On the use of Artificial Neural Networks to monitor a pharmaceutical freeze-drying process

This is an Accepted Manuscript of an article published by Taylor & Francis Group in *Drying Technology* on 10/01/2013 (Volume 31, Issue 1, pages 72-81, 2013), available online: http://www.tandfonline.com/doi/abs/10.1080/07373937.2012.718308.

On the use of Artificial Neural Networks to monitor a pharmaceutical freeze-drying process

Elena Niculina Drăgoi[1], Silvia Curteanu[1], Davide Fissore[2]

[1]*Department of Chemical Engineering,*
"Gheorghe Asachi" Technical University of Iasi, Str. Prof.dr.doc. Dimitrie Mageron No. 73, 700050 Iasi, Romania

[2]*Dipartimento di Scienza Applicata e Tecnologia,*
Politecnico di Torino, corso Duca degli Abruzzi 24, 10129 Torino (Italy)

Abstract

This paper is focused on the use of artificial neural networks to monitor a pharmaceutical freeze-drying process.

A detailed phenomenological model of the process is used to provide the dataset for the learning phase of the neural network. Then, a methodology based on a self-adaptive differential evolution scheme is combined with a back-propagation algorithm, as local search method, for the simultaneous structural and parametric optimization of the neural network which models the freeze-drying process.

Using some experimentally available measurements at a generic time t, the neural network is able to estimate the temperature of the product and the thickness of the dried cake (the amount of residual ice, as well as the sublimation flux can be easily calculated from the cake thickness) at a future time $t + \Delta t$, for the given operating conditions (the temperature of the heating shelf and the pressure in the drying chamber). Also, the duration of the primary drying phase and the maximum product temperature in the future are predicted, in case the operating conditions are not modified. By this way it is possible to understand if it is necessary to modify the operating conditions, if the product temperature should trespass the limit value before the ending point of the primary drying.

Despite the artificial neural network is obtained using a learning set determined for specific values of heat transfer coefficient (between the heating shelf and the product at the bottom of the container) and of mass transfer resistance (of the dried cake to vapor flow), reliable and accurate estimations are obtained also in case the sensor is used to monitor a process characterized by different values of heat and mass transfer coefficients.

Key words

Freeze-drying; Mathematical model; Monitoring; Soft-sensor; Artificial neural network; Differential evolution

Introduction

In a pharmaceutical freeze-drying process it is required to monitor the dynamics of the product, i.e. the evolution of the temperature and the residual amount of ice in the product, both in the phase of recipe design and during manufacturing. In fact, it is mandatory to avoid product damage due to overheating, in case the product temperature trespasses a limit value that is a characteristic of the formulation being freeze-dried. Besides, the operating conditions (the temperature of the heating shelf and the pressure in the drying chamber) used during primary drying (when ice sublimation occurs) have to be modified to allow desorption of the bound water (secondary drying) only when all the ice has been removed, in order to avoid collapse of the dried cake, or product meltback. During manufacturing, a properly designed sensor is thus required to point out if the selected operating conditions allow fulfilling the operating constraint about product temperature, and when primary drying is completed. Similarly, in the phase of recipe design, such sensor should be able to identify the optimal operating conditions that allow minimizing the duration of the process.[1-3]

The search for a sensor having these characteristics is motivated also by the Guidance for Industry PAT (Process Analytical Technology), that encourages the design of innovative tools for in-line product quality monitoring, thus avoiding testing product quality at the end of the process.[4] A wide literature, presenting numerous devices that can be used in lab-scale freeze-dryers and, in some cases, also in industrial-scale units, is available to freeze-drying practitioners. Among these devices, those using some kind of physical measures (e.g. product temperature using a thermocouple, or chamber pressure using a capacitance manometer) and a mathematical model of the process appear to be particularly promising. These devices can be grouped into two categories: those based on the pressure rise test and the observers. The pressure rise test consists of closing the valve in the duct connecting the drying chamber to the condenser for a short time interval: the pressure in the chamber increases and the measured values are compared to those calculated using a mathematical model of the process. The unknown parameters (e.g. the ice temperature and the residual moisture) are retrieved by looking for the best fit between the measured and calculated values of the chamber pressure.[5-10] The observer is a mathematical algorithm that uses some experimentally available measurements from a process (in this case, product temperature measured with a thermocouple) to “correct” model calculations and to estimate some unknown variables. Both the high gain technique and the Kalman filter were proposed to design the observer to monitor the primary drying of a vial freeze-drying process.[11-13] The observer allows monitoring the process taking into account the non-uniformity of the batch of vials (due to various reasons, e.g. non-uniform shelf temperature, uncontrolled nucleation, and the pressure gradients into the drying chamber [14,15]) using the measurement of product temperature in various vials in the batch. Besides, the observer can be used in a control loop with the goal to minimize the duration of the primary drying.[16]

Drăgoi et al.[17] proposed a new soft-sensor, based on artificial neural network (ANN), that was able to monitor the freeze-drying process using the temperature measurement at the bottom of the vial. The sensor had the goal to estimate the thickness of the dried cake and, thus, the ending point of the primary drying.

The ANNs (which are one of the many tools the Artificial Intelligence field puts at the disposal of the researcher) are mathematical models that simulates the behavior of the human brain, being composed of neurons organized into layers that perform specific calculations. Thus, the ANN models are not based on the physical and chemical laws governing the processes, and they can be particularly useful when process dynamics is too complex or not perfectly known. Moreover, the calculations required by ANN models are very simple and this can be of outmost importance when the calculation speed plays an important role as in process monitoring and control.[18-22] These very important characteristics are the main motivation for choosing the ANNs over other artificial intelligence techniques.

The open literature presents several approaches related to the use of neural networks as soft sensors. For instance, a feed-forward neural network-based software sensor is applied for coagulation control in a water treatment plant, having as key feature the ability to take into account various sources of uncertainty such as atypical input data, measurement errors, and limited information content of the training set.[23] An artificial neural network was built for real-time prediction of the moisture and fat content in olive pomace, using two-phase olive oil processing. The results obtained indicate a very good predictive capacity of the three-layer ANN model.[24] In the freeze-drying field, the neural networks are scarcely used to model and predict the behavior of various food products (carrots[25], strawberries[26], and apples[27]). Recently, Todorov et al.[28] and Todorov and Tsvetkov[29] have proposed a Volterra Fuzzy-Neural model in a nonlinear model predictive controller with the goal to reduce the duration of the sublimation step.

