C_3 ', that is, the minimum value of ' η_3 '. Having this in mind, the determinant Δ of the matrix of derivatives of the response function in order to the model parameters was built for three time levels t_1 , t_2 and t_3 . For the sake of simplification, Δ was again represented in terms of a system response η , the fraction by which the yield ' C ' falls short of its equilibrium value.

Therefore, the values of η_1 and η_2 that maximise the determinant in eqn (6) were also numerically obtained for different fixed values of η_3 , thus obtaining sub-optimal designs.

To have a better insight into the loss of precision in the parameter estimates due to the deviation of η_3 from its optimum value (zero: $C_i = C_\infty$), relative design efficiencies (E.D.) were computed for each value of η_3 , as defined by Neter *et al.* (1966):

$$\text{E.D.}_i = \left(\frac{|(F^T F)^{-1}|_{\min}}{|(F^T F)^{-1}|_i} \right)^{1/p} * 100 = \left[\left(\frac{\Delta_i}{\Delta_{\max}} \right)^2 \right]^{1/p} * 100 \quad (8)$$

Set of isothermal experiments conducted over a range of temperatures

As a rate parameter, α is temperature sensitive and $1/\alpha$ can be expected to follow an Arrhenius-type behaviour, as often found (Shepherd & Bhardwaj, 1988; Ilincau *et al.*, 1995; Machado *et al.*, 1997):

$$\frac{1}{\alpha_i} = \frac{1}{\alpha_0} e^{\left[-\frac{Ea}{(RT_i)} \right]} \quad (9)$$

The β parameter should indicate the kinetic pattern and thus be temperature independent, particularly within a limited range of temperatures. This result has been verified in published works (Chhinnan, 1984; Ilincau *et al.*, 1995; Machado *et al.*, 1997). Therefore, the kinetic parameters of the model for practical situations in food processing are α_0 , Ea and β , as non-isothermal conditions are the most usual. These parameters can however be determined by performing isothermal experiments at different temperatures covering the range of interest. For this case, the model is:

$$\eta_i = c \left\{ -\left(\frac{t_i}{\alpha_0} e^{\left[-\frac{Ea}{(RT_i)} \right]} \right)^\beta \right\}, i = 1, 2, 3 \quad (10)$$

In this situation, optimal designs require not only the selection of sampling times but also the selection of the temperatures that lead to a higher precision of the estimates. If isothermal experiments are performed, then one should select the most appropriate temperatures for the experiments and, for each temperature, the adequate sampling times. Non-isothermal experiments require the definition of both the temperature history and the sampling times but this case will not be analysed in this work.

The sampling times for C_{ini} and C_∞ , in case they are taken as model parameters, were already approached ($t = 0$ and $t \rightarrow \infty$, respectively). It is eventually necessary to analyse whether C_∞ is temperature independent or not. The subsequent analysis refers only to the determination of the three kinetic parameters and therefore C_{ini} and C_∞ are taken as known constants in the mathematical handling. Following the previous procedure, the determinant, Δ , of the matrix of derivatives of the response function in order to the model parameters, α_0 , Ea and β , was determined for three time levels t_1 , t_2 and t_3 and three temperature levels T_1 , T_2 and T_3 , respectively, yielding:

$$\Delta = \text{mod} \left(\frac{\beta(C_{ini} - C_\infty)^3 \eta_1 \ln(\eta_1) \eta_2 \ln(\eta_2) \eta_3 \ln(\eta_3)}{\alpha_0 R T_1 T_2 T_3} \times \left[T_1 T_2 \ln \left(\frac{\ln(\eta_2)}{\ln(\eta_1)} \right) + T_2 T_3 \ln \left(\frac{\ln(\eta_3)}{\ln(\eta_2)} \right) + T_1 T_3 \ln \left(\frac{\ln(\eta_1)}{\ln(\eta_3)} \right) \right] \right) \quad (11)$$

The optimal experimental design leading to a minimum confidence region for any combination of temperatures (T_1 , T_2 and T_3) was found by determining numerically the conversion values η_1 , η_2 and η_3 that maximise the modulus of Δ . Substituting these values back into eqn (10), the corresponding optimal sampling times: t_1 , t_2 and t_3 , were obtained.

