How to increase the performance of models for process optimization and control

R. Simutis, R. Oliveira, M. Manikowski, S. Feyo de Azevedo [1], A. Lübbert *

Institut für Bioverfahrenstechnik und Reaktionstechnik, Martin-Luther-Universität, Halle-Wittenberg, D-06099 Halle, Germany

Abstract

Some key aspects of obtaining hybrid process models which perform well and that can be used in process supervision, optimization and control are discussed from the point of view of the benefit/cost-ratio. The importance of starting with a clear definition of the problem and a corresponding quantitative objective function is shown. In order to enhance the benefit/cost-ratio above the threshold of acceptance, a series of procedures is proposed: in the beginning an exploratory process data analysis is suggested to classify the process variables according to their importance and to facilitate the development of black- and grey-box models. Efficient validation of the model is shown to be indispensable. Hybrid model approaches proved to have to significant advantages, since they allow the activation of a larger portion of the available a-priori knowledge. Applications of hybrid models with respect to process optimization require new techniques, since the classical approaches are too difficult to use and are restricted to well-performing models. Finally, powerful software tools are required to implement the different algorithms at the production plants and to allow the efficient conversion of the ideas to real benefits. © 1997 Elsevier Science B.V.

Keywords: Performance; Process optimization; Process control; Hybrid process models

1. Introduction

As biotechnology develops and extends all over the world, competition between different companies grows. The general consequence to be noticed is that companies are taking a closer look at the economics of their production processes.

From the biochemical engineering point of view, the straightforward way of improving the economics is to invest in process optimization and control.

In industrial practice, improvements are often achieved by more- or less-educated trial and error methods, i.e. by empirical methods guided by intuition and experience. In this way, luck is also an essential condition of success.

Others prefer systematic approaches, which, on average, lead to a faster rate of improvements.

---

* Corresponding author.
[1] Current address: Department of Chemical Engineering, University of Porto.

The systematic approach is characterized by a consequent utilization of the a-priori knowledge about the process. With this approach the number of experiments required can usually be kept low. Thus, we expect a significant reduction of expenses.

Most often, mathematical process models are considered to represent the a-priori knowledge. When we speak about models in biochemical engineering and, in particular, with respect to process optimization and control, we think of relationships which describe the basic aspects of the real process that are most important to process performance. One main demand is that process models allow the representation of the relevant process properties quantitatively and therfore make them accessible to efficient computational analysis and optimization techniques. In biotechnology, classical mathematical models are constructed using some mechanistic concepts, which are embedded into mass balances, i.e. a relatively simple set of differential equations (Sonnleitner and Käppeli, 1986). Recently, some alternative methods of representing bioprocesses were discussed in literature, in particular, artificial neural networks and fuzzy expert systems (Lübbert and Simutis, 1994; Thomson and Kramer, 1994) are used to extend the classical approaches.

In order to solve the task of improving the process performance, it is indispensable to define beforehand how to measure progress. What we need is a quantitative criterion formulated by means of an objective function. Such an objective function may be a rather complicated function, since it should not only consider the main objective, e.g. the optimization of the volumetric productivity with respect to the main product, but also the practical boundary conditions which must be met during production.

Once we have such an objective function for the process performance, we are also able to define what we mean by model performance. The performance of a model is simply the advantage it brings in improving the process performance with respect to the objective function.

Evidently, the number of industrial production reactors, where model-supported process supervision, optimization and control is applied, is very low. There are a number of reasons. The most significant ones arise in connection with the problem of activating enough knowledge about the process under consideration. Moreover, too complicated modeling and optimization procedures led to high development times. Therefore, in most practical attempts made so far the benefit/cost-ratio appeared to be too low. Essentially, both factors—the cost as well as the benefit—did not justify model-supported optimization and control (Royce, 1993).

In this paper we address the problem of systematically improving process models from the point of view of increasing the benefit/cost-ratio of model-supported optimization and control.