The ANN models are very easy to set up and use, but their performance is highly dependent on their structure and, thus, the determination of the optimal topology (number of inputs, outputs, hidden layers and neurons in each hidden layer) and of the optimal internal parameters (weights, biases and activation functions) are crucial steps. Various methods were proposed [30] to achieve these results, and evolutionary techniques appears to be particularly useful due to the ability to escape local minima, to their robustness, and to the ability to adapt in a changing environment.[31-37]

Drăgoi et al.[17] combined a self-adaptive differential evolution with back-propagation algorithm as a local search method to improve the best solution (neural network) obtained locally. This approach is used in this paper to set-up a soft sensor that not only estimates the residual amount of ice in the product (or the dried cake thickness), but also the maximum product temperature and the drying time in case the operating conditions are not modified in the future. A detailed model of the process can be used to calculate the learning patterns, with the goal to minimize the experimental effort. [38-43] In this work, the detailed mono-dimensional model of Velardi and Barresi is used.[43] Besides, process simulation using the same detailed model is used for validation purposes. In fact, mathematical simulation allows precisely knowing the values of product temperature and of ice content in the vial, thus being possible to accurately evaluate the error of the estimated variables. Such precise evaluation would be almost impossible to perform experimentally. This methodology poses a problem: the learning set is calculated using a specific set of values of the heat transfer coefficient (between the heating shelf and the product at the bottom of the container) and of the mass transfer resistance (of the dried cake to vapor flow) that can be obtained either from a previous experimental investigation (with small or high uncertainty), either by estimations. Consequently, it is necessary to evaluate if reliable and accurate estimations can be obtained when the sensor is used to monitor a process characterized by different values of both coefficients.

The structure of the paper is the following: first, the ANN algorithm based on the combination of differential evolution and back-propagation algorithms is briefly described, and the proposed artificial neural networks designed to monitor the freeze-drying process are described. Then, a section for the suggested training procedure and the case study is inserted. Finally, results demonstrating the effectiveness of the proposed sensor are presented and extensively discussed.

Methods and Materials

Description of the ANN algorithm

In order to create a simple, flexible, and easy to use sensor for the primary drying of a freeze-drying process, a bio-inspired computing technique, based on artificial neural networks and an auto-adaptive hybrid version of the Differential Evolution (DE) algorithm, is here applied. This approach was chosen due to the specific characteristics of each algorithm and of the advantages coming from the combination between them. The ANNs are highly nonlinear and their structure can be more flexible and representative compared to other types of models.[44] Also, they have the ability of estimating system behavior based on incomplete data.[45] The main disadvantage of the ANNs consist in the difficulty of selecting the optimal topology and choosing the best training algorithm (which also includes the training parameters). An alternative possibility to solve these problems is represented by the use of DE algorithm for developing and training ANNs.

DE is a population based, simple and straightforward meta-heuristic; the key of its success consists in spontaneous self-adaptability, diversity control, and continuous improvement.[34] An own software was developed using C# and .NET Framework 4.0 to model and monitor the freeze drying process. The main aspects of the general methodology considered in this paper are:

i) automation of the DE parameter selection based on an auto-adaptive technique in which the control parameters are included into the algorithm itself;
ii) evolutionary automation of the topological and internal parameters optimization for ANNs, by considering a population of encoded ANNs as the population on which DE works;
iii) improvement of the optimization by a local search procedure performed by back-propagation algorithm, combined with global DE search;
iv) flexibility and high generalization properties of the final neural model.

The general steps of the algorithm are the followings:

1. Generation of a set of potential solutions, where each solution is a neural network encoded using a real value vector approach.
2. Set the current generation to 1.
3. While the current generation is smaller than the maximum allowed value, the followings are repeated:
   a. Evolve the population by means of mutation, crossover and selection.
   b. Select the best individual in the population.
   c. Improve the best individual by applying the back-propagation algorithm.
   d. Replace the current population with the evolved one.
   e. Increase the value of the current generation with 1.

4. Select the best individual from the current population as the determined ANN model for the problem at hand.

The mutation and crossover procedures from the step 3.a are the means through which automatic optimization of ANNs is performed, having a significant influence on the overall results of the methodology. In the step 3.b, the best individual in the population is chosen based on its fitness function that measures the adaptation of each ANN to the environment, represented in this case by the system describing the primary drying. Since the scope of the methodology is to find the best possible model of the considered process, the fitness function is based on the mean square error (MSE) between the predicted and the desired values. Into the methodology, a combined stop criteria (consisting in a predefined number of iteration or MSE on a simple data set reaching a low value equal to 10^{-8}) was used. Throughout the evolutionary process, each neuron of the network can have one of the eight types of activation functions: Linear, Hard Limit, Bipolar Sigmoid, Logarithmic Sigmoid, Tangent Sigmoid, Sinus, Radial Basis, and Triangular Basis functions. Further details about the algorithm can be found in Ref.[17]

Description of the proposed artificial neural networks

A general structure of the ANN is first considered (Figure 1A) and then, based on the process particularities, it is refined in order to find the best model (Figure1B). This implies the determination of the inputs, outputs, and number of time delayed recurrences of each output. The time delayed inputs are considered because the model is used not only for predicting the system at a future specified moment, but also for determining the maximum temperature and the drying time. This means that it must predict various variables for the duration of the entire process without having any kind of information provided, except for the state of the system at the chosen moment of time. In order to generate the entire dynamic of the system based on known information from a specific moment, the ANN must be very precise, the incorporation of time delay inputs allowing much better results than the simple case when the ANN is used without recurrences. In order to determine the optimal number of inputs, a series of simulations with different combinations were considered (Table 1). In Table 1, for the best neural network obtained in each case, the following parameters are listed: the input variables, topology, MSE in the training phase ($MSE_{training}$), and MSE in the initial testing phase ($MSE_{testing}$), where a simple dataset was used. The methodology provides an optimized neural network from both topological and parametric point of view. Consequently, different comparisons (in different testing situations) between the models obtained in each case are redundant since the difference between $MSE_{testing}$ is quite large (as it can be observed from Table 1).

Figure 1B indicates the best neural network has 9 inputs, 1 hidden layer with 6 neurons, and 2 outputs. From considered inputs of the network 3 are fixed (time, temperature of the heating shelf, T_{shelf}, pressure in the drying chamber, $P_{chamber}$) and 6 are time recurrent (3 for each output). The 3 fixed parameters were chosen based on the idea that they have a major influence on the status of the process and can be relatively easy to measure. On the other hand, it is difficult to measure product temperature (T) and the thickness of the dried cake (L_{dried}). Also, the number of recurrences for the ANN must be chosen carefully, too low value having small effect on the outputs and too high value rising the complexity of the network. Consequently, a set of simulations was performed and, based on the obtained results, it was concluded that a number of 3 time delays for each output is a good trade-off.