RESULTS

Experiments conducted at a single temperature

As discussed in the 'Mathematical methods' section, the complexity of eqn (7) prevented an analytical solution and the optimal design for the Weibull probabilistic model was numerically computed for a large number of different sets of the model parameters α (between 0.1 and 600 min) and β (between 0.1 and 5). It was found that the fractional concentrations at the sampling times that maximised Δ were independent of the model parameters. In all cases, the solution of the optimisation problem was a pair of irrational numbers, identical to those obtained by Cunha *et al.* (1998) for the optimal experimental design for systems following the Thermal Death Time model under nonisothermal conditions: $\eta_1 = 0.70322\dots$ and $\eta_2 = 0.19245\dots$, which are such that their product is equal to $1/e^2$. Figure 2 shows how the determinant Δ from eqn (7) typically evolves as a function of the fractional concentration values, with its maximum occurring at the referred values of η_1 and η_2 . The two sampling points are symmetrical in a logarithmic scale in relation to $1/e$ (see Fig. 3), this point corresponding to the optimal sampling point obtained by Box and Lucas (1959) for the simple exponential model (equal to the Weibull model for β equal to 1). It was further verified that if β is a known constant and α is the only parameter to be estimated, the optimal sampling time would also correspond to $\eta = 1/e$, independently of the value of β . It is noteworthy that the time lag between the two optimal sampling times decreases as the value of the shape parameter β increases (see Fig. 3). Figure 4 shows the time lag between the optimal sampling times of the Weibull probabilistic model as a function of the model parameters α and β . Each curve drawn indicates all pairs of α and β that have the same time lag. Published

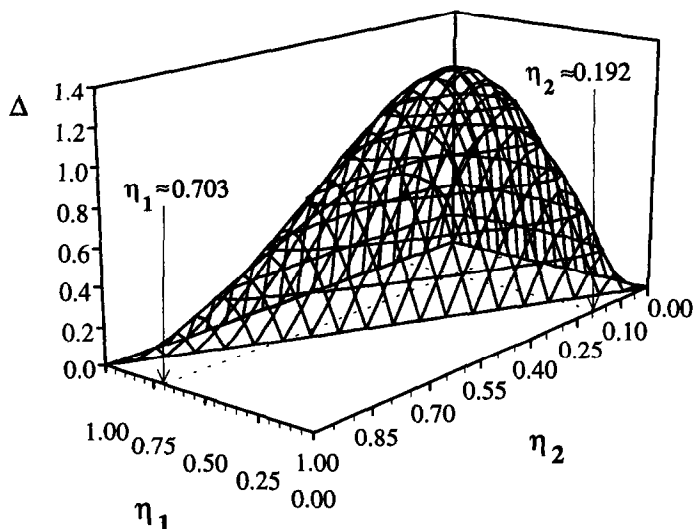


Fig. 2. Surface plot of Δ as a function of η_1 and η_2 (eqn (7)). Maximum value of Δ is attained for $\eta_1 \cong 0.703$ and $\eta_2 \cong 0.192$.

values of α and β are also included. In general, the kinetics reported on literature show a value of β lower than 1 and a value of α such that the time lag is, in most cases, larger than 20 min. However, for Heinz and Knorr (1996) data and for some of the data from Machado *et al.* (1997), the β is so high and/or α is so small, that this period either becomes unpractical or may lead to a considerable experimental error in the independent variable, t , that is not accounted for in the present analysis (see Fig. 4). A case study will be presented in the next section to illustrate how to handle this problem.

In growth kinetics, the equilibrium value (C_∞) is a very important parameter and should be evaluated at a sampling time near infinity, that is, at $\eta = 0$. However, sometimes the maximum attainable value of C is smaller than C_∞ . In these situations η_3 is fixed by practical constraints and new optimal η_1 and η_2 values need to be calculated. It was found that as η_3 increases, the optimal values of η_1 and η_2 also increase and get closer (see Fig. 5(a)). The design efficiency (as computed from eqn (8)) strongly decreases when η_3 increases (see Fig. 5(b)). It is also possible to translate this loss of efficiency in terms of the number of experimental data required to have the same precision of the estimates (i.e., same Δ). It is noteworthy that if,