The essential point to note at the beginning of any attempt to construct process models for optimization and control is that it makes no sense to begin without having a clearly formulated problem statement and objective function. This is simply because the model's value cannot be determined. Unfortunately, most papers on process modeling do not consider this fact seriously enough. The importance of the problem oriented objective can be illustrated by an elementary example, the fed-batch *Saccharomyces cerevisiae* cultivation process. Its basic behavior may be described by the following model, using simple Monod-like kinetics:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu(S)X - \frac{F}{V}X$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\frac{1}{Y_{XS}}\mu(S)X - mX + \frac{F}{V}(S_F - S)$$

$$\frac{\mathrm{d}V}{\mathrm{d}t} = F, \quad \text{where} \quad \mu(S) = \frac{\mu_{\max}S}{K_S + S + S^2/K_I} \tag{1}$$

X and S are biomass and substrate concentration; S_F the substrate concentration in the feed, V the reactor volume; F the substrate feed flow rate and K_S, K_I, $\mu_{\max}$, Y_{XS} and m are model parameters. The parameter values used for this specific example are: $K_S = 0.1$ g l^{-1}; $K_I = 50$ g l^{-1}; $\mu_{\max} = 0.3$ h^{-1}; $Y_{XS} = 0.5$; $m = 0.03$ g g^{-1} h^{-1}, $S_F = 500$ g l^{-1}.

Regarding the accuracy by which the model parameters can be determined from experimental data, it is well known that the K_S-parameter appearing in the Monod kinetics, cannot be identified with high accuracy from batch cultivation data (Munack, 1991). Inaccurate parameter values lead to modeling errors. Assuming that the true concentrations of biomass and substrate in this process are $X_{\text{true}i}$ and $S_{\text{true}i}$, then in the simulations using Eq. (1), we get $X_{\text{sim}i}$, $S_{\text{sim}i}$ upon assuming a parameter set (μ_{max}, K_S), which deviates from the set assumed to be true. The model error

$$J = f(K_S, \mu_{\text{max}})$$

$$= \sum \left[(X_{\text{true}i} - X_{\text{sim}i})^2 + 0.5(S_{\text{true}i} - S_{\text{sim}i})^2 \right] \qquad (2)$$

resulting from this assumption has the shape of a flat valley along a certain direction in the (μ_{max}, K_S)–phase plane. Thus, by choosing parameter values (μ_{max}, K_S) along this direction, the corresponding values of J will not change significantly. For a batch cultivation example calculated with $X_0 = 0.2$ g 1^{-1}, $S_0 = 10.0$ g 1^{-1}, $F = 0.0$, $t_f = 13$ h^{-1}, $\Delta t = 0.01$ h^{-1} some part of the error surface is shown in Fig. 1a. As can be seen from the figure, the sensitivity of the modeling error to changes in K_S is roughly 100 times smaller than to changes of μ_{max}.

The conclusion most often drawn from this fact is that special experiments must be performed in order to identify K_S more accurately. This generally produces costs.

Our objective, however, is usually not to obtain accurate kinetic parameters, but instead to optimize the process, e.g. to optimize the total biomass in a fed-batch cultivation within substrate limited/inhibited environments. With simple numerical calculation procedures we can analyze the influence of K_S on the optimal feeding profile and the total biomass. The optimal growth conditions for cultivation process described using Eq. (1) are:

$$\frac{\partial \mu}{\partial S} = 0, \qquad S_{\text{opt}} = \sqrt{K_S K_I} \qquad (3)$$

Hence, the optimal feeding profiles for different values of K_S, μ_{max} can easily be calculated by

$$F_{\text{opt}} = \left(\frac{\mu}{Y_{XS}} + m \right) \frac{XV}{(S_F - S_{\text{opt}})} \qquad (4)$$

In yeast (biomass) production, the main objective is to obtain as much biomass as possible at the end ($t = t_f$) of the cultivation. Hence, the process performance index J_{p} may be defined by

$$J_{\text{p}} = f(F_{\text{opt}}) = XV, \text{ at } t = t_f. \qquad (5)$$

For every set (μ_{max}, K_S) we obtain a feeding profile F_{opt} according to Eq. (4). These feeding profiles will lead in the real process (that with the true parameter set (μ_{max}, K_S)) to a performance index J_{p} defined by Eq. (5).

Corresponding simulation results are shown in Fig. 1b. Here one can clearly see, that the optimal feeding strategy and therefore the process performance index is essentially independent of the K_S. Hence, in this case it does not make sense to determine this parameter to a high level accuracy.

A clear-cut formulation of the objective helps to avoid unnecessary activities. Consequently, the concrete statement of the problem may save considerable time and costs.

