Global Optimization of a Dryer by Using Neural Networks and Genetic Algorithms

A. Hugget, P. Sébastian, and J.-P. Nadeau
L.E.P.T.-ENSAM, Esplanade des Arts et Métiers, 33405 Talence Cedex, Bordeaux, France

For many optimum design problems, the objective function is the result of a complex numerical code and may not be differentiable and explicit. The first aim is to propose a way of solving such complexity on an example problem. A novel and global strategy involving artificial neural networks and a genetic algorithm is presented and validated for an industrial convective dryer. To begin with, a method to represent a drying model using artificial neural networks is defined. This method is tested and the results are compared with those obtained with classic numerical methods. This approach drastically reduces simulation times and maintains good accuracy and generalization properties. Second, the associated optimal design problem is considered. This optimization appears as a difficult combinatorial problem with a complex objective function that involves different economical criteria. The second aim is to present the methodology to solve this problem using genetic algorithms. Final results illustrate the efficiency of this global approach.

Introduction

Solid-particulate drying is a simultaneous heat and mass-transfer operation that is present in important industrial processes involving mineral, agricultural products, forest, and polymer treatment. In this field, diversity of equipment, heterogeneous feed characteristics, as well as complex heat and mass-transfer mechanisms, and evolving properties of the material being dried make the design of dryer systems very difficult (Nadeau and Puiggali, 1995). A typical design problem may be defined as follows: given the large amount of equipment that is used in a dryer, how do we quickly and economically optimize both the equipment (type and size) and process variables needed to produce the best quality product using the least amount of energy and taking the relevant safety and environmental requirements into account?

The optimization of such a complex system is a more generic problem and usually involves two successive steps. The first step consists of representing total system behavior by the use of a very fast and accurate method. Considering the drying field, the computation times required to accurately represent the process, the behavior of the different equipment, and their interactions are very large, particularly, for wood drying. The classic numerical methods do not appear to be suitable for the optimal design problem. The second step, induced by the first, is to find the best architecture for this system so as to minimize the objective function representation of the industrial requirements. Using classic local reduction methods proved to be inefficient for performing global optimization with a complex, nondifferentiable, objective function that involves several criteria.

The aim of this article is to suggest a way to improve the treatment of such problems. The global strategy is to use artificial neural networks (ANNs) to represent system behavior, and, for the optimal design problem, to use well-adapted genetic algorithms (GAs). The first major interest of this work is to define a new method for quickly representing a mathematical model using ANNs. The second interest is to solve the associated optimization problem, taking dryer equipment and process variables into account simultaneously, and using several economic criteria. ANNs and GAs used in this work are still classic, even if different new developments are added to increase their performances. The more important point is the methodology meant to solve the design problem under consideration. This global approach is applied for only one type of convective industrial dryer, but is accurately described for efficient use in numerous other problems. The details of the procedure are described.

Current address of A. Hugget: Dept. of Chemical Engineering, University of Technology, STW 0.39, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

From a survey of the literature we find that different methods exist to represent the way complex systems such as dryers behave. Simplified models appear to be suitable for performing quick numerical simulations (Kisakurek, 1982; Kaminski et al., 1989; Kiranoudis et al., 1996a,b) in spite of their inability to take into account all nonlinear interactions between elements and to supply the objective criteria necessary for design optimization. Furthermore, a more complex model would prohibit the optimization step due to the extensive CPU time required for the numerical simulations. Recently, ANNs have been used for design and control as well as in modeling a large variety of processes. ANNs include layers of fully interconnected parallel processing units (neurons). They are inherently parallel and they hold promise because of their ability to approximate any continuous nonlinear function through supervisory learning, provided that the number of hidden units is sufficient and all the activation functions of the hidden units are continuously differentiable (Cybengo, 1989; Hornick et al., 1989). ANN abilities have been studied by many researchers and used for many applications in industry, business, and science (Widrow et al., 1994). Most of the time, multilayer feedforward neural networks are used to approximate complex functions. In this way, ANNs are chosen rather than other approximation methods, such as Fourier series or splines, because this method actually seems to be the more efficient (Barron, 1993; Hornick et al., 1994). Many researchers (Bhat and McAvoy, 1990; Chen and Weigand, 1994; Thompson and Kramer, 1996; Van Can et al., 1996) have illustrated that ANNs appear to be suitable for modeling complex nonlinear processes. They use ANNs to map input/output signal sets with minimum error. Others, like Huang and Mujumdar (1993), Heyd et al. (1996), Jay and Oliver (1996), and Zbicinski et al. (1996a), have shown that ANNs can be used as a black box for drying modeling on the basis of experimental data and without any detailed knowledge of the system. This approach leads to accurate models, but appears to be specific for each dryer configuration, so it cannot be used for design optimization. Furthermore, Psichogios and Ungar (1994). Cubillos et al. (1996), Zbicinski et al. (1996a,b) have used hybrid neural models, referred to as gray-box models, to study drying processes. This method is based on mathematical models including ANNs to estimate uncertain parameters like heat or mass-transfer coefficients. The hybrid approach is used to obtain accurate models featuring extrapolated properties derived from experimental results, but remains limited due to the large CPU time involved in mathematical models.

In this article, a design operation is assumed to quickly evaluate different equipment. The simplified models and black-box or gray-box methods just presented cannot be used directly for this purpose. The approach proposed in this work is quite different and uses ANN as a way of representing a mathematical model. This method greatly reduces simulation time and maintains the accuracy and generalization properties of the initial model. There are two differences between classic ANN models and the ANN build in the present approach. First, ANN inputs are dimensionless numbers that represent both the technological characteristics of the dryer and process variables. These numbers are general enough to simulate various dryer configurations operating in different conditions. Furthermore, data sets used during the ANN training phase are accurately computed using classic mathematical models rather than experimental data. A similar approach has been developed to study the steady-state behavior of a tunnel dryer operating on polycarbonate mud (Hugget and Sébastian, 1996). This new methodology for representing a mathematical model using ANN is explained in detail.

The speed of the ANN simulation code makes it reasonable to consider solving the associated optimal design problem. Optimization parameters represent all the equipment involved in the dryer—fans, the heating device, drying sections—and operating conditions, such as inlet air temperature or fan regulation. Dryer equipment and process variables are optimized simultaneously in order to obtain dryer configurations really suited to the problem. The objective function combines different criteria, such as total annual cost (energy, equipment, personal, insurance, etc.), product flow rate, and quality of dried material, and is based on the concept of maximum efficiency (Kaminsky et al., 1989). This function results from a numerical code using ANN and is not explicit. Attempts to solve the preceding optimization problem using classic local descent methods, such as quasi-Newton or conjugate gradient algorithms, proved to be inefficient, because good estimates of the function derivatives are required. Unfortunately for this type of problem, the objective function is not continuous and differentiable, and can be extremely irregular. In this context, the optimization of a dryer and its design variables, such as fan or heating device, is still an opened question.

For this reason, we feel is necessary to adopt a different optimization strategy capable of dealing with such cases. We have consequently turned out attention to a statistical approach using genetic algorithms (GAs). The main concepts of this method are derived from biological evolution in a competitive environment (Holland, 1975). GAs are mathematical algorithms that transform populations of individual mathematical objects—typically fixed length genes—into new populations using operations patterned after genetic operations such as proportionate reproduction following the Darwinian theory of survival of the fittest. Due to the recent progress of GA theory, many industrial applications have been developed with the aid of this tool. These applications include gas pipeline design, process control, VLSI circuit layout, and so forth (Senders, 1994; Goldberg, 1989). From a survey of the literature, confirmed by our own experience, it appears that genetic programming is a very efficient, globally convergent, and robust optimization method (Senders, 1994). Two major characteristics contribute to this feature. First, GAs have a distinctive pattern in the search sample. Instead of a point-to-point search, the search sample is expanded to a group in which a search point moves among peaks. Consequently, the probability of finding a near-global optimal solution is increased. Second, the use of an objective function does not require derivatives (that is, sensitivity analysis) or other auxiliary knowledge. As a consequence, they are not restricted to the conditions of continuity, sensitivity, convexity, and nonlinearity of both the objective function and constraints. The different steps involved to define optimization parameters well, and the different criteria of the objective function are described in detail. The classic GAs and different new developments, such as the hill-climbing operator or ranking method —added to increase the GA performances—are described below.