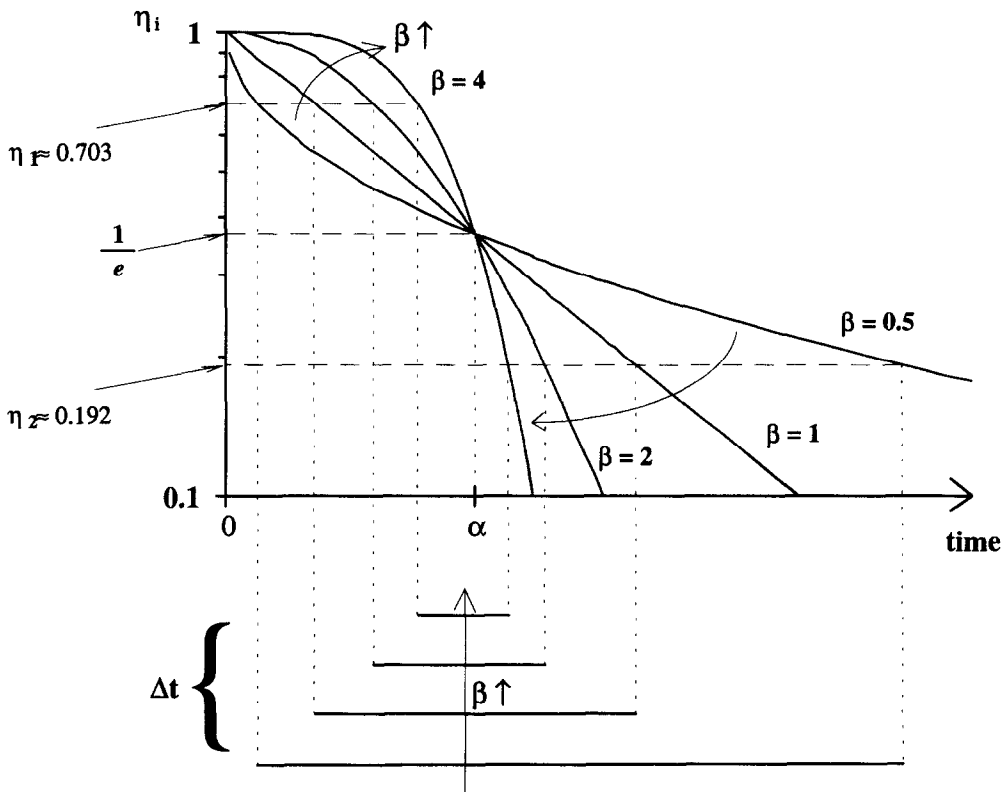


Fig. 3. Identification of the optimal sampling points for the Weibull probabilistic model, evaluated at different values of the shape parameter β .

for example, one can only follow up a certain process until 50% of the total conversion ($\eta_3 = 0.5$), design efficiency will drop so strongly that in order to have the same precision the number of sampling points required would be approximately 18 times larger, when compared to the full optimal design (see small graph in Fig. 5(b)). Thus, even for relatively low values of η_3 , the application of optimal designs may not be enough to prevent significant loss of precision in the estimation of the model parameters.

Set of isothermal experiments conducted over a range of temperatures

The results of Box and Lucas (1959) for the first-order reaction model and those of Cunha *et al.* (1997) for the isothermal Thermal Death Time model, showed that when analysing a given system over a range of temperatures, an experiment should

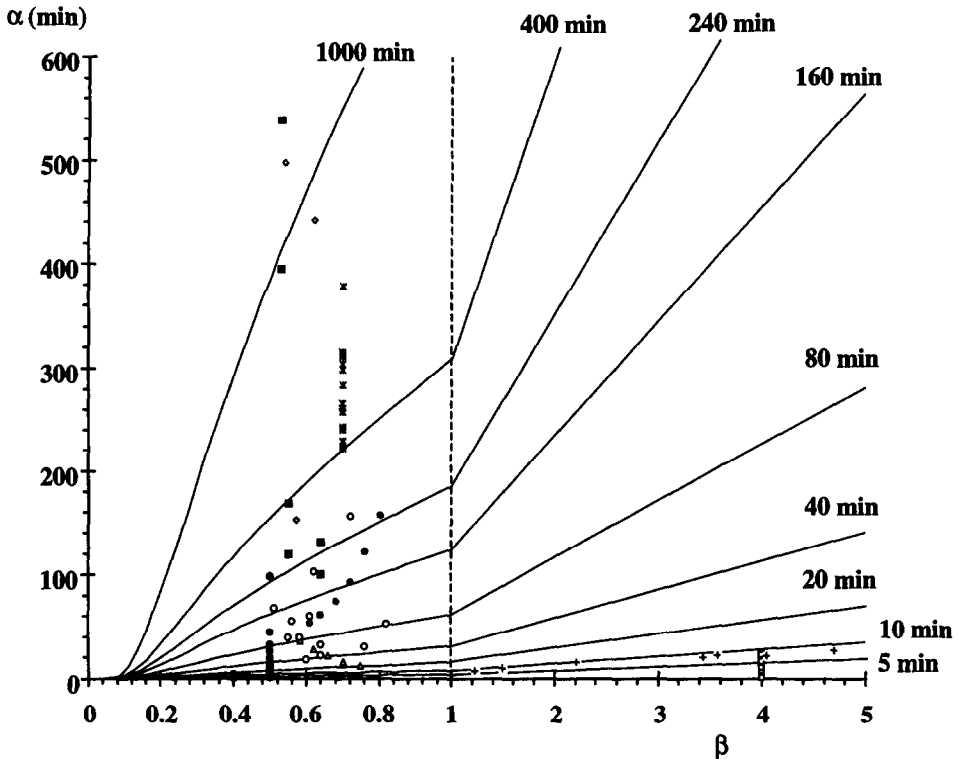


Fig. 4. Time lag between the optimal sampling times of the Weibull probabilistic model as a function of the model parameters α and β . Each curve drawn indicates all pairs of α and β that have the same time lag (iso-time lag curves). Published values of α and β are indicated according to the following legend: *, Chhinnan (1984); \square , Heinz and Knorr (1996); \circ , Ilincanu *et al.* (1995); Δ , Li and Morey (1984); \bullet , Machado *et al.* (1997); \diamond , Misra and Brooker (1980); +, Peraza *et al.* (1986); \bullet , Shepherd and Bhardwaj (1988); and \blacksquare , White *et al.* (1981).