2. Exploratory analysis

We start the discussion with a situation which often happens in industry. Consider a process that is already producing, however, a comprehensive process model which could be used to optimize the process is still missing. What we can immediately learn from the example with the saturation constant K_S is that, before we can really start with modeling, we must understand: (1) the clear problem formulation and corresponding objective function and; (2) which of the state or abiotic variables have a significant influence on this function.

In order to determine those variables, we propose a procedure which could be named an exploratory analysis of the available process data (Tukey, 1977; Simutis and Lübbert, 1997a).

As a first approach to representing the relationships describing this influence, process engineers often use simple black-box models. Most often
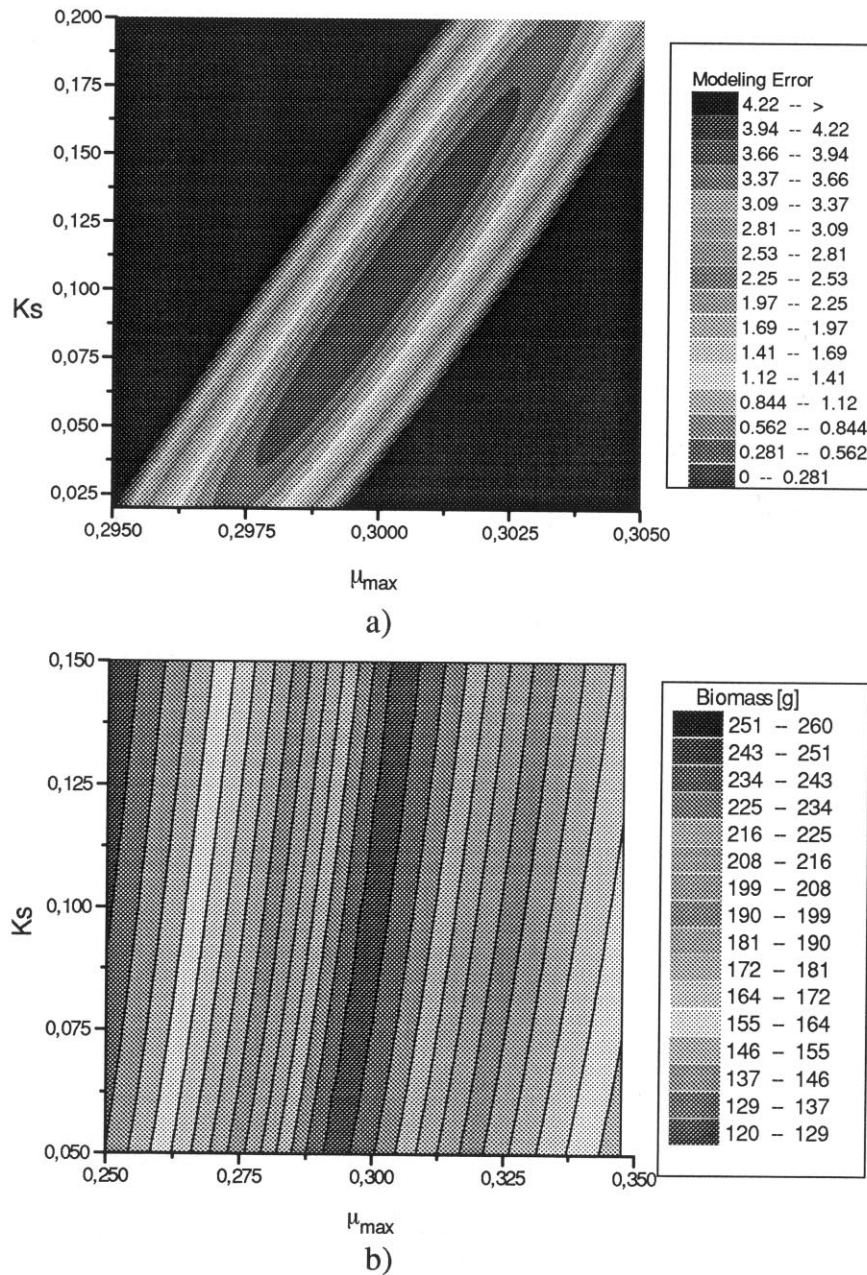
Fig. 1. (a) Modeling error as a function of the parameters K_S and μ_{max}. (b) Dependency of the total biomass calculated for an optimized fed-batch process as a function of the parameters K_S and μ_{max}.

they use power-law approaches called engineering correlations.