The outputs of the networks (T and L_{dried}) are chosen based on the following considerations: i) during freeze drying, the temperature of the product must be in a specific interval, thus knowing that its estimated value over long periods of time is an important aspect in choosing the best operating conditions; ii) the thickness of the dried layer is an indication of the status of the process and, based on the difference between this value and the thickness of the product after freezing (that corresponds to the thickness of the dried cake at the end of primary drying, in case no shrinkage or collapse occurs), the ANN knows when the process is finished. Consequently, when the predicted value of the dried layer is equal to the product thickness, the model extracts the time input used to make the prediction and provides it to the end user as the drying time. It has to be remarked that once the dynamics of the dried cake thickness is known, then it can be possible to calculate the sublimation flux from a simple mass balance at the interface of sublimation:

$$J_w = \left(\rho_{frozen} - \rho_{dried}\right)\frac{dL_{dried}}{dt} \tag{1}$$

and, from this, the dynamics of the frozen layer thickness (i.e. of the residual amount of ice in the product) can be calculated:

$$\frac{dL_{frozen}}{dt} = -J_w \tag{2}$$

Description of the training procedure

In order to design the ANNs, a set of data from which the network can learn is required. Although the ANNs can provide good models when using incomplete or high error data, the degree of performance is not the same as when using small error data. Taking into consideration that the scope of the ANN model is to estimate the evolution of the product in different conditions, the required performance for good predictions is quite high. Consequently, different aspects of the training data must be taken into account: i) reduction of the measurement errors; ii) good coverage of the search space, and iii) enough number of training exemplars.

In this case, the learning data are provided by a detailed mono-dimensional model which considers all the chemical and physical laws that govern the process.[43] The goal of using an already existing model is to eliminate the need for experimental procedures which are time consuming and to provide a set of data with small measurement errors.

In order to obtain a good coverage of the search space, the training data was selected based on the fact that the most important operating parameters are $P_{chamber}$ and T_{shelf}. The representation of this pair is realized in a two dimensional plan (a rectangle) determined by the physical limits imposed by the equipment and characteristics of the drying product. If the third parameter (represented by the time) is considered, the $P_{chamber}$ and T_{shelf} variation in time can be delimited in a three dimensional box. Since the primary drying was recorded from start to finish, and for each pair ($P_{chamber}$, T_{shelf}) a set of data was collected at an interval of 600s, a set of 9 pairs (4 from the corners of the search space delimiting the box and 5 randomly chosen from inside) was considered to be enough for the network learning. This value was chosen based on the complexity of the process and on the author experience related to the training and testing optimal neural network models. It is considered as a useful practice to have a good coverage of the search space, but a high number of training data can lead to high computational time. The methodology used here for determining the neural model, in its turn, has problems related to the high computational time required. One significant advantage of our software is that once a good model is generated (although its determination is time consuming), it can be easily used for modeling or monitoring the process in various conditions.

Case study description

The case study investigated in this paper is the freeze-drying of a 10% w/w sucrose aqueous solution. The values of the heat transfer coefficient (K_v), used to model the dependence of the heat flux from the shelf to the product at the bottom of the vial (J_q) on the driving force:

$$J_q = K_v \left(T_{shelf} - T_B \right) \qquad (3)$$

and of the resistance of the dried layer (R_p), used to model the dependence of the sublimation flux from the interface of sublimation to the drying chamber (J_w) on the driving force:

$$J_w = \frac{1}{R_p}\left(P_{\mathrm{w,i}} - P_{w,c}\right) \tag{4}$$

are taken from Ref.[46] In particular, the dependence of K_v on $P_{chamber}$ is modeled by means of the following equation:

$$K_v = 3.25 + \frac{0.74 \cdot P_{chamber}}{1 + 0.02 \cdot P_{chamber}} \tag{5}$$

and the dependence of R_p on L_{dried} is modeled with Equation (6):

$$R_p = 10^4 + \frac{1.4 \cdot 10^8 \cdot L_{dried}}{1 + 542 \cdot L_{dried}} \tag{6}$$

These values have been obtained in a LyoBeta 25 freeze-dryer (Telstar, Spain); it is characterized by a vacuum-tight chamber (volume of 0.2 m^3) equipped with four shelves, that provide a total area of 0.5 m^2. The vacuum system consists of an external condenser (maximum ice capacity $\approx$ 40 kg) and a vacuum pump to remove the incondensable gases. The formulation is processed in tubing vials having an internal diameter of 14.25 mm. The thickness of the frozen product at the beginning of the drying is equal to 10 mm. The intervals of values of the shelf temperature and of the chamber pressure used in the training phase are the followings: [243.15, 273.15] K and [5, 20] Pa, respectively. Product temperature can be measured by means of a thermocouple, inserted in one of the vials of the batch and put in contact with the vial bottom.

Results and discussion

As it has been pointed out in the introduction section, a detailed phenomenological model of the process is used to calculate the dynamics of the product in the vial and, thus, the learning data for the neural networks. By this way, it is possible to correctly evaluate the performance indexes of the neural network model when computing the evolution of product temperature and of cake thickness, as well as the maximum product temperature and the drying time for a given set of operating conditions. Besides, the goal of the neural network model is to replace the detailed model in the control loop, e.g. for the in-line optimization of the process, as the calculation time (for predictions) is significantly reduced. In order to assess the performance of the proposed model, different cases in which various values of the difference between the K_v used in the training phase ($K_{v,training}$) and in the monitored process ($K_{v,monitoring}$), and between the R_p used in the training phase ($R_{p,training}$) and in the monitored process ($R_{p,monitoring}$) were considered.

Figure 2 shows a comparison between the values estimated using the soft-sensor based on the neural network model and the true values of the product temperature and of the frozen layer thickness during a cycle run with T_{fluid} = -20°C and $P_{chamber}$ = 10 Pa. In this case, the difference between the parameters of the monitored process and the data used for training is -20% for K_v and +20% for R_p. Nevertheless, the accuracy of the sensor is very high for both monitored variables, thus evidencing the capacity of the ANN model to provide fairly accurate estimations even when K_v and R_p values are different from the ones used in the training phase. This is of outmost importance because: i) the experimentally determined values of both parameters can be affected by significant errors; ii) in some cases, the experimental investigation can be unfeasible, e.g. in industrial-scale freeze-dryers. Thus, the training of the ANN model is based on the values of K_v and R_p taken from the literature and these values are affected, among the others, by the type of freeze-dryer and vial used, as well as by the nucleation phenomena occurring in the freezing phase (completely stochastic phenomena).