The remainder of this article is organized in the following manner. First, the mathematical model of the drying process and the classic numerical methods to solve it are presented. Comments are provided on the limitations of this approach. In the second step, the ANN representation of a mathematical model is presented and applied to the dynamic drying process under consideration. Next, GAs are combined together with ANNs to optimize the design and process variables of the complete industrial dryer. Finally, the results we obtained and the global strategy is discussed.

Mathematical Model

Convective dryers are some of the most important equipment involved in the wood industry, and appear to be complex to design. A dryer-scale approach is implemented in order to write the classic differential/algebraic equations through parameters such as heat-transfer coefficients or drying kinetic, expressed as mass flow, Fm. The mathematical modeling of convective dryers requires the formulation of material and energy balances to describe the combined heat and mass-transfer phenomena. Taking the four characteristic variables into account, temperature (T) and moisture content (X) for both air (a) and product (p), the drying process is described by the following nonlinear system of equations (Eqs. 1 to 4) (detailed in Sébastian et al., 1993a).

• Humidity balance in section Δx is expressed by the following equation:

$$q_a \cdot \frac{\partial X_a}{\partial x} \cdot dx + dm_a \cdot \frac{\partial X_a}{\partial t} = - dm_p \cdot \frac{\partial X_p}{\partial t} \qquad (1)$$

• Heat balance is expressed as follows:

$$q_a \cdot C_a \cdot \frac{\partial T_a}{\partial x} \cdot dx + dm_a \cdot C_a \cdot \frac{\partial T_a}{\partial t} + dm_p \cdot C_p \cdot \frac{\partial T_p}{\partial t} = - H \cdot Fm \cdot L \cdot dx \qquad (2)$$

• Mass transfer is given by the following equation:

$$dm_p \cdot \frac{\partial X_p}{\partial t} = Fm \cdot L \cdot dx \qquad (3)$$

• Heat transfer is controlled by the heat-transfer coefficient, and is expressed by means of the following equation:

$$q_a \cdot C_a \cdot \frac{\partial T_a}{\partial x} \cdot dx + dm_a \cdot C_a \cdot \frac{\partial T_a}{\partial t} = h \cdot \left(T_a - T_p \right) \cdot L \cdot dx \qquad (4)$$

This system must be completed with the drying kinetic and the expression of the heat-transfer coefficient, both of which are specifics of the product and its configuration in the dryer. These equations are given later in the validation section.

This system is usually solved by traditional numerical schemes, such as finite differences methods.

Sébastian and colleagues (1993a) proposed a different approach to the numerical simulation of drying systems. To build a set of highly integrated algebraic equations, they combined modules in a transfer network. Links between the modules of transfer networks can be complex due to the complex circulation of liquids, solids, and gas in a drying system. Transfers between fluids and solids occur in modules of transfer unit (MTU), which have one inlet and one outlet for both air and product (Figure 1). The nonlinear system is then integrated using the number-of-transfer-unit (NTU) method. This approach requires us to analytically integrate transfer equations when it is possible. It usually requires much shorter computation times than with classic methods but, on the other hand, involves complex analytical computations. Usually used to design heat exchangers (Kandlilar and Shah, 1989), this method becomes unusable if the mathematical model is too complex.

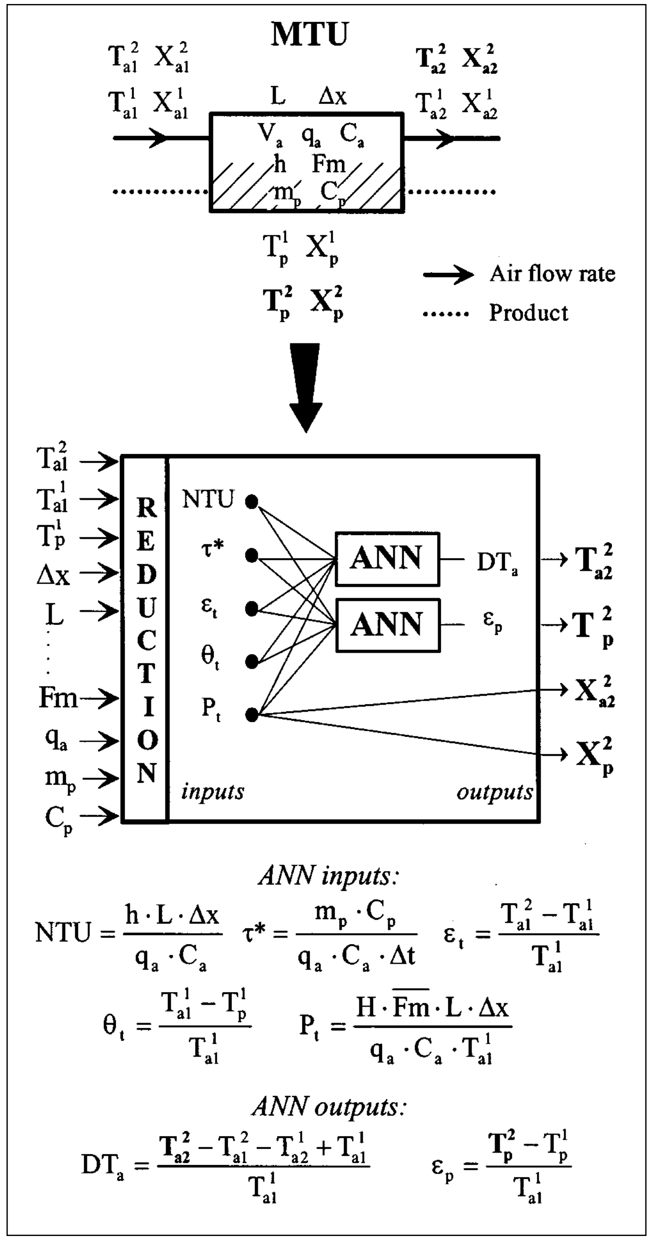

Figure 1. MTU and ANN representation of a model.

Table 1. Comparison of the Methods

Method	Δx (m)	Δt (s)	CPU Time (s)
Finite differences	[0.005; 0.1]	[0.1; 10]	[200; 2000]
NTU	[0.005; 0.1]	[0.1; 50]	[100; 2000]
ANN	[0.25; 1]	[500; 1000]	[5; 80]

All the details of this method and its applications are given in Sébastian et al. (1993b).

Heat and mass-transfer phenomena are indeed complex, and their solution requires extensive computation times even using the NTU method. During numerical computations, problems of stability and accuracy occur that require the use of very small integration variables ($\Delta x < 0.1$ m and $\Delta t < 50$ s; see Table 1). The classic numerical methods do not appear to be suitable for the optimal design problem.

ANN Representation of a Mathematical Model

The model involved in design operation must be fast and relatively accurate, but it also requires good generalization properties. In this light, Chen and Weigand (1994) proposed a very interesting method combining ANN models with a universal dynamic matrix control algorithm. This approach leads to very small computation times and good generalization properties for dynamic chemical processes. The approach proposed in this article is quite different and uses ANN as a representational tool of the mathematical model that greatly reduces the simulation time and maintains accuracy and the generalization properties of the initial model. The following five steps are involved in this approach, steps 3 and 4 being specific to using the ANN.