Recently, however, a much more powerful variant of black-box representations has been dis-cussed in the literature under the headline artifcial neural networks (ANN) (Morris et al., 1994; Sjöberg et al., 1995). In particular, complex non-linear relationships can be represented in a much

more accurate and flexible way with such networks. Even dynamic processes can be represented by neural nets (Narendra and Parthasarathy, 1990; Haykin, 1994).

Artificial neural networks learn from data measured at the process under consideration. The more relevant data one can use in the training procedure, the better ANN perform in reproducing the input–output relationships of the process. However, one should stress the term relevant. One should be careful with supplying neural networks with all the input data available. Then, one cannot expect good model identification results.

A simple analogue, the learning procedure of a child, helps to understand this problem: when you confront a child with information that does not clearly show the message you would like to transmit, then the child will become confused and does not make good progress in learning. Later on, after some basic education, you can come up with details and specialties. The same applies to neural networks. With complex multi-dimensional inputs and limited data records an efficient network design becomes more difficult. Every additional input variable drastically reduces the chance of learning the true relationships between the input and output quantities of the process (Kell and Sonnleitner, 1996).

In order to dixcover which variables are most important to the current objectives, we suggest a simple exploratory procedure.

We start with the choice of basic variables (e.g. substrate concentration) which we definitively know (a-priori) belong to the key variables influencing the specific growth rate, or the specific product development rate of the biochemical process. Process performance and the objective function for process optimization are indirectly determined by those specific rates in most cases. Hence, the essential practical problem is to determine the specific rate expressions using the state variables and some additional variables characterizing the abiotic phase of the culture like pH, pO_2, temperature, etc. The specific rate expressions are represented by means of an ANN and are immediately used as input quantities in the mass balance equation system. During the training of this basic artificial neural network the root mean square error (RMSE) between predicted and measured state variables is minimized (Simutis and Lübbert, 1997b).

All other variables than the basic ones are then systematically, i.e. one after the other, tested for their influence on the RMSE. This test is performed by adding the individual variables to the input variables of the trained basic neural network and by extending the training of the entire net with the additional variable.

In this way we obtain a classification of the importance of the variables tested in terms of the RMSE. This information helps to improve the efficiency of process modeling, as it is straightforward in constructing relationships like process models by dealing with the most important factors first.

One can now proceed one step forward in the direction of model generation.

If there is at all a variable that, added to the basic network, leads to a significant improvement with respect to the RMSE, then, in a second global step, the best of these extended neural nets can be taken as a new second generation basic network. The first procedure can then be repeated once more. This, once again, will lead to a classification of the variables available, but not considered, in the basic network. The procedure can be continued until no further variables are available which could lead to a significant improvement in the RMSE.

The reason for this kind of re-evaluation of the rank of the variables is that co-operative effects of variable pairs can be considered which cannot be detected with the simple first classification procedure.

With this exploration procedure we finally obtain a neural network that contains only those variables as input quantities which really make an improvement with respect to the objective function of the process and not merely a contribution to the noise. This will not only improve the significance of the network but also reduce the complexity of the network and the expenses for training and maintenance. Moreover, the result is a highly valuable information for all further model development efforts.

An essential problem of all knowledge representations is that they need to be validated by independent measurement data. With other words, in order to assure that the model works correctly we must test it with independent data that have not been used during the model parameter identification.

An example is the representation of diacetyl development during beer fermentation. Diacetyl is a key component determining the taste of beer. With the data available from many brewing runs a neural network was trained (Manikowski, 1996). The modeling result shows quite a good agreement with the data (Fig. 2a). However, during the validation procedure, the same neural network applied to different data records from the same bioreactor that had not been used during the model development showed a quite unsatisfactory result (Fig. 2b).

Hence, it must be clearly stated that it is not sufficient to perform a fit of some model to available data. The obtained representation must be validated. The recommended procedure is known under the name cross validation technique (Pollard et al., 1992; Haykin, 1994). With respect to the exploration procedure described above it should be noted that in every step of the classification or neural network development procedure, it is indispensable to determine the values of the objective functions with independent validation data sets in the sense of the cross validation technique.