Figure 3 summarizes the errors on the estimations of product temperature and on the drying time as a function of the difference between $K_{v,monitoring}$ and $K_{v,training}$ (graphs A and B). Even in case the difference on the value of the heat transfer coefficient is very large (+50%), the error on the estimation of product temperature is almost negligible. With respect to the duration of primary drying, in all cases, the error can be considered negligible (lower than 1.5%). In addition, it can be evidenced that when the difference between $K_{v,monitoring}$ and $K_{v,training}$ increases, the error on the drying time reduces, while that on product temperature increases. In fact, the values of product temperature, sublimation flux (that determines the drying time), K_v and R_p are strictly related as they have to satisfy the energy balance at the interface of sublimation, that means that all the heat transferred by the shelf to the product is used for ice sublimation. This means that in case the error on K_v is high, the error on product temperature is expected to increase (as this is the variable estimated by the neural network trained for a different value of the parameter) and, thus, the drying force for the heat transfer to the product (given by the difference between the shelf temperature and the product temperature) is expected to decrease. This means that the sublimation flux will remain almost unaffected (in this case there are no errors on the value of dried cake resistance), or that it will exhibit a small variation (increasing or decreasing as a function of the error, depending on the values of K_v and of the driving force), as it is shown in graph B of Figure 3.

A similar analysis has been carried out with respect to the effect of the difference between $R_{p,monitoring}$ and $R_{p,training}$ (Figure 3, graphs C and D). It appears evident that the effect of the error on this variable is responsible for a higher error on product temperature, even if very low values are obtained in all cases. This means that the ANN model has to be used for the same type of product considered in the training phase (as limited variation of the R_p values are expected as a consequence of the variation in the nucleation temperature). Therefore, great care has to be paid: i) when using the ANN model with a different product (as the values of R_p would be significantly different); or ii) when using this sensor for the same type of product, but with a different filling volume or a different concentration (both variables can affect the values of resistance of the dried cake).

Figure 4 summarizes the joint effect of the difference between $K_{v,monitoring}$ and $K_{v,training}$ and of the difference between $R_{p,monitoring}$ and $R_{p,training}$. While the error on the duration of primary drying ranges from 0.7 % and 1.7 %, and, thus, it is fully acceptable, the error on product temperature remains bounded from $-2\cdot10^{-4}$ K to $+2\cdot10^{-4}$ K only in case the difference between the value of $R_{p,monitoring}$ and the value used in the training phase is below 20-25 %. If the difference on R_p is +50%, and there is also a significant difference on K_v, then the error on product temperature can increase to 5 K.

Results shown in Figures 3 and 4 evidence that the soft-sensor based on the ANN model can effectively applied to monitor a process where the values of the heat transfer coefficient from the shelf to the product at the bottom of the vial, and the values of the resistance of the dried cake to vapor flow are different from those used in the training phase. The use of experimental measurements from the process is able to reduce the error on product temperature and on duration of primary drying. In particular, the difference on the value of $K_{v,training}$ has a limited effect on the accuracy of the estimations. Thus, the same sensor can be used to monitor the evolution of the product in various vials of the same batch, as well as in different freeze-dryers (as these issues can affect the value of K_v). Instead, the effect of the difference on R_p is much more relevant, and the sensor can be used to monitor only the type of product considered in the training phase.

Further simulations have been carried out to test the effectiveness of the sensor designed to estimate the maximum product temperature and the duration of the drying time on the basis of the measurement of some process variables at a generic time instant *t*. In this case the capacity of the sensor to estimate product dynamics was tested, not only up to the following time instant ($t + \Delta t$) when other experimental measurements become available, but up to the end of primary drying (according to the structure of Figure 1). As it is shown in

Figure 5, the errors of the estimations of maximum product temperature and of the drying time become smaller as time goes on. In particular, the error on the estimation of the maximum product temperature is very small (about 0.6 K in this case) at the beginning of primary drying, while the accuracy of the estimations of drying time is improved in the second half of primary drying. As a consequence, at the onset of primary drying, the proposed soft-sensor allows to understand if the product temperature remains below the limit value or not and, thus, if the operating conditions have to be modified. With respect to the duration of primary drying, the uncertainty on the estimated values can be significant. Nevertheless it has to be remarked that the estimated duration is always higher than the current time and, at time t, the sensor will never provide the (erroneous) information that the drying is concluded, but in all cases the indication will be that more time is required. Then, when primary drying is approaching to the end, the estimated value of this variable becomes very accurate. As in previous cases, there is a remarkable difference between $K_{v,monitoring}$ and $K_{v,training}$, and between $R_{p,monitoring}$ and $R_{p,training}$, but the estimation of the maximum product temperature is very accurate.

Similar conclusions can be obtained when investigating the effect of the difference between $K_{v,monitoring}$ and $K_{v,training}$ (Figure 6, graphs A and B). The error on the maximum product temperature is higher at the beginning of primary drying, where it ranges from -0.5 K and +0.2 K, and it is negligible near the end of the primary drying. This can be perfectly accepted. With respect to the error on the duration of primary drying, when about 50% of the ice has disappeared, the error on the estimation of drying time is roughly -20% and it reduces as far as drying goes on. Finally, it can be evidenced that the effect of the difference of K_v has a poor effect on both the estimated variables. The study has been repeated to investigate the effect of the difference between $R_{p,monitoring}$ and $R_{p,training}$, and similar conclusions are obtained (Figure 6, graphs C and D). While the error on drying time is of the same order of magnitude of that caused by K_v, the error on the estimated values of maximum product temperature can be fairly high (about 2 K) at the beginning of primary drying, in case the R_p difference is higher than 20%. This means that also in this case the resistance of the dried cake is the most critical parameter for the performance of the sensor, and that the sensor has to be used only with the same type of product whose parameters have been used in the training phase.

Figure 7 summarizes the errors on maximum product temperature and on drying time as a function of the difference between $K_{v,monitoring}$ and $K_{v,training}$ and between $R_{p,monitoring}$ and $R_{p,training}$, in the middle and at the end of primary drying. As the results indicate, the previous conclusions are confirmed.

Conclusions

Artificial neural networks can be effectively used to design a soft-sensor to monitor a freeze-drying process, with the goal to estimate the dynamics of the dried cake thickness (and, then, of the residual amount of ice, as well as the sublimation flux), the product temperature, the duration of primary drying and the maximum product temperature for a given set of operating conditions. A detailed phenomenological model was used to calculate the data required in the training phase of the neural network design. In order to reduce the time necessary for designing the sensor, a specific couple of values of the overall heat transfer coefficient from the heating shelf to the product in the container and of the dried cake resistance to vapor flow is selected. The error committed by the sensor when monitoring a process characterized by different values of both coefficients is generally small, in particular with respect to product temperature. Thus, the proposed sensor could be used for both in-line process monitoring and in-line process optimization, as the time required by calculations is very small when using artificial neural networks.