Step 1. Determination of the Necessary and Sufficient ANN Inputs/Outputs. When applying ANN to represent a drying model, one of the main problems is determining the inputs that are necessary and sufficient for the ANN to learn heat and mass-transfer laws. Figure 1 presents the 17 parameters and variables involved in the numerical description of the transfers within an MTU. Assuming that four learning points per input (which is the minimum) are sufficient for the ANN to learn the drying process, the learning data set includes 4^{17} events and the learning phase appears to be impossible to perform because of memory-requirement problems. It is hence necessary to decrease the input number, and this can be performed as follows.

- Sébastian et al. (1993b) proved that this set of input variables could be reduced by using the dimensionless characteristic numbers of both the dryer configuration and the drying phenomena. These numbers appear in dimensionless equations derived from Eqs. 1 to 4. A similar system appears in Hugget and Sébastian (1996). For instance, the NTU is significant for the intensity of the heat transfer in an MTU. Other problems may involve different dimensionless numbers, like Reynolds or Prandtl numbers, without modifying the methodology. Furthermore, system outputs are thermal efficiencies denoted by DT_a and ϵ_p in Figure 1. They are used to approximate the temperature change of both air and product. In this article, moisture content is calculated directly by using the dimensionless number P_t, which is characteristic of the water mass flow.
- In the second step, a numerical code based on the finite difference scheme is used to compute the system output for any given set of inputs. In order to reach enough accuracy and to avoid stability problems, MTU length Δx and interval time Δt are lower than 10^{-3} m and 10^{-1} s, respectively. The numerical method used in this code is not important, but it is necessary to obtain outputs of the system rather quickly.
- Using this code, a sensitivity study can be performed and different inputs, like T_{a2}^{1}, the output air temperature of an MTU considered at time t, are eliminated because they have very little effect on the ANN outputs. At the end of this step, we obtain a system with only five inputs and two outputs (Figure 1).

Step 2. Data Set Computing. A complex function approximation is mainly data driven, and the resulting functions are not believed to have any extrapolation properties. Therefore, data used for identification should cover the whole simulation domain in order to avoid any extrapolation. Moreover, dynamic process simulation requires great precision for interpolation and, as a consequence, a very efficient identification phase. When considering drying processes with multiple input/output, this requires a large amount of computational data, perhaps as many as 6000 events. Thus, computation of the data set required for the learning phase took more than two days of CPU time on an Ultra Sparc Sun station. This phase only occurs once in building an ANN.

Before the computation of the data set, each input point is determined, taking into account its influence on the output set. For example, NTU and P_t represent heat and mass-transfer intensity, respectively. Therefore, these inputs control process behavior and require numerous well-chosen discrete points (Table 2).

On the other hand, the cross-validation data set is also computed in order to assess the efficiency of the approximation from the learning set.

Step 3. ANN and Training Algorithm Characteristics. Using an ANN involves two different tasks: the determination of network parameters (structure, number of neurons, threshold function, etc.), and weights estimation.

Determination of the relevant network structure is a key issue. Indeed, the ANN structure governs the capability of the network to provide an adequate approximation of input–output relationships. Drying processes are very complex, and a four-layer feed-forward neural network (two hidden layers) appears to be suitable for learning such complex behaviors. All tests performed with one hidden layer and a large number of neurons failed. In order to reach a network topology that avoids overfitting phenomena, a practical approach is used (Wang et al., 1992; Z. Wang et al., 1994). This method optimizes the size of the output subspace of each hidden layer with the independent δ-linear criterion, and has

Table 2. ANN Inputs

Input	Input Interval	Number of Discrete Points
τ^*	[0.01; 12]	5
NTU	[0; 12]	8
P_t	[0; 0.25]	8
θ_t	[0; 0.2]	5
ϵ_t	[−0.15; 0.15]	4

been validated in a previous work (Hugget and Sébastian, 1996). The threshold functions of each hidden layer are tanh functions, and the output is a linear function. Each output is approximated by a different ANN in order to obtain a better accuracy.

The backpropagation algorithm appears to be inefficient in estimating the weights of the ANN, due to the number of input–output events in training data. The Levenberg–Marquardt method, an approximation of Newton's method, is used in this article, because it achieves better results from the training phase (Hagan and Menhaj, 1994). This optimization technique is more powerful than gradient descent ones, but requires more memory.

Step 4. Learning Phase. In order to increase the training efficiency, the learning algorithm uses several different sets of data points (Hugget and Sébastian, 1996). Each one is trained during five or ten cycles and allows an increase in the number of learning points with the same size as the training matrix. Furthermore, stopped training is used (Finof et al., 1993; C. Wang et al., 1994). Stopped training implies that we must monitor the performance of the network during the training phase. Training is stopped when the error on the cross-validation set reaches a minimum. Networks trained using stopped training appear to be very different from those that we obtain using convergent training and generally return better results. Finally, multiple training appears to be necessary because network parameter optimization is constrained by many local minimums. Varying initial conditions for network training can lead to different solutions (Kolen and Pollack, 1990). For each function to learn, the best solution of multiple training is kept. Training the ANN involved in the applications of this article generally requires two days of CPU time on a SUN Ultra Sparc station. Considering a more complex process, stacked ANN (Sridhar et al., 1996) could be used in order to increase the precision of the ANN representation.

After these different steps, we obtain a very good approximation of the two drying functions on and/or off the training points.

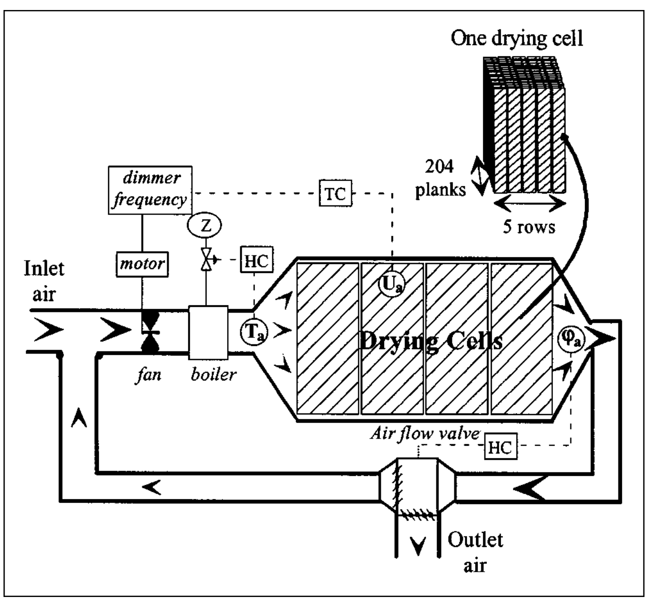

Figure 2. Dryer configuration.

Step 5. Validating ANN Representation of a Model. The validation step is performed by comparing ANN and classic differential model results for the entire industrial dryer under consideration. ANN representation of a mathematical model is used just to represent the drying process. The proposed methodology was applied to the drying of poplar planks (600 $\times 50 \times 5$ mm) for tray manufacture, with an initial material moisture content level of 1.8 kg/kg dB down to 0.2 kg/kg dB. The input values of air temperature and relativity humidity are, respectively, 293 K and 60%. This type of dryer includes drying cells placed in series, one or several fans for air circulation throughout the product, recirculation equipment for outlet air, and a heating device such as a gas boiler. A diagram of the dryer is presented in Figure 2. The final control elements are the gas flow valve, the recirculation air flow valve, and the dimmer frequency that regulates the boiler combustion, the flow rate of recirculation, and the fan, respectively. In practical use, air flow is often reversed in order to homogenize the moisture content of the product along drying cells.