3. Extrapolation problems

A model cannot immediately lead to benefits in a production process. But it can be used to improve the process performance. For instance, it can be used in a fed-batch process to determine an improved substrate feeding profile.

Assume for the moment that the optimization itself does not pose a problem. Then the result of an optimization procedure might be a feeding profile that leads the process through areas of the state space which have never been explored in practice. Hence, the question arises as to whether or not the model covers this area at all, or in other words, whether or not the model has sufficient extrapolation capabilities.

For the artificial neural networks discussed so far, it is well known that their performance outside the areas covered by the training data cannot be guaranteed. Hence, we must recognize there is a problem (Leonard et al., 1992; Simutis et al., 1995a).

In order to compensate for this deficiency we need additional information or knowledge about the process. The source, which obviously must be exploited first, is literature about the same or similar processes. This primarily leads to classical process models which one can at least assume have better extrapolation capabilities. The parameters of these models must be adapted to the experimental data of the particular process under consideration. Additionally, any further information or knowledge that could help to describe situations about which we do not have enough measurement data is welcome. A rich source of such additional information is the heuristics accumulated by process experts.

The straightforward way to proceed at this stage is to take all the information or models one can get within the time available for solving the task, and after identifying their parameters, use these models or this information simultaneously. This is what we understand by hybrid modeling. Combining different process models or parts of them in a parallel or sequential way and using their individual strengths.

Since the different parts of the hybrid model obviously own different performance characteristics, the question arises as to how to weight their individual results. In case of a parallel application the most practical solution is to build a weighted average from the individual results.

An example for the case of a parallel combination of a Monod-like model and an artificial neural network in determining the specific biomass production rate is the following.