be performed at each of the limit temperatures in order to obtain the maximum precision of the parameters of the model. Under those circumstances, sampling times corresponding to a conversion value equal to $1/e$, should be taken at each limit temperature. By analogy, and taking into consideration that the Weibull model

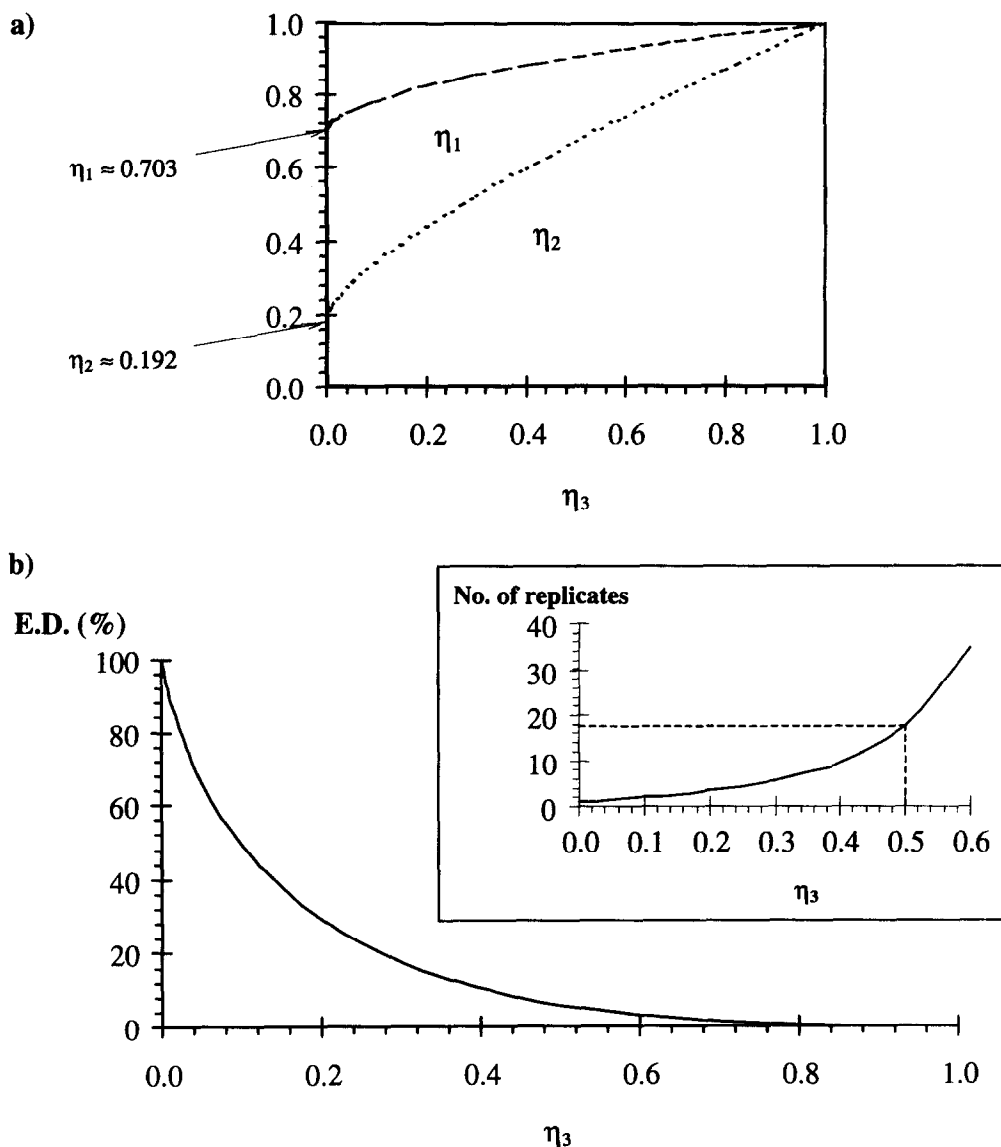


Fig. 5. *D*-optimal design for the Weibull probabilistic model when the maximum conversion value (η_3) falls short of equilibrium ($\eta_3 = 0$): (a) optimal sampling values; (b) efficiency of the sub-optimal designs. The number of replicates required to have the same precision (small plot) is computed as the reciprocal of the efficiency index.