Acknowledgements

This work was partially done by financial support provided by EURODOC "Doctoral Scholarships for research performance at European level" project, financed by the European Social Found and Romanian Government and by the project PN-II-PT-PCCA-2011-3.2-0732 financed by Romanian Research Agency UEFISCDI.

List of Symbols

J_q	heat flux, W m^{-2}
J_w	vapor flux, kg m^{-2} s^{-1}
K_v	overall heat transfer coefficient between the shelf and the product at the bottom of the container, W $m^{-2}K^{-1}$
$K_{v,monitoring}$	value of K_v in the monitored system, W $m^{-2}K^{-1}$
$K_{v,training}$	value of K_v used in the training phase, W $m^{-2}K^{-1}$
L_{dried}	dried cake thickness, m
L_{frozen}	frozen layer thickness, m
L_0	product thickness after freezing, m
$P_{chamber}$	pressure in the drying chamber, Pa
$P_{w,c}$	partial pressure of water into the drying chamber, Pa
$P_{w,i}$	partial pressure of water at the interface, Pa
R_p	resistance of the dried layer to vapor flow, m s^{-1}
$R_{p,monitoring}$	value of R_p of the monitored product, m s^{-1}
$R_{p,training}$	value of R_p used in the training phase, m s^{-1}
T	product temperature (at the interface of sublimation), K
T_B	product temperature at the vial bottom, K
T_{shelf}	temperature of the heating shelf, K
T_{max}	maximum product temperature (at the interface of sublimation), K
t	time, s
t_d	drying time, h

Greek letters

ρ_{dried}	apparent density of the dried product, kg m^{-3}
ρ_{frozen}	density of the frozen product, kg m^{-3}

Abbreviations

ANN	artificial neural network
DE	differential evolution
PSO	particle swarm optimization

References

1. Liapis, A.I.; Pikal, M.J.; Bruttini, R. Research and development needs and opportunities in freeze drying. Drying Technology **1996**, *14*, 1265–1300.
2. Sadikoglu, H.; Ozdemir, M.; Seker, M. Freeze-drying of pharmaceutical products: Research and development needs. Drying Technology **2006**, *24*, 849-861.
3. Barresi, A.A.; Fissore, D.; Product quality control in freeze drying of pharmaceuticals. In Modern Drying Technology: Product Quality and Formulation; Tsotsas, E., Mujumdar, A. S., Eds.; Wiley-VCH: Weinhein, Germany, 2011; 91-154.
4. US Food and Drug Administration. PAT guidance for industry – a framework for innovative pharmaceutical development, manufacturing and quality assurance. US Department of Health and Human Services, Food and Drug Administration, Center for drug evaluation and research, Center for veterinary medicine, Office of regulatory affairs: Rockville, 2004.
5. Milton, N.; Pikal, M.J.; Roy, M.L.; Nail, S.L. Evaluation of manometric temperature measurement as a method of monitoring product temperature during lyophilisation. PDA Journal of Pharmaceutical Science and Technology **1997**, *5*, 7-16.
6. Liapis, A. I.; Sadikoglu, H. Dynamic pressure rise in the drying chamber as a remote sensing method for monitoring the temperature of the product during the primary drying stage of freeze-drying. Drying Technology **1998**, *16*, 1153-1171.
7. Chouvenc, P.; Vessot, S.; Andrieu, J.; Vacus, P. Optimization of the freeze-drying cycle: a new model for pressure rise analysis. Drying Technology **2004**, *22*, 1577-1601.
8. Velardi, S.A.; Rasetto, V.; Barresi, A.A. Dynamic Parameters Estimation Method: advanced Manometric Temperature Measurement approach for freeze-drying monitoring of pharmaceutical. Industrial Engineering Chemistry Research **2008**, *47*, 8445-8457.
9. Fissore, D.; Pisano, R.; Barresi, A.A. On the methods based on the Pressure Rise Test for monitoring a freeze-drying process. Drying Technology **2011**, *29*, 73-90.
10. Pisano, R., Fissore, D., Barresi, A.A. Innovation in monitoring food freeze-drying. Drying Technology **2011**, *29*, 1920-1931.
11. Velardi, S.A.; Hammouri, H.; Barresi, A.A. In-line monitoring of the primary drying phase of the freeze-drying process in vial by means of a Kalman filter based observer. Chemical Engineering Research & Design **2009**, *87*, 1409-1019.
12. Velardi, S.A.; Hammouri, H.; Barresi, A.A. Development of a High Gain observer for in-line monitoring of sublimation in vial freeze-drying. Drying Technology **2010**, *28*, 256-

268.
13. Bosca, S.; Fissore, D. Design and validation of an innovative soft-sensor for pharmaceuticals freeze-drying monitoring. Chemical Engineering Science **2011**, *66*, 5127-5136.
14. Rasetto, V.; Marchisio, D. L.; Fissore, D.; Barresi A. A. On the use of a dual-scale model to improve understanding of a pharmaceutical freeze-drying process. Journal of Pharmaceutical Sciences **2010**, *99*, 4337-4350.
15. Barresi, A.A.; Pisano, R.; Rasetto, V.; Fissore, D.; Marchisio, D.L. Model-based monitoring and control of industrial freeze-drying processes: effect of batch non-uniformity. Drying Technology **2010**, *28*, 577-590.
16. Fissore, D.; Velardi, S.A.; Barresi A.A. In-line control of a freeze-drying process in vials. Drying Technology **2008**, *26*, 685-694.
17. Drăgoi, E. N., Curteanu, S., Fissore D. Freeze-drying modeling and monitoring using a new neuro-evolutive technique. Chemical Engineering Science **2011**, *72*, 195-204.
18. Ochoa-Martinez, C.I.; Ramaswamy, H.S.; Ayala-Aponte, A.A. Artificial neural network modeling of osmotic dehydration mass transfer kinetics of fruits. Drying Technology **2007**, *25*, 85-95.
19. Mujumdar, A.S.; Zhonghua, W. Thermal drying technologies - Cost-effective innovation aided by mathematical modeling approach. Drying Technology **2008**, *26*, 145-153.
20. Mihajlovic, T.; Ibric, S.; Mladenovic, A. Application of design of experiments and multilayer perceptron neural network in optimization of the spray-drying process. Drying Technology **2011**, *29*, 1638-1647.
21. Polat, K.; Kirmaci, V. A novel data preprocessing method for the modeling and prediction of freeze-drying behavior of apples: multiple output-dependent data scaling (MODDS). Drying Technology **2012**, *30*, 185-196.
22. Keshani, S.; Daud, W.R.W.; Woo, M.W.; Talib, M.Z.M.; Chuah, A.L.; Russly, A.R. Artificial neural network modeling of the deposition rate of lactose powder in spray dryers. Drying Technology **2012**, *30*, 386-397.
23. Valentina, N., Denoeux, T. A neural network-based software sensor for coagulation control in a water treatment plant. Intelligent Data Analysis **2001**, 5, 23–39.
24. Jimenez, A., Beltran, G., Aguilera, M. P., Uceda, M. A sensor-software based on artificial neural network for the optimization of olive oil elaboration process. Sensors and Actuators B **2008**, *129*, 985–990.
25. Erenturk, S., Erenturk, K. Comparison of genetic algorithm and neural network

approaches for the drying process of carrot. Journal of Food Engineering **2007**, *78*, 905-912