Experience with neural networks shows that they usually are capable of representing non-linear relationships between multiple variables in a complex process more accurately than mathematical equation systems identified on the same data set, provided the information contained in the
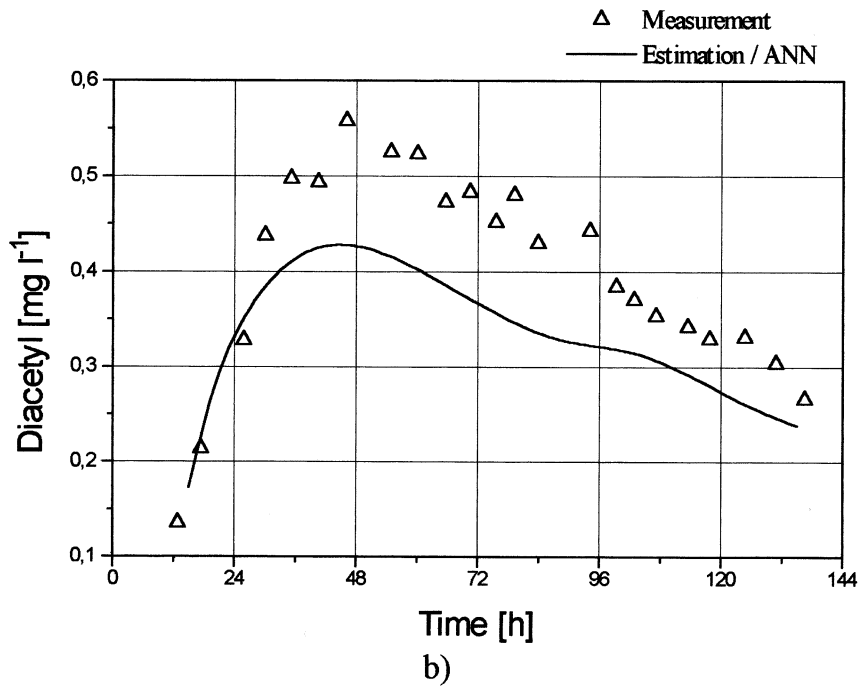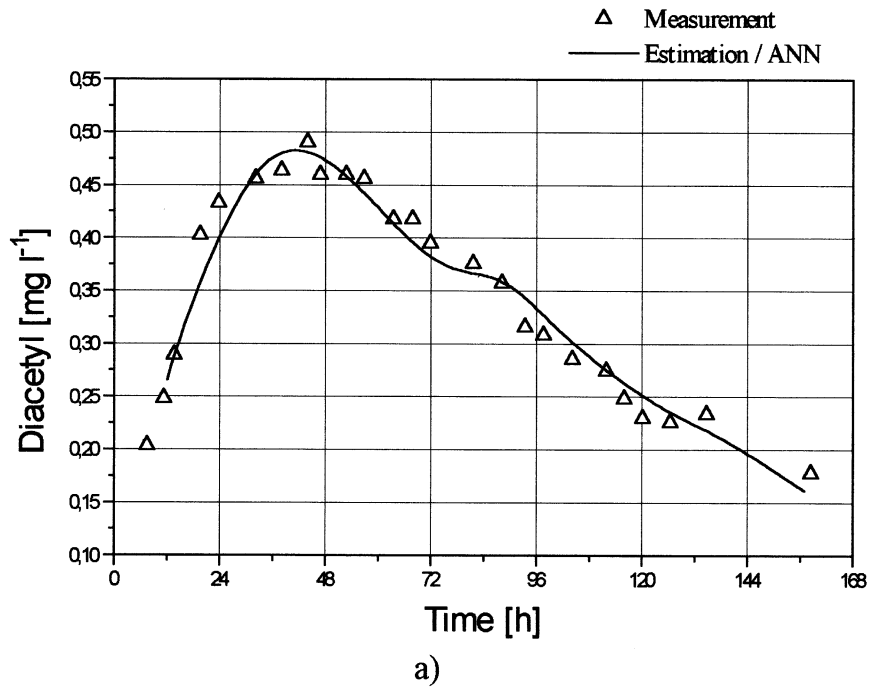
Fig. 2. The quality of the diacetyl-estimator during the industrial beer brewing process. (a) For one of the training sets. (b) For one of the validation sets.
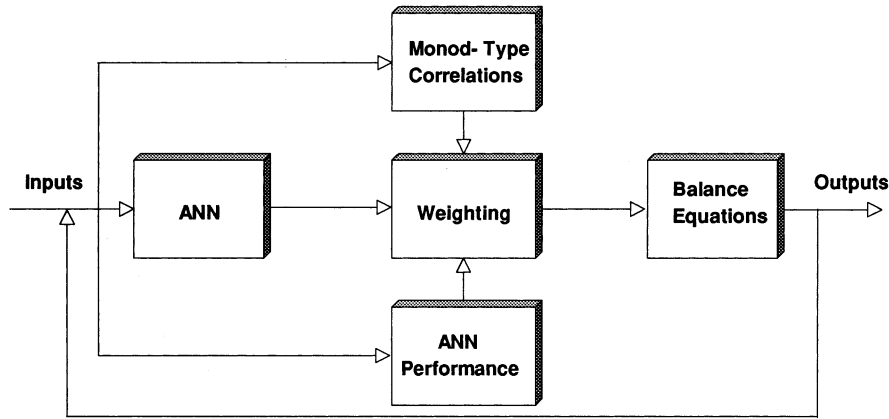
Fig. 3. Combination of the ANN with mechanistic model.

available data is sufficiently high. Thus, we prefer to use neural networks to represent the kinetic relationships wherever the conditions concerning the size of the training data set are met.

From the degree the condition is met, we can derive a weight for the neural network component. The key information is an amount of training data available around the different points in the state space of the input variables of the network. In order to provide an algorithm for determining the weights, which can be used in on-line applications of the ANN, the experimental data state space was structured by means of a cluster analysis and the working area of the i cluster was determined by following Gaussian bell-shaped functions

$$N_i(\boldsymbol{x}_{ci}, \boldsymbol{\sigma}_{ci}, \boldsymbol{x}) = \exp -\frac{(\boldsymbol{x} - \boldsymbol{x}_{ci})^2}{\boldsymbol{\sigma}_{ci}^2} \qquad (6)$$

where $\boldsymbol{x}$ is a vector of ANN input variables, $\boldsymbol{x}_{ci}$ is a vector of the i cluster center and $\boldsymbol{\sigma}_{ci}$ is a vector of the i cluster shape. The cluster centers can be easily determined with the 'k-means' algorithm and the cluster shapes, heuristically or using the 'global first nearest-neighbour' algorithm (Moody and Darken, 1989).

In order to determine the extrapolation measure (EM), i.e. the extrapolation capacity of the ANN for an arbitrary input point $\boldsymbol{x}$, we search for the training data cluster centered closest to $\boldsymbol{x}$ and calculate EM $= N$, the relative weight assigned to the ANN according to Eq. (6). Then, in a pro-

posed combination of the neural network with any mathematical or mechanistic model, the other components get the complementary weight (1-EM).