under these conditions has three parameters, it would be expected that there should be an experiment at each of the limit temperatures, with a third data point being taken at any given temperature between T_{\min} and T_{\max} . In order to analyse this situation, Δ (from eqn (11)) was maximised considering $T_1 = T_{\min}$, $T_3 = T_{\max}$ and covering different values of T_2 , ranging from T_{\min} to T_{\max} . Furthermore, the effect of changing T_1 from T_{\min} to T_{\max} was also assessed. Figure 6 shows how the design efficiency typically evolves with sampling temperature using $T_3 = T_{\max}$ and it can be clearly seen that the best design is $T_1 = T_2 = T_{\min}$ with $T_3 = T_{\max}$, $T_1 = T_{\min}$ with $T_2 = T_3 = T_{\max}$ or $T_2 = T_{\min}$ with $T_1 = T_3 = T_{\max}$. Lower efficiency values were obtained whenever $T_{\min} < T_3 < T_{\max}$, whatever the values of T_1 and T_2 . Maximisation of Δ in order to η_1 , η_2 and η_3 , for $T_1 = T_2$, yielded the following optimal design: $\eta_1 = 0.70322\dots$, $\eta_2 = 0.19245\dots$ and $\eta_3 = 1/e$. Thus, when considering a range of temperatures, two sampling points should be taken at an extreme temperature, corresponding to fractional concentrations of $0.70322\dots$ and $0.19245\dots$. The remaining sample should be taken at the other extreme temperature at the time when the fractional concentration of $1/e$ is expected.

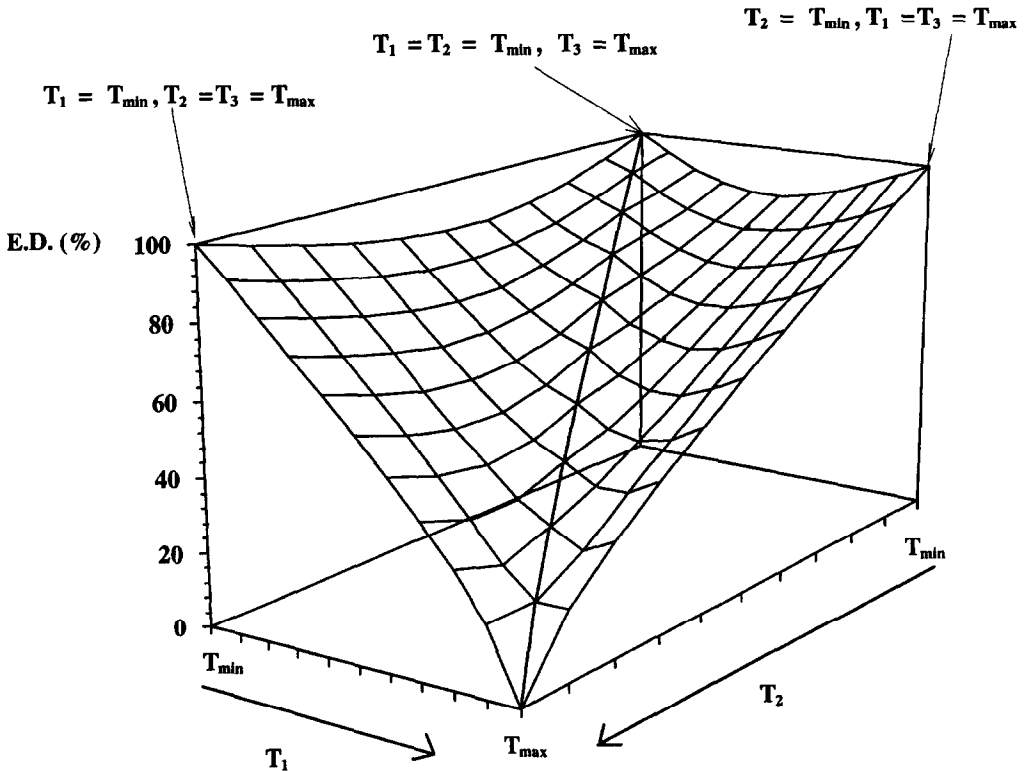


Fig. 6. Efficiency of the D -optimal design for the Weibull probabilistic model with an Arrhenius temperature dependency of the scale parameter α , as a function of the sampling temperatures T_1 and T_2 for $T_3 = T_{\max}$.

CASE STUDIES

To clarify the application of some of these concepts in food research, some case studies are presented using data from literature.