The drying kinetic for this product, given by Nadeau and Puiggali (1995), is expressed by means of the mass flow, Fm, given by the following equation:

$$Fm = Fmis \cdot \left(\frac{X_p}{Xcr} \right)^{\beta} \tag{5}$$

where

$$Xcr = 1.28 \cdot 10^{-2} \cdot T_a - 3.36$$

and

$$Fmis = 4.5 \cdot 10^{-4} \cdot \left(1 - \frac{2}{1 + e^{-20 \cdot (\varphi_a - 1)}} \right) \cdot V_a^{0.7} \cdot e^{-1250/T_a}$$

where β is the slackness coefficient and is equal to 0.9 if $X_p < X_{cr}^a$, and is equal to 0 if not.

The heat-transfer coefficient is given as follows:

$$h = \frac{Nu \cdot \lambda_a}{Dh} \tag{6}$$

where $Nu = f(Re, Pr)$ is the Nusselt number, D_h is the hydraulic diameter, and λ_a, the air conductivity.

One of the main problems occurring during the simulation of a drying sequence is the interaction between fan and product. Air velocity is mainly influenced by both the moisture content of the product and the number of drying cells due to the pressure drop involved by air flow within the product. During drying, the product moisture content decreases, yielding product retraction and a decrease in the pressure drop in the wood stack. Thus, the air velocity, which controls the drying kinetic with air temperature, greatly depends on the fan behavior. This point is extremely important and controls the global dryer behavior. In this article, 21 centrifugal fans were tested for the optimal design problem. The behavior of the fans and their electrical consumption, which also depends on the pressure drop, are given by specific ANNs trained as in black-box approaches on the basis of experimental data that

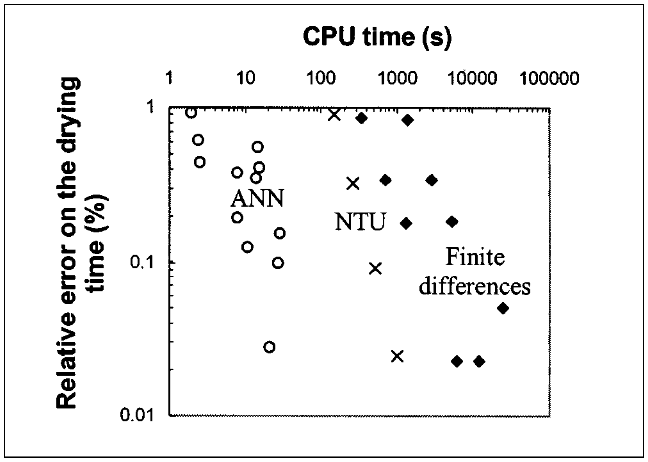

Figure 3. Comparison of the different methods.

Fan = CRTT 1000; $\eta_{fan} = 1250$ tr/min; 8 drying cells; $T_{a1} = 353$ K.

are themselves provided by the industry. These complementary black-box models compute the air flow within the product used to define some inputs of the ANN representation of the drying model. Furthermore, the air recirculation is regulated by means of a classic PID controller, depending on the relative humidity of the output air.

A comparison of our ANN approach is made in Figure 3 with a classic differential model involving a finite difference scheme or the NTU method. The comparison is based on a reference solution given by a mathematical model using the finite difference method with very small time steps and control volumes. The global dryer being modeled involves a specific fan (CRTT 1000) and eight drying cells. Simulations reveal the efficiency of the ANN drying representation. The main advantages are the following:

- ANN greatly reduces the computation time compared to differential models that use classic methods. This is because of the ability of the ANN to highly integrate nonlinear phenomena and thus to use a very large discrete interval ($\Delta x > 0.25$ m and $\Delta t > 500$ s) whatever the transfer intensity is (Table 1).
- The simulation results show that the proposed approach has a good accuracy compared with other methods. This is because the data series are accurately computed with a very small discrete interval ($\Delta x < 0.001$ m and $\Delta t < 0.1$ s). The efficiency of the learning phase is also necessary to achieve such a precision.
- ANN, as defined earlier, can be easily extended to a wide range of products and equipment, and appears as a very good tool for the optimal design problem. The problem identification is included in the determination of the dimensionless numbers, which are computed before the ANN representation is used. For instance, if the wood type or plank size changes, only the kinetics and heat transfer coefficients have to be changed, and these coefficients are only used to compute ANN inputs.
- ANNs can accurately learn a dynamic process, even when the conventional models are not available. In this study, the mathematical model is used to compute the learning data set, but the same methodology could be applied with experimental data. In this case, this one needs a large amount of experimental data.

Design and Operation Optimization of a Convective Industrial Dryer

The speed and accuracy of the simulation code using ANN make it reasonable to consider solving the associated optimal design problem. The first design step consists of finding both optimization parameters and an objective function suitable for an industrial need. In this field, using a multicriteria function (as in Zbicinski, 1992) appears to be the best approach. The second step involves the optimal design determination using GA.

Step 1. Optimization Parameters. The eight optimization parameters are given in Table 3. Dryer equipment and process variables are optimized simultaneously in order to obtain dryer configurations that are best suited to the problem. For instance, if one of the dryer elements, such as a fan, is very well suited, but does not operate at the correct speed, the entire dryer will return poor results and will not be selected.

The first four parameters represent different devices, such as heating utilities, that can use electricity, gas, or sawdust; two types of fans with different diameters are available (21 fans in all); and the number of drying cells of the dryer. In practice, the air flow rate throughout the product is commonly controlled by means of the fan dimmer frequency to maintain V_a constant in the wood stack, but this control involves supplementary investment costs. Therefore, we also consider the dryer without this control. The reason for simultaneously considering all these equipment parameters is be-

Table 3. Design and Operating Parameters

	Parameters	Range of Changes	Number of Choice
1	Heating device	1. Electrical resistance 2. Gas boiler 3. Sawdust boiler	3
2	Fan type	$1 \rightarrow 15$ model VCRAT $16 \rightarrow 21$ model CRTT	21
3	Number of drying cells N_c	[1; 20]	20
4	Regulation type Reg_t	1. V_a is constant (η_{fan} vary) 2. η_{fan} is constant (V_a vary)	2
5	Air temperature T_a (K)	[293; 353]	21
6	Outlet-air relative humidity φ_a (%)	[20; 100]	9
7	Inversion time for air flow rate t_{inv} (s)	[0; 20000]	8
8	Fan speed E_f	[0; 100]	11

cause of their complex interactions, for instance, between fan type, number of drying cells, and air flow control. The heating device is also very important because we will design the dryer using different imposed productivities.

Considering process variables, poplar planks are not to be dried at temperatures exceeding 353 K for thermal degradation reasons. The relative humidity of the outlet air is controlled by air recirculation, while the inversion time defines the number of inversions occurring during a drying sequence. The last parameter is characteristic of the fan speed (0, corresponding to the minimal speed of the fan under consideration, and 100, to the maximum one). The last four parameters are defined on a continuous interval and are discretized in order to simplify the problem. We thus give a number of discrete choices for each one. Fine discretization leads to a tougher optimization, but the optimization solution is more accurate.

Due to this discretization, the optimization becomes a combinatorial problem, with, in this case, 42 million solutions. Some optimization parameters, such as fan or heating device, which are not numerical parameters, prohibit the use of many optimization methods (that is, quasi-Newton or conjugate gradient algorithms). In this field, a statistical approach such as GA appears very efficient. These algorithms are not restricted to the conditions of continuity, sensitivity, convexity, and nonlinearity of both objective function and constraints.