26. Menlik,T., Kirmaci, V.,Usta, H. Modeling of freeze drying behaviors of strawberries by using artificial neural network. Journal of Thermal Science and Technology **2009**, *29*, 11-21
27. Menlik,T., Ozdemir, M. B., Kirmaci, V. Determination of freeze-drying behaviors of apples by artificial neural network **2010**, *37*, 7669-7677.
28. Todorov, Y. V., Terziyska, M., Petrov, M. Nonlinear model predictive controller using a Fuzzy-Neural Wiener-Hammerstein model, In: Proceedings of 16th International Conference on Process Control, 2007, Article No 216, 1-7.
29. Todorov, Y. V., Tsvetkov, T. D. Volterra Model Predictive Control of a lyophilization plant, In: Proceedings of 4th International IEEE Conference – Intelligent Systems, 2008, 3, 2013-2018.
30. Rojas, R. Neural Networks. A systematic introduction. Springer: Berlin, 1996.
31. Xin, Y., Evolving artificial neural networks. In: Proceedings of the IEEE 87, 1999; 1423-1447.
32. Abbass, H A. A memetic Pareto evolutionary approach to artificial neural networks. In: Proceedings of the 14th Australian Joint Conference on Artificial Intelligence, 2001; 2256, 1-12.
33. Olafsson, S., Metaheuristics. In: Henderson, S. G., Nelson, B. L., Eds. Simulation. Handbooks in operations research and management science. Elsevier: Amsterdam, 2006; 633-654.
34. Feoktistov, V.. Differential Evolution: in search of solutions. Springer-Verlag: Berlin, 2006.
35. Floreano, D.; Durr, P.; Mattiussi, C. Neuroevolution: from architecture to learning. Evolutionary Intelligence **2008**, *1*, 47-62.
36. Bedri Ozer, A. CIDE: Chaotically Initialized Differential Evolution. Expert System with Applications **2010**, *37*, 4632-4641.
37. Curteanu, S.; Cartwright, H.. Neural networks applied in chemistry. I. Determination of the optimal topology of multilayer perceptron neural networks. Journal of Chemometrics **2011**, *25*, 527-574.
38. Lombraña, J.I.; De Elvira, C.; Villaran, M.C.; Izcara, J. Simulation and design of heating profiles in heat controlled freeze-drying of pharmaceuticals in vials by the application of a sublimation semispherical model. Drying Technology **1993**, *11*, 471-487.

39. Lombraña, J.I.; De Elvira, C.; Villaran, M.C.; Izcara, J.. Simulation and design of heating profiles in heat controlled freeze-drying of pharmaceuticals in vials by the application of a sublimation cylindrical model. Drying Technology **1993**, *11*, 85-102.
40. Sadikoglu, H.; Liapis, A.I. Mathematical modelling of the primary and secondary drying stages of bulk solution freeze-drying in trays: parameter estimation and model discrimination by comparison of theoretical results with experimental data. Drying Technology **1997**, *15*, 791-810.
41. Gan, K.H.; Crosser, O.K.; Liapis, A.I.; Bruttini, R. Lyophilisation in vials on trays: effects of tray side. Drying Technology **2005**, *23*, 341-363.
42. Hottot, A.; Peczalski, R.; Vessot, S.; Andrieu, J. Freeze-drying of pharmaceutical proteins in vials: modeling of freezing and sublimation steps. Drying Technology **2006**, *24*, 561-570.
43. Velardi, S. A., Barresi, A. A.. Development of simplified models for the freeze-drying process and investigation of the optimal operating conditions. Chemical Engineering Research and Design **2008**, *87*, 9-22.
44. Himmelblau, D. Applications of artificial neural networks in chemical engineering. Korean Journal of Chemical Engineering **2000**, 17, 373-392.
45. Kisi, O. River suspended sediment concentration modeling using a neural differential evolution approach. Journal of Hydrology **2010**, 389, 227-235.
46. Giordano, A.; Barresi, A.A.; Fissore, D. On the use of mathematical models to build the design space for the primary drying phase of a pharmaceutical lyophilization process. Journal of Pharmaceutical Sciences **2011**, *100*, 311-324.

List of Tables

Table 1. Effects of various input combinations used for determining the best neural network for the primary drying phase on the $MSE_{training}$ and on the $MSE_{testing}$.

List of Figures

Figure 1 General structure (graph A) and best model (graph B) of the neural network model used to estimate the maximum product temperature and the drying time.

Figure 2 Comparison between the values estimated using the soft-sensor based on ANN (solid lines) and the true values (symbols) of product temperature (A) and of the thickness of the frozen layer (B), in case the difference on K_v is -20% and the difference on R_p is +20%, with respect to the parameters used in the training phase.

Figure 3 Maximum errors on the estimations of product temperature (A) and of drying time (B) as a function of the difference between $K_{v,monitoring}$ and $K_{v,training}$ (graphs A and B, respectively), and of the difference between $R_{p,monitoring}$ and $R_{p,training}$ (graphs C and D, respectively),

Figure 4 Maximum errors on the estimations of product temperature (A) and of drying time (B) as a function of the difference between $K_{v,monitoring}$ and $K_{v,training}$, for different values of error on R_p (■ = -50 %, ● = -20 %, ▲= -10 %, □ = +10 %, ○ = +20 %, Δ = +50 %).

Figure 5 Comparison between the time evolution of the estimations (solid lines) of maximum product temperature (A) and of drying time (B), and the correct values (dashed lines), in case the difference on K_v is -20% and the difference on R_p is +20%, with respect to the parameters used in the training phase.