Separation of grape pomace in a fluidised bed

To model the simultaneous effect of drying and separation of grape pomace in a fluidised bed, Peraza *et al.* (1986) have used the Weibull model, as mentioned in the introduction. In this work, they have measured as a response function the grape skin fraction in the bed at any time (η) and modelled grape pomace elutriation as a function of fluidisation velocity (u_f) and length/diameter ratio (L/D) of the fluidised bed. For the case where $u_f = 5.0$ m/s and $L/D = 4$, the Weibull model yielded the following parameters: $\alpha = 26.275$ min and $\beta = 4.701$. If one would intend to perform the same study using an optimal design, samples should be taken at $t_1 \approx 21.0$ min (time at which $\eta \approx 0.703$) and at $t_2 \approx 29.2$ min (time at which $\eta \approx 0.192$), the time lag between t_1 and t_2 (Δt) being therefore 8.2 min. From an experimental point of view, these sampling times are perfectly feasible. The heuristic experimental design used by Peraza *et al.* (1986), was based on six sampling points equally spaced in time, with a time lag of 5 min ($t = 5, 10, 15, 20, 25$ and 30 min); comparing this design with the corresponding optimal sampling procedure (three replicates taken at $t = 21.0$ min and three at $t = 29.2$ min), its efficiency (eqn (8)), would be approximately 52%. This means that the heuristic design used by Peraza *et al.* (1986) would require twice as many points as the optimum design to have roughly the same precision of the parameters estimated.

Soluble solids losses during rehydration

Considering a different situation, Machado *et al.* (1997) have modelled the soluble solid losses by puffed breakfast cereals immersed in water at 55°C, where the Weibull model yielded the following parameters: $\alpha = 1.04$ min and $\beta = 0.4$. To perform the same study using an optimal design, according to those values, samples should be taken at $t_1 \approx 0.08$ min and at $t_2 \approx 3.62$ min, with a resulting time lag of 3.54 min. While the time lag between sampling times appears to be experimentally manageable, the first sampling time may be unfeasible. If this requirement becomes an experimental hurdle, the optimal design cannot be applied. However, sub-optimal designs may be defined, for a pre-selected first sampling time (t_1), if the determinant in eqn (7) is maximised as a function of t_2 only (and thus as a function of the time lag between samples). Figure 7 shows the results obtained for values of t_1 up to 5 min, together with the respective design efficiency (eqn (8)). If, for example, one decides to establish a first sampling time equal to 2 min, the second sampling point should be taken at approximately 10.4 min ($\Delta t \approx 8.4$ min; see Fig. 7), with corresponding fractional concentration values of 0.27 and 0.08, respectively. The resulting efficiency would however be around 40% (see Fig. 7).

Microbial death kinetics under high pressure

Heinz and Knorr (1996) used the Weibull model to describe microbial death kinetics under high pressure conditions. Taking the example at a temperature of

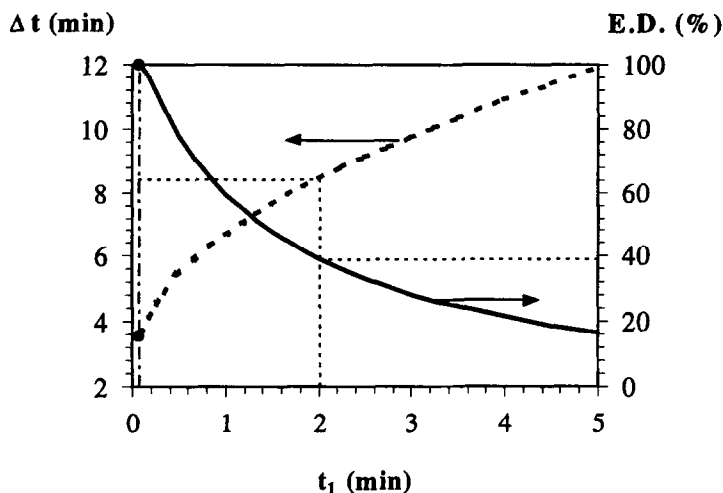


Fig. 7. Sub-optimal designs when fixing the values of t_1 , for Machado *et al.* (1997) parameters concerning the loss of soluble solids by puffed breakfast cereals immersed in water at 55°C. The plot gives the Δt ($t_2 - t_1$) values for any t_1 in the left axis and the corresponding efficiencies of the design in the right axis. (●) Optimal sampling conditions: $t_1 = 0.08$ min; $\Delta t = 3.55$ min and E.D. = 100%.