Individuals involved in the GA search have the eight genes defined by the optimization parameters described earlier. An example of an individual would be the following:

1	14	21	1	347	92	2000	90

and represents a dryer configuration with an electrical-resistance fan that is a 900-mm-diameter VCRAT, has 12 drying cells, and so on.

Step 2. Individual Evaluation. In order to apply GA, an evaluation value is computed for each individual according to its fitness. For this design problem, this evaluation involves three successive phases:

- Determination of local characteristics for one drying sequence;
- Calculation of various economic criteria;
- Determination of the objective function.

First, the equipment and information about process variables included in individual genes are tested by simulating one drying sequence using ANN. This simulation returns such local characteristics of the configuration as the drying time, the product moisture content along drying cells, and the different energy costs (kWh) for fans or heating devices. Results are obtained with good accuracy, taking into account nonlinear temporal variation of the behavior of each element and their interactions.

Second, these local characteristics of the dryer and the gene information are then used to obtain the following economic criteria.

Computation of the product capacity P (cells/day) requires the number of possible drying sequences in one production day. In this day, the drying system was operating from 8 a.m. to 6 p.m. If the drying time is less than 9 hours, there can be a drying sequence during the night under the control of one person.

The total annual cost T_c includes operational and financial costs. The operational cost of the plant consists of the costs of the thermal and electrical energy consumed by heat exchangers and fans; the cost of the human labor required to move drying cells with a fork-lift truck or to control the drying process; maintenance of the heating device; and the cost involved in using the fork-lift truck (Table 4). Financial costs are very specific from one case to another. One configuration available for such a drying process (Poitrat, 1991) is given in Table 4.

The investment cost I is affected by the number of drying cells, the heating device capacity, the installed power of fans, and the cost of the control equipment for air inversion, recirculation, and flow rate. Human labor needed to install the different equipment is included in this cost. It is necessary to consider I separately from T_c for later control of the importance of each parameter in the objective function.

The final product quality requires very homogeneous product moisture content along drying cells. In fact, these wood planks are used to manufacture crates. The heterogeneity of moisture content along the drying cells generates a difference in crate weight, and finally, an error in the product weight measured in the crate. In order to optimize the market value of the poplar planks, this design problem requires that the difference in moisture content along the drying cells be minimized. For this purpose, a quality function Q is defined as follows. The maximum accepted difference in moisture content between planks is 0.02. Hence, the quality function does not remain 0.02 and involves two linear functions on both sides of this point (Figure 4a). Slopes of these functions will be useful during the GA optimization to generate dryer configurations that produce the best product quality. Notice that the product quality cannot be considered as a constraint in

Table 4. Operational and Financial Costs

Operational Costs							
Human Labor	Energy			Maintenance			Fork-Lift Truck
	Gas	Sawdust	Electricity	Gas	Sawdust	Electricity	
11.5 \$/h	1.7 ¢/kWh	0.5 ¢/kWh	8.6 ¢/kWh	0.55 I/yr	0.75 I/yr	0.4 I/yr	13.8 \$/h

Financial Costs				
Taxes	Insurance		Depreciation	Loans Repayment
	Electric resistance	Boiler		
0.03 R/yr	(0.001 I + 0.002 R/yr)	(0.002 I + 0.002 R)/yr	0.1 I/yr	0.35 I/yr

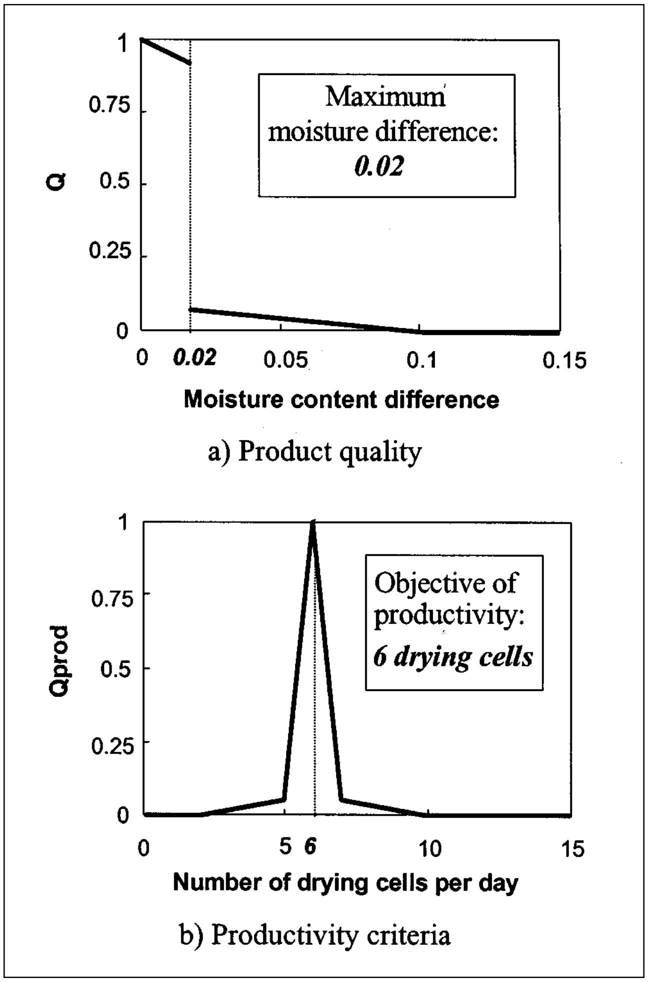

Figure 4. Quality and productivity criteria.

this problem, since we have to optimize it. More to the point, this quality function can be changed (such as in taking into account the final shape of wood planks or cracks). This kind of modification involves other works, such as a mechanical study of stresses and deformations, and is not necessary for understanding the methodology developed in this article.

In order to design a drying system for one prescribed product, another function Q_{prod} is considered. This function is equal to one if the production capacity P of the dryer configuration is equal to the imposed one. Otherwise, Q_{prod} has smaller values (Figure 4b). Just like the function Q, Q_{prod} is useful during the GA search to find dryer configurations with a production capacity equal to the prescribed one.

Finally, the whole objective function is determined using these economical criteria. This function is a formulation of the criterion vector (quality measurement) with the introduction of an ordering relation between each component of this vector. In practice, it is impossible to choose or even recommend a universal method, and in this study, the concept of maximum efficiency (Kaminsky et al., 1989), given as follow, is used:

$$\max_{x \in D} [F(x)] = \prod_{k=1}^{N} f_k(x) \tag{7}$$

where f_k represents different economic criteria and D is the solution space.

Such a multicriteria procedure requires *a priori* information on both these criteria f_k and their hierarchy of importance (Vanderpooten and Vinckle, 1991), but appears well suited to numerous optimization problems.

Step 3. Genetic Optimization. At this point, GA can be used because this method only requires the objective function value for each tested dryer configuration. There are three operators in a simple GA: selection/reproduction, crossover, and mutation. In this article, a hill-climbing operator is added to increase GA performance, as in Senders (1994). The optimization process involves classic steps of GA search, as well as more specific ones (Figure 5). This process can be summarized as follows:

Step 1. First, GAs create a population of several individuals of the same system with parameters (or genes) assigned arbitrarily or randomly.

Step 2. The different individuals undergo a set of trials in order to compare their performances (see the preceding section). In this way, each individual i receives an evaluation value F_i according to its fitness.