Figure 6 Errors on the estimations of the maximum product temperature and of drying time at the beginning of primary drying (●), in the middle of this phase (▲), and at the end of primary drying (■) as a function of the difference between $K_{v,monitoring}$ and $K_{v,training}$ (graphs A and B respectively), and of the difference between the value of $R_{p,monitoring}$ and $R_{p,training}$ (graphs C and D respectively),

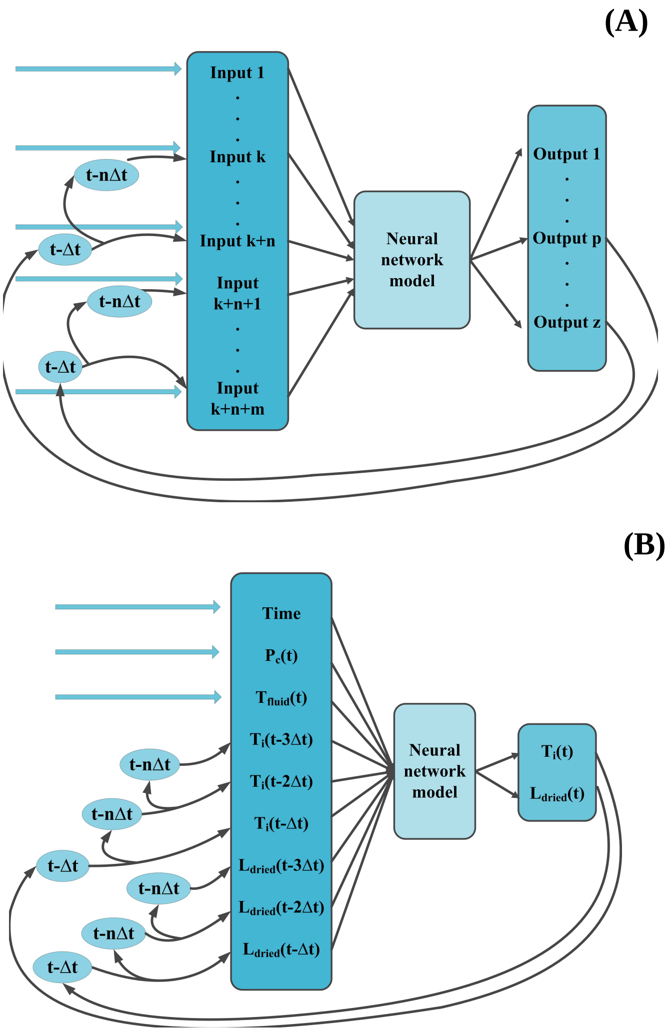
Figure 7 Errors on estimations of the maximum product temperature (A) and of drying time (B) as a function of the difference between $K_{v,monitoring}$ and $K_{v,training}$, for different

values of error on R_p (■ = -50 %, ● = -20 %, ▲ = -10 %, □ = +10 %, ○ = +20 %, Δ = +50 %) in the middle (left hand side) and at the end (right hand side) of primary drying.

Table 1

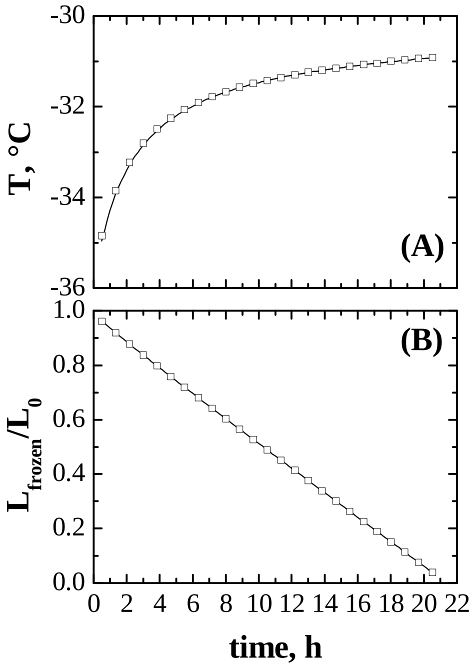
Case	Inputs									Outputs		Topology of the best ANN	$MSE_{training}$ of the best ANN	$MSE_{testing}$ of the best ANN
	Time	T_{shelf}	P_c	T_i @ t-3Δt	T_i @ t-2Δt	T_i @ t-Δt	L_{dried} @ t-3Δt	L_{dried} @ t-2Δt	L_{dried} @ t-Δt	T_i @ t	L_{dried} @ t			
1	x	x	x			x				x	x	4:9:2	$4.82\cdot10^{-4}$	$5.63\cdot10^{-3}$
2	x	x	x			x			x	x	x	5:4:4:2	$2.78\cdot10^{-4}$	$5.4\cdot10^{-3}$
3	x	x	x		x	x			x	x	x	6:7:2	$6.38\cdot10^{-5}$	$8.61\cdot10^{-4}$
4	x	x	x		x	x		x	x	x	x	7:5:3:2	$3.27\cdot10^{-5}$	$3.69\cdot10^{-5}$
5	x	x	x	x	x	x			x	x	x	7:8:2	$1.46\cdot10^{-6}$	$4.86\cdot10^{-6}$
6	x	x	x	x	x	x	x	x	x	x	x	9:6:2	$9.37\cdot10^{-7}$	$3.79\cdot10^{-7}$

Figure 1



(A)
Input 1
Input k
Input k+n
Input k+n+1
Input k+n+m
t-nΔt
t-Δt
t-nΔt
t-Δt
Neural network model
Output 1
Output p
Output z
(B)
Time
$P_c(t)$
$T_{fluid}(t)$
$T_i(t-3\Delta t)$
$T_i(t-2\Delta t)$
$T_i(t-\Delta t)$
$L_{dried}(t-3\Delta t)$
$L_{dried}(t-2\Delta t)$
$L_{dried}(t-\Delta t)$
t-nΔt
t-nΔt
t-Δt
t-nΔt
t-nΔt
t-Δt
Neural network model
$T_i(t)$
$L_{dried}(t)$

Figure 2



-30
-32
-34
-36
T, °C
(A)
1.0
0.8
0.6
0.4
0.2
0.0
L_{frozen}/L_0
(B)
0 2 4 6 8 10 12 14 16 18 20 22
time, h