40°C and a pressure of 250 MPa the parameters reported were: $\alpha = 4.819$ min and $\beta = 4$. To perform experiments using an optimal design, samples should be taken at $t_1 \approx 3.71$ min and at $t_2 \approx 5.46$ min, with a resulting Δt of 1.75 min. This time lag is relatively small and thus a larger lag may be preferred. Sub-optimal conditions were studied for this situation, constraining the value of the time lag and maximising the determinant in eqn (7) in order to one of the sampling times only, for values of Δt up to 5 min. Figure 8 shows the results obtained. If, for example, one decides to select a time lag of 2.5 min, the first sample should be taken at ≈ 3.26 min (therefore, $t_2 = 5.76$ min), and the efficiency of this sub-optimal design would be around 85% (see Fig. 8).

For many experimental studies, the initial concentration (C_{ini}) is not necessarily measured at $t = 0$, as the factor under study shows high stability over a given period of time. However, there are systems and/or variables for which C_{ini} must be measured exactly at $t = 0$, as the factor will change over a short period of time. Under these circumstances, care should be taken when the minimum Δt is too high, as the first sampling time might be lower than Δt and t_1 is a time lag to zero. This is shown by the shaded part of the graph.

CONCLUSIONS

The optimal experimental design for determining the Weibull model parameters with the minimum possible error region for a number of points equal to the number of parameters at a constant temperature T consists of taking a sample at the time such that the fractional concentration is $\eta_1 = 0.70322\dots$ and another when

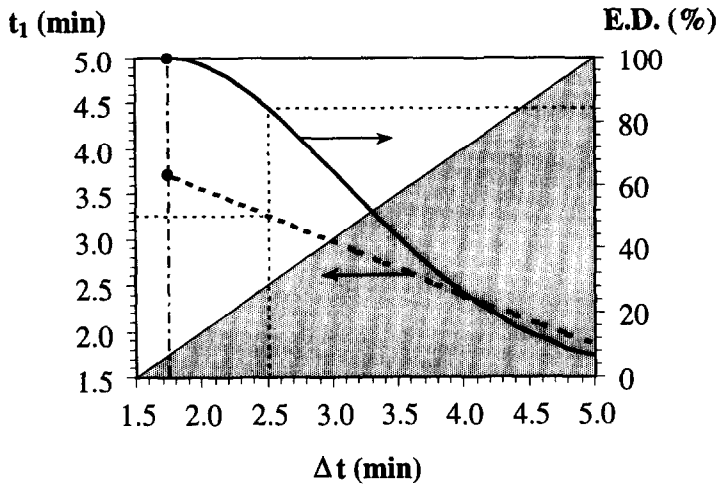


Fig. 8. Sub-optimal designs when fixing time lag between the two sampling points, for Heinz and Knorr (1996) parameters concerning the death kinetics of *Bacillus subtilis* at 40°C and 250 MPa. The plot gives the t_1 -value for any Δt ($t_2 = t_1 + \Delta t$) in the left axis and the corresponding efficiencies of the design in the right axis. The shaded area indicates designs that may not be logical, as t_1 is lower than Δt . (●) Optimal sampling conditions: $t_1 = 3.71$ min; $\Delta t = 1.75$ min and E.D. = 100%.

$\eta_2 = 0.19245\dots$, with $\eta_1^* \eta_2 = 1/e^2$. Assuming that $1/\alpha$ follows an Arrhenius-type temperature dependency and β is independent of temperature, the optimal experimental design to determine the system parameters for temperature dependent situations with isothermal experiments consists of two experiments, performed at the limit temperatures of the range of interest. In one, two samples should be taken, one when the conversion is $\eta_1 = 0.70322\dots$ and another when $\eta_2 = 0.19245\dots$, while in the other experiment one sample should be taken where $\eta_3 = 1/e$.

The initial concentration and the equilibrium concentration (if different from zero) should be determined by analysing samples, respectively, at time zero and at a significantly large time so that η is constant ($C_i \rightarrow C_\infty$).

ACKNOWLEDGEMENTS

The authors acknowledge financial support from the EU projects AIR1-CT92-0746 and STD3-CT94-0333. The first author acknowledges financial support from Fundação para a Ciência e a Tecnologia (FCT) through 'Sub-programa Ciência e Tecnologia do 2º Quadro Comunitário de Apoio'. We also thank Andreia Pinheiro Torres for her helpful advice.

REFERENCES

- Agrawal, Y. C. & Singh, R. P. (1977). Thin-layer drying studies on short-grain rough rice. ASAE Paper No. 77-3531, ASAE, St Joseph, MI.