Step 3. Selection/reproduction can then occur. Each individual is duplicated according to its objective function values. The values can be thought of as some measure of a maximized profit, utility, or goodness, such as, minimum weight. In that case, "copying individual according to their fitness values" means that individuals with a lower value have a higher probability of contributing one or more offspring in the next generation. The selection/reproduction operator can be implemented in algorithmic form in a number of ways. The easiest way is to create a biased roulette wheel, where each individual currently in the population has a roulette-wheel slot sized according to its fitness. In this article, the ranking method (Baker, 1985) is used to determine slot size both for its robustness and its particular handling of rapid convergence. Each time a new offspring is required, a simple spin of the weighted roulette wheel yields the reproduction candidate. In this way, the more fit individuals have a higher number of offspring in the succeeding generation, and an exact replica of the individual is made. This individual is then

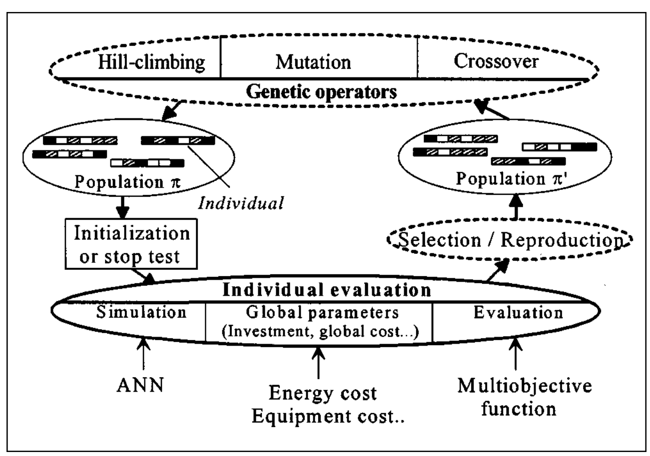

Figure 5. Dryer optimization algorithm.

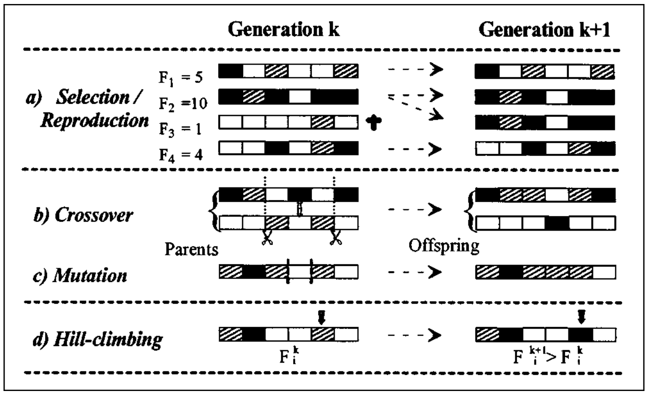

Figure 6. Example of genetic algorithm operators.

entered into a mapping pool, an intermediate new population (π'), for further genetic operation action (Figure 6a).

Step 4. Following the selection/reproduction, crossover, mutation, and hill-climbing operators are successively used with probabilities p_c, p_m, and p_h, respectively.

Crossover randomly mates part of the parents strings and generates pairs of offspring to replace the parents. Crossover is the most important search operator and may proceed in two steps. In the first step, it randomly pairs members of the newly reproduced individuals in the mating pool with uniform distribution, while in the second step, genes in each pair of individuals are randomly interchanged (Figure 6b).

Mutation plays a secondary role in the operation of traditional GA: it prevents the occasional loss of potentially useful genetic material from reproduction and crossover. The function of the mutation operator is the occasional random alteration of the value of a gene (Figure 6c).

Hill-climbing allows a significant increase in the algorithm performance. This operator is a point-to-point search that generates a better offspring when it is possible (Figure 6d). In this study, hill-climbing is particularly efficient, because it uses information given by the functions Q and Q_{prod}.

Different methods exist to adapt operator probabilities in GA (Davis, 1989; Spears, 1991). In this article, operators are used with the following probabilities $p_c = 0.6$, $p_m = 0.05$, and $p_h = 0.5$ in order to favor the crossover search.

Step 5. Once the first trial is completed, iterations will not stop until the convergence criterion is satisfied.

Two different cases are studied as examples of applications. The first one involves an optimization without productivity constraint and uses the relative influence of the two cost criteria. As a second case, we choose to design a drying system for a different prescribed productivity. In all the optimizations, the population is composed of 50 individuals. The GA generates 200 generations and then stops.

Application 1. Optimization Without Imposed Productivity. In the first example, we want to optimize both the product quality, Q, and the production capacity, P, and, at the same time, we also want to minimize the entire cost, including the total annual cost, T_c, and the total investment cost, I. The objective function is defined as follow:

$$F = Q \cdot P \cdot \left(\frac{1}{a_1 \cdot T_c + a_2 \cdot I} \right) \tag{8}$$

where a_1 and a_2 are weights, and $a_1 + a_2 = 1$. These weights are used to change the relative importance of T_c and I in the total cost criteria.

Considering this objective function, the optimum reached will depend on the relative contribution of each component. The product quality function, Q, does not pose any problem because it has been defined in order to control its influence: Q returns a value nearest to 1 if the difference in moisture content is less than 0.02, and a value nearest to 0 if the difference in moisture content is greater than 0.02. Furthermore, we cannot control the numerical contribution of the global cost and the production capacity criteria. After a quick study, we found that both criteria made approximately the same contribution. Therefore, the objective function was defined by multiplying these criteria. We can note that the use of weighted components in the objective function could have yielded to another solution, highlighting the fact that the definition of well-suited weights is then a very difficult task.

Results of the optimization procedure for this objective function are given in Table 5. Five cases were studied, varying a_1 and a_2. We note that quite large differences appear in the optimal dryer configuration due to the cost criteria. When we consider only the total annual cost, energy consumption is preponderant. The optimization procedure returns a dryer configuration with a relatively small number of drying cells and a fan that does not use its maximum power and the control of V_a in the wood stack appears to be necessary. If the

Table 5. Dryer Optimization Without Imposed Productivity

	Optimization Variables								Economical Criteria			
	Heating Device	Fan	N_c	Reg_t	T_a (K)	φ_a (%)	t_{inv} (s)	E_f	Q	P (cells/day)	T_c (k$/yr)	I (k$)
Case 1 ($a_1 = 1$, $a_2 = 0$)	Sawdust boiler	VCRAT 450	10	0	353	90	11,000	85	0.992	10	921	746
Case 2 ($a_1 = 0.75$, $a_2 = 0.25$)	Sawdust boiler	VCRAT 400	13	0	353	90	14,000	100	0.972	13	1,204	785
Case 3 ($a_1 = 0.5$, $a_2 = 0.5$)	Sawdust boiler	VCRAT 560	8	0	350	90	11,000	100	0.995	16	1,574	739
Case 4 ($a_1 = 0.25$, $a_2 = 0.75$)	Gas boiler	VCRAT 560	8	0	347	90	3,000	100	0.986	16	2,137	381
Case 5 ($a_1 = 0$, $a_2 = 1$)	Electrical resistance	VCRAT 560	9	0	353	90	10,000	100	0.995	18	3,692	218