Figure 3

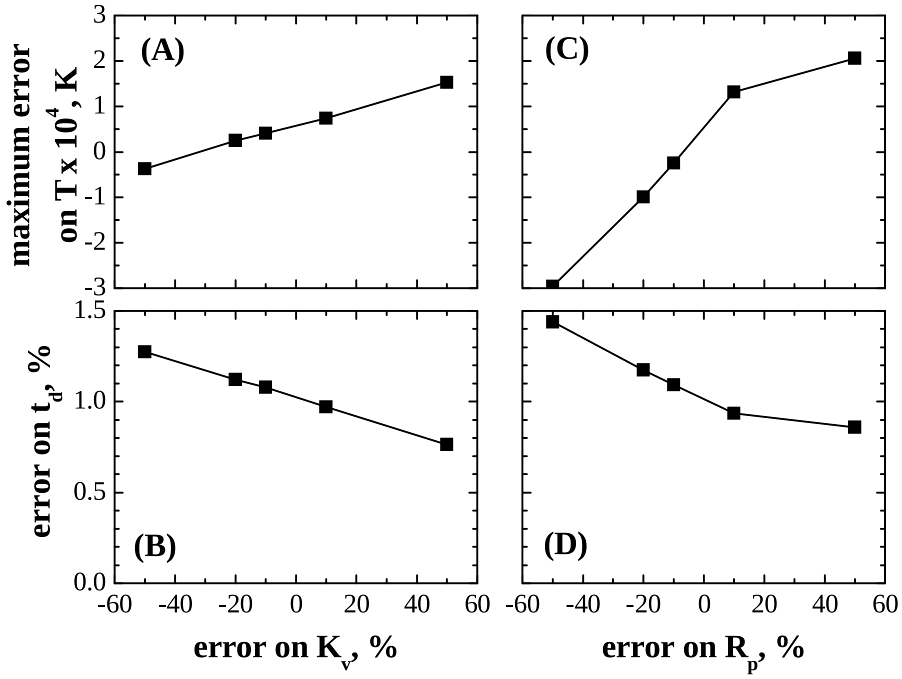
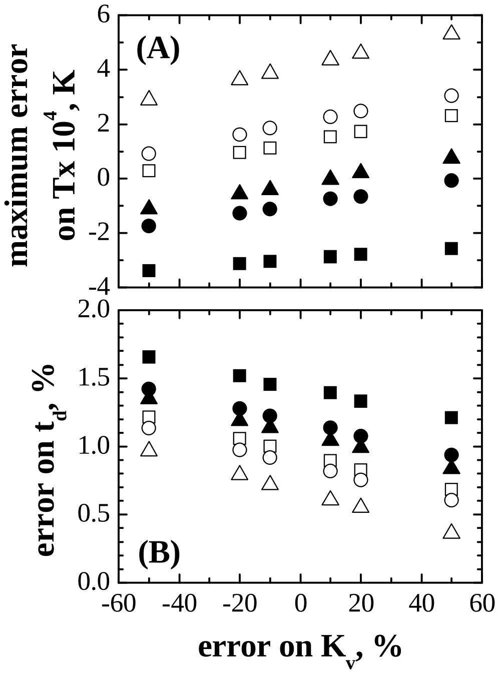
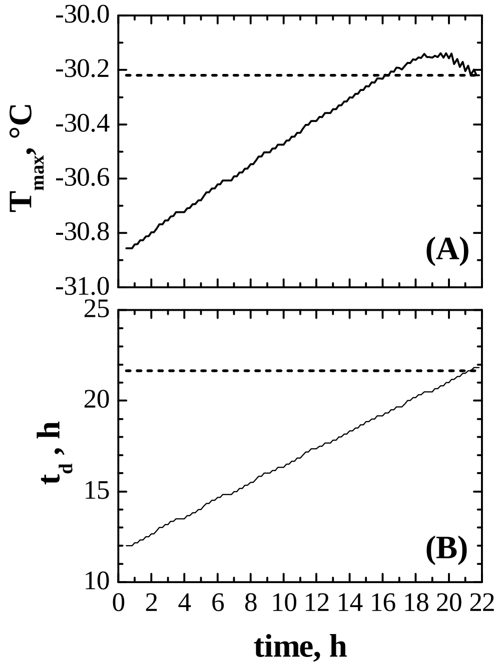

(A)
(B)
(C)
(D)
maximum error on T x 10^4, K
error on t_d, %
error on K_v, %
error on R_p, %
3
2
1
0
-1
-2
-3
1.5
1.0
0.5
0.0
-60
-40
-20
0
20
40
60

Figure 4



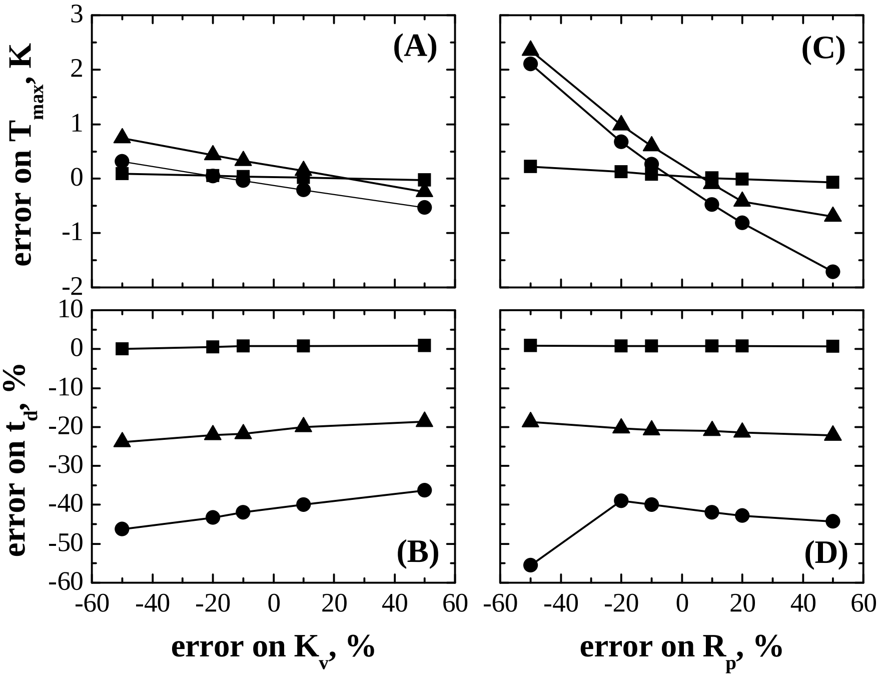
(A)
(B)
maximum error on Tx 10^4, K
error on t_d, %
error on K_v, %
6
4
2
0
-2
-4
2.0
1.5
1.0
0.5
0.0
-60
-40
-20
0
20
40
60

Figure 5



-30.0
-30.2
-30.4
-30.6
-30.8
-31.0
T_{max}, °C
(A)
25
20
15
10
t_d , h
(B)
0 2 4 6 8 10 12 14 16 18 20 22
time, h

Figure 6



(A)
(B)
(C)
(D)
error on T_{max}, K
error on t_d, %
error on K_v, %
error on R_p, %
3
2
1
0
-1
-2
10
0
-10
-20
-30
-40
-50
-60
-60
-40
-20
0
20
40
60

Figure 7