- Atkinson, A. C. & Hunter, W. G. (1968). The design of experiments for parameter estimation. *Technometrics*, **10**, 271–289.
- Bates, D. M. & Watts, D. G. (1988). *Nonlinear Regression Analysis and Its Applications*, pp. 121–127. Wiley, New York.
- Box, G. E. P. & Lucas, H. L. (1959). Design of experiments for non-linear situations. *Biometrika*, **46**, 77–90.
- Chhinnan, M. S. (1984). Evaluation of selected mathematical models for describing thin-layer drying of in-shell pecans. *Transactions of the ASAE*, **27**, 610–615.
- Cunha, L. M., Oliveira, F. A. R., Brandão, T. R. S. & Oliveira, J. C. (1997). Optimal experimental design for estimating the kinetic parameters of the Bigelow Model. *Journal of Food Engineering*, **33**, 111–128.
- Gacula, M. C. Jr & Kubala, J. J. (1975). Statistical models for shelf life failures. *Journal of Food Science*, **40**, 404–409.
- Hahn, G. J. & Shapiro, S. S. (1967). *Statistical Models in Engineering*. Wiley, New York.
- Heinz, V. & Knorr, D. (1996). High pressure inactivation kinetics of *Bacillus subtilis* cells by a three-state-model considering distributed resistance mechanisms. *Food Biotechnology*, **10**, 149–161.
- Ilincanu, L. A., Oliveira, F. A. R., Drumond, M. C., Machado, M. F. & Gekas, V. (1995). Modelling moisture uptake and soluble solids losses during rehydration of dried apple pieces. Proceedings of the Copernicus First Main Meeting—*Process Optimisation and Minimal Processing of Foods*, December 1995, ESB, Porto, Portugal.
- Li, H. & Morey, R. V. (1984). Thin-layer drying of yellow dent corn. *Transactions of the ASAE*, **27**, 581–585.
- Lu, R., Siebenmorgen, T. J. & Archer, T. R. (1994). Absorption of water in long-grain rough rice soaking. *Journal of Food Process Engineering*, **17**, 141–154.
- Machado, M. F., Oliveira, F. A. R. & Gekas, V. (1997). Modelling water uptake and soluble solids losses by puffed breakfast cereal immersed in water or milk. *Proceedings of the Seventh International Congress on Engineering and Food*, Brighton, UK.
- McLaughlin, N. B. & Pitt, R. E. (1984). Failure characteristics of apple tissue under cyclic loading. *Transactions of the ASAE*, **27**, 311–320.
- Misra, M. K. & Brooker, D. B. (1980). Thin-layer drying and rewetting equations for shelled yellow corn. *Transactions of the ASAE*, **23**, 1254–1260.
- Nelson, W. (1969). Hazard plotting for incomplete failure data. *Journal of Quality Technology*, **1**, 27–52.
- Neter, J., Kutner, M. H., Nachtsheim, C. J. & Wasserman, W. (1966). *Applied Linear Statistical Models*, 4th ed., pp. 1289–1293. Irwin, Chicago.
- Oliveira, F. A. R., Silva, T. R. & Oliveira, J. C. (1995). Optimal experimental design for estimation of mass diffusion parameters using non-isothermal conditions. *Proceedings of the First International Symposium on Mathematical Modelling and Simulation in Agriculture and Bio-Industries of IMACS/IFAC*, 9–12 May, Brussels, Belgium.
- Page, G. E. (1949). Factors influencing the maximum rates of air drying shelled corn in thin layers. M.Sc. thesis, Purdue University, Lafayette, IN.
- Peleg, M. (1988). An empirical model for the description of moisture sorption curves. *Journal of Food Science*, **53**, 1216–1219.
- Peraza, A. L., Penã, J. G., Segurajauregui, J. S. & Vizcarra, M. (1986). Dehydration and separation of grape pomace in a fluidized bed system. *Journal of Food Science*, **51**, 206–210.
- Schmidt, K. & Bouma, J. (1992). Estimating shelf-life of cottage cheese using hazard analysis. *Journal of Dairy Science*, **75**, 2922–2927.
- Shepherd, H. & Bhardwaj, R. K. (1988). Thin layer drying of pigeon pea. *Journal of Food Science*, **53**, 1813–1817.
- Simon, J. C. & Woeste, F. E. (1980). A system-independent numerical method to obtain maximum likelihood estimates of the three parameter Weibull distribution. *Transactions of*

the ASAE, **23**, 955–958, 963.

Tagawa, A., Kitamura, Y. & Murata, S. (1996). Thin layer drying characteristics of adzuki beans. *Transactions of the ASAE*, **39**, 605–609.

Wang, C. Y. & Singh, R. P. (1978). A single layer drying equation for rough rice. ASAE Paper No. 78-3001. ASAE, St. Joseph, MI.

White, G. M., Bridges, T. C., Loewer, O. J. & Ross, I. J. (1981). Thin-layer drying model for soybeans. *Transactions of the ASAE*, **24**, 1643–1646.

Wolfram, S. (1993). *Mathematica*®. *A System for Doing Mathematics by Computer*, 2nd ed., pp. 703–704, 794. Addison-Wesley Publishing Company, Reading, MA.