Table 6. Dryer Optimization with Imposed Productivity

Objective of Productivity (cells/day)	Optimization Variables								Economical Criteria			
	Heating Device	Fan	N_c	Reg_t	T_a (K)	φ_a (%)	t_{inv} (s)	E_f	Q	Q_{prod}	T_c (k$/yr)	I (k$)
4	Gas boiler	VCRAT 500	4	0	353	80	8,000	65	0.994	1	665	282
6	Gas boiler	VCRAT 500	6	0	353	70	5,000	95	0.994	1	868	328
8	Gas boiler	VCRAT 450	8	1	353	80	11,000	60	0.992	1	1,084	361
10	Sawdust boiler	VCRAT 450	10	1	353	90	8,000	60	0.997	1	932	735
12	Sawdust boiler	VCRAT 400	12	0	353	90	9,000	80	0.977	1	1,120	766
14	Sawdust boiler	VCRAT 400	14	0	353	90	13,000	100	0.980	1	1,336	812
16	Sawdust boiler	VCRAT 560	8	0	350	90	11,000	100	0.995	1	1,574	739
18	Sawdust boiler	VCRAT 560	9	0	353	90	10,000	100	0.963	1	1,862	761

investment becomes as important as or more important than the total annual cost—cases 3, 4, and 5—the dryer involves only 8 or 9 drying cells, but this configuration dries more quickly and two sequences per day are possible. Thus, energy consumption is very important, but production capacity is much more so. When the investment cost is preponderant, there is no special control to keep V_a constant, and the optimal heating device is shown to be a gas boiler or an electrical resistance, which are less expensive. GA always proposes the VCRAT fan type with diameters ranging from 400 to 560 mm. This is due to their ability to perform better when a high reduction in the pressure occurs in the wood stack. The use of a 560-mm diameter fan can provide a high air velocity when the number of drying cells is small enough. The drying process is then very fast and allows two sequences of drying per day. When energy consumption is not preponderant, these fans are always running at their maximum power capacity. The control of relative humidity for the outlet dryer air is always 90%, which allows both a good mass flow and a decrease in energy consumption by the heating device. In almost all cases, the optimal air temperature reaches its upper bound, where it increases the mass flow. If air recirculation becomes impossible for technological reasons, this optimal air temperature will be quite different.

Application 2. Optimization with Prescribed Productivity. The second example deals with the dryer optimization with a different prescribed productivity. The objective function is then the following:

$$F = Q \cdot Q_{prod} \cdot \left(\frac{1}{0.5 \cdot T_c + 0.5 \cdot I} \right) \tag{9}$$

Considering this example, the influence of each component of the objective function is controlled by means of Q, Q_{prod}, and weights a_1 and a_2. The second application allows us to understand the influence of productivity on the optimal dryer configuration, but is less realistic than the preceding one. The results of the optimization are given in Table 6 and in Figure 7. The productivity objective varies from four drying cells per day to 18.

The imposed productivity has a great influence on the optimal dryer configuration. We note that the optimal heating device directly depends on this productivity with the gas boiler and the sawdust boiler for an average and maximum productivity, respectively. Furthermore, average productivity involves fans that do not run at their maximum power.

In Figure 7 we can see the slope breaks for the investment cost and the total annual cost curves. For a required productivity larger than 8 cells per day, the optimal heating device becomes a sawdust boiler that involves both less energy and higher investment costs. When the productivity demand is larger than 14 drying cells per day, the optimal configuration does not involve more drying cells, but allows two drying sequences per day and requires large energy consumption. For more than 20 drying cells per days, one dryer is not sufficient.

On the average, GA evaluates 3000 individuals to find several near-global optimal solutions. Each drying simulation depends on the number of drying cells and requires CPU times of more than 20 seconds. The required CPU times to

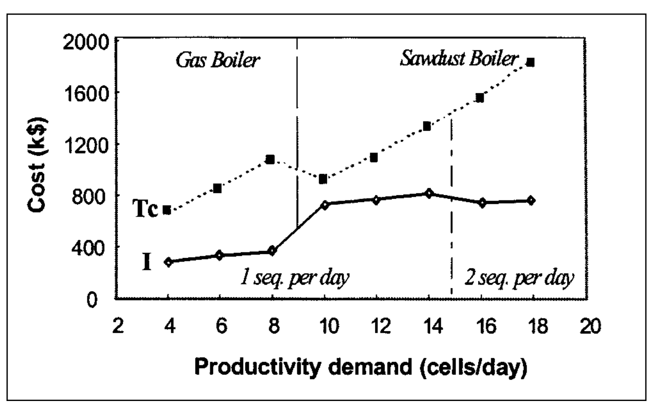

Figure 7. Optimization with a productivity constraint.

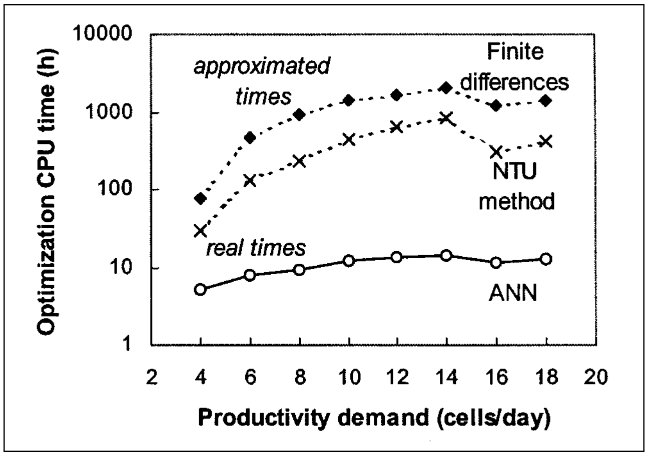

Figure 8. CPU times required for the optimizations.

determine the solution are given in Figure 8. They vary from 5 hours to 20 hours. The same optimization approach does not appear to be feasible if classic modeling is used.

Conclusions

Optimal design for complex systems is a universal problem. The principal aim of this article was to propose, in an example problem, a way to solve such a complexity. A novel approach involving ANN representation of a mathematical model and stochastic optimization was presented and validated for an industrial convective dryer.

Considering the representation of the behavior of this convective dryer, the methodology proposed in this work is quite different compared to classic ones. It uses ANN as a representative tool for the mathematical model. This approach significantly reduces simulation time compared to differential models, and maintains good accuracy and generalization properties. Using an ANN representation, a drying sequence of 20 hours is simulated in less than 20 seconds CPU time.

The associated optimal design problem was then studied. This optimization is a complex combinatorial problem with 42 million potential solutions and a multiobjective function that is not continuous, not differentiable, and not explicit—it is computed by a program. In the cases considered in this work, a stochastic method such as genetic algorithms appeared to be well suited to this kind of problem. A few excellent solutions for different problems were obtained rather quickly, that is, in less than 20 hours.

The methodology proposed in this work was applied to represent and optimize a complex system. It is rapid, accurate, and maintains general properties. The descriptions provided in this article are intended to provide enough detail for possible use of the technique to solve many other kinds of problems.

Notation

C = heat capacity, J/kg·K
F_{mis} = isenthalpic mass flow, kg/m^2·s
h = heat-transfer coefficient, W/kg·k
H = latent heat of vaporization of water, J/kg
L = contact length between air and product in a drying cell, m
m = mass, kg
p_c = crossover probability
p_h = hill-climbing probability
p_m = mutation probability
q = flow rate, kg/s
R = annual revenues, \$/yr
T = temperature, K
V = velocity, m/s
X = moisture content, kg/kg db
X_{cr} = critical moisture content, kg/kg db

Greek letters

ϵ_t = ANN input characteristic of the air temperature evolution (x) between instants t and $t+\Delta t$
η_{fan} = fan velocity, tr/min
θ_t = ANN input characteristic of the temperature difference between air and product
τ^* = ANN input characteristic of the MTU thermal capacity

Superscripts

1 = instant t
2 = instant $t+\Delta t$

Subscripts

1 = x
2 = $x+\Delta x$

Literature Cited

Baker, J. E., "Adaptive Selection Methods for Genetic Algorithms," *Proc. Int. Conf. on Genetic Algorithms*, p. 101 (1985).

Barron, A. R., "Universal Approximation Bounds for Superpositions of a Sigmoidal Function," *IEEE Trans. Inform. Theory*, **IT-39**, 930 (1993).

Bhat, N., and T. J. McAvoy, "Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems," *Comput. Chem. Eng.*, **14**(4/5), 573 (1990).

Chen, Q., and W. A. Weigand, "Dynamic Optimization of Nonlinear Processes by Combining Neural Net Model with UDMC," *AIChE J.*, **40**(9), 1488 (1994).

Cubillos, F. A., P. I. Alvarez, J. C. Pinto, and E. L. Lima, "Hybrid-Neural Modeling for Particulate Solid Drying Processes," *Powder Technol.*, **87**, 153 (1996).

Cybengo, G., "Approximating by Superposition of Sigmoidal Function," *Math. Control Signals Syst.*, **2**, 303 (1989).

Davis, L., "Adapting Operator Probabilities in Genetic Algorithms," *Proc. Int. Conf. on Genetic Algorithms*, J. D. Schaffer, ed., p. 61 (1989).

Finof, W., H. Hergerdt, and H. G. Zimmerman, "Improving Model Selection by Nonconvergent Methods," *Neural Networks*, **6**, 771 (1993).

Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, New York (1989).

Hagan, M. T., and M. B. Menhaj, "Training Feedforward Networks with Marquardt Algorithm," *IEEE Trans. Neural Networks*, **5**(6), 989 (1994).

Heyd, B., G. Rodriguez, and J. Vasseur, "Modelling of Drum Drying Using Neural Network and Polynomial Model," *Proc. Int. Drying Symp.*, Vol. A, p. 323 (1996).

Holland, J. H., *Adaptation in Natural and Artificial Systems*, Univ. of Michigan Press, Ann Arbor (1975).

Hornick, K., M. Stinchcombe, and H. White, "Multilayer Feedforward Network are Universal Approximators," *Neural Networks*, **2**(5), 359 (1989).

Hornick, K., M. Stinchcombe, H. White, and P. Auer, "Degree of Approximation Results for Feedforward Networks Approximating Unknown Mapping and Their Derivatives," *Neural Comput.*, **6**, 1262 (1994).

Huang, B., and A. S. Mujumdar, "Use of Neural Network to Predict Industrial Dryer Performance," *Drying Technol.*, **11**(3), 525 (1993).

Hugget, A., and P. Sébastian, "Modélisation de Séchoirs à Tapis:

Utilisation des Réseaux de Neurones," *Rev. Gén. Therm.*, **35**, 599 (1996).

Jay, S., and T. N. Oliver, "Modelling and Control of Drying Processes Using Neural Networks," *Proc. Int. Drying Symp.*, Vol. B, p. 1393 (1996).

Kaminski, W., I. Zbicinski, S. Grabowski, and C. Strumillo, "Multiobjective Optimization of Drying Processes," *Drying Technol.*, **7**, 1 (1989).

Kandlilar, S. G., and R. K. Shah, "Asymptotic Effectiveness—NTU Formulas for Multipass Plate Heat Exchangers," *J. Heat Transfer*, **111**, 314 (1989).

Kiranoudis, C. T., Z. B. Maroulis, and D. Marinos-Kouris, "Design and Operation of Convective Industrial Dryers," *AIChE J.*, **42**(11), 3030 (1996a).

Kiranoudis, C. T., Z. B. Maroulis, D. Marinos-Kouris, and G. D. Saravacos, "Comparison of Various Types of Industrial Dryers: Design and Operation," *Proc. Int. Drying Symp.*, Vol. B, p. 463 (1996b).

Kisakurek, B., "Retention Time in a Rotary Dryer," *Proc. Int. Drying Symp.*, J. C. Ashworth, ed., Woovelhampton, England, p. 148 (1982).

Kolen, J. F., and J. B. Pollack, "Backpropagation is Sensitive to Initial Conditions," *Complex Syst.*, **4**, 269 (1990).

Nadeau, J. P., and J. R. Puiggali, *Séchage: des Processus Physiques aux Procédés Industriels*, TEC&DOC, Paris (1995).

Poitrat, E., "Des Procédés Performants de Séchage du Bois pour Accroître la Rentabilité et Améliorer la Qualité des Produits," *Proc. Int. Conf. EDG*, CTBA, Paris (1991).

Psichogios, D. C., and L. H. Ungar, "A Hybrid Neural Network—First Principle Approach to Process Modelling," *AIChE J.*, **38**(10), 1499 (1992).

Sébastian, P., J. P. Nadeau, and J. R. Puiggali, "Les Réseaux de Modules d'Unités de Transferts, Thermiques, Massiques, Outils Pertinents d'Analyse de Séchoirs," *Int. J. Heat Mass Transfer*, **36**, 1763 (1993a).

Sébastian, P., J. P. Nadeau, and J. R. Puiggali, "Les Nombres d'Unités de Transferts, Thermiques, Massiques, Outils Rapides de Simulations de Séchoirs," *Int. J. Heat Mass Transfer*, **36**, 1773 (1993b).

Senders, J. M., *Algorithmes Génétiques et Réseaux de Neurones*, Hermès, Paris (1994).

Spears, W. M., "Adapting Crossover in a Genetic Algorithm," *Proc. Int. Conf. on Genetic Algorithms*, R. K. Belew and L. B. Booker, eds., Morgan Kaufmann, p. 138 (1991).

Sridhar, D. V., R. C. Seagrave, and E. B. Bartlett, "Process Modelling Using Stacked Neural Networks," *AIChE J.*, **42**(9), 2529 (1996).

Thompson, M. L., and M. A. Kramer, "Modeling Chemical Processes Using Prior Knowledge and Neural Networks," *AIChE J.*, **40**(8), 1328 (1994).

Van Can, H. J. L., C. Hellinga, K. C. A. M. Luyben, J. J. Heifjnen, and H. A. B. te Braake, "Strategy for Dynamic Process Modeling Based on Neural Networks in Macroscopic Balances," *AIChE J.*, **42**(12), 3403 (1996).

Vanderpooten, D., and P. Vinckle, "Description and Analysis of Some Representative Multicriteria Procedure," *J. Oper. Res. Soc.*, **42**, 3 (1991).

Wang, C., S. S. Venkatesh, and J. S. Judd, "Optimal Stopping and Effective Machine Complexity in Learning," *Adv. Neural Inf. Process. Sys.*, **6**, 303 (1994).

Wang, Z., M. T. Tham, and A. J. Morris, "Multilayer Feedforward Neural Networks: A Canonical Form Approximation of Nonlinearity," *Int. J. Control*, **56**(3), 655 (1992).

Wang, Z., C. Di Massimo, M. T. Tham, and A. J. Morris, "A Procedure for Determining the Topology of Multilayer Feedforward Networks," *Neural Networks*, **7**(2), 291 (1994).

Widrow, B., D. E. Rumelhart, and M. A. Lehr, "Neural Networks: Applications in Industry, Business and Science," *Commun. ACM*, **37**, 93 (1994).

Zbicinski, I., "Application of Multiobjective Optimization to Spray Dry," *Drying '92*, A. S. Mujumdar, ed., p. 515 (1992).

Zbicinski, I., W. Kaminski, K. Ciesielski, and P. Strumillo, "Dynamic and Hybrid Neural Model of Thermal Drying in a Fluidized Bed," *Proc. Int. Drying Symp.*, Vol. B, p. 1433 (1996a).

Zbicinski, I., P. Strumillo, and W. Kaminski, "Hybrid Model of Thermal Drying in a Fluidized Bed," *Comput. Chem. Eng.*, **20** (Suppl.), S695 (1996b).

Manuscript received May 4, 1998, and revision received Feb. 21, 1999.