Elementary examples for the case of a recombinant mammalian cell cultivation are shown in Figs. 3 and 4. For the prediction of specific cell growth rate an ANN and the following Monod-like equation including biomass inhibition effects are used

$$\mu(S,X) = \frac{\mu_{\max}S}{K_S + S + S^2/K_I} \frac{K_X}{K_X + X} \qquad (7)$$

The inclusion of the biomass concentration inhibition effect on the specific growth rate is taken as a coarse approximation of the ammonium and lactate inhibition, that typical appears in mammalian cell cultivation processes (Goergen et al., 1994). Fig. 4 shows the pattern of the weighted specific growth rate, using the ANN extrapolation measure approach.

With the utilization of the classical model, which obviously must also be identified with the same data set, the question of the benefit of the combination arises. Does the hybrid model really have significant advantages?

Classical and hybrid modeling approaches were compared by Psichogios and Ungar (1992), as well as by Thomson and Kramer (1994), Simutis et al. (1995b). Fig. 5 shows a typical estimations of biomass concentrations in fed-batch yeast cultivation obtained with these methods (Simutis et

al., 1995b). The difference clearly indicates that the hybrid model performs better in this case, however, it should be made clear that the advantage lies in the amount of a-priori knowledge incorporated into the hybrid model as compared to the mathematical model. Usually, it is possible to activate a larger part of the available knowledge with the hybrid models. However, when it would be possible to provide a comparable amount of relevant knowledge in the mechanistic mathematical model, then the difference would be diminished. The hybrid models allow the direct use of knowledge as it becomes available and thus, avoids transformation losses. Additionally, this technique saves development time and hence cost.

There are many different possibilities to make use of the hybrid modeling technique. The choice of a particular variant depends essentially on the amount of a-priori knowledge available at the beginning (Psichogios and Ungar, 1992; Su et al., 1992; Schubert et al., 1994; Thomson and Kramer, 1994).

If there is a classical process model available that helps to solve the problem at hand, it may be of advantage to use this as a basic model and to improve the model performance by adding a neural network in a parallel fashion, as schematically depicted in Fig. 6, in order to compensate for the classical model's deficiencies.

4. Optimization problem

Optimization problems for biochemical reactors have been discussed by a number of investigators (Arkun and Stephanopoulos, 1980; Lim and Lee, 1991). Classical optimization techniques (e.g. Pontryagin's maximum principle) can be used to obtain optimal control profiles with simple process models. However, a considerable degree of understanding the complex optimization procedures is necessary to arrive at acceptable solutions.

Hybrid models take advantage of the combination of knowledge and information from different sources. Consequently, they are more complex than classical ones. This may lead to problems with classical optimization procedures, like dynamic programming or the application of Pontryagin's maximum principle for process optimization. Such difficulties might reduce the applicability of the model in practice and hence, its benefit.

Since the mentioned classical optimization methods are generally difficult to apply in practice they are seldom found in industry. Thus, it makes sense to look for simple-to-use alternatives that can also be applied to the complex hybrid models.

Today, where computing power becomes available to everybody at steadily decreasing costs, an alternative becomes more and more attractive. The random search procedures (Bohachevsky et al., 1986; Bremermann and Anderson, 1989; Montague and Ward, 1994; Fogel, 1995; Cardoco et al., 1996). Generally, they are simple to apply, are not restricted to simple mathematical models with very special properties and can cope efficiently with the various constraints (equalities and inequalities).

There are different algorithms by which the random search procedure can be implemented. They differ by their strategy to come out of a local minimum. Simutis and Lübbert (1997a) compared three different random search techniques:
1. Chemotaxis algorithm,
2. Simulated annealing algorithm,
3. Evolutionary programming algorithm,
for open-loop optimization of different biochemical conversion processes. They used polynomial and artificial neural network representations for constructing the optimum control variable trajectories.

Open loop optimization, where the evolutionary programming algorithm procedure is used in combination with back-propagation neural networks based on sigmoidal response functions, were found to lead to sufficient accuracy of the control variable profiles within acceptable computing times.

The benefit of this random search technique is that it can be applied by biochemical engineers without a very special education in optimization techniques. Its main advantage is that it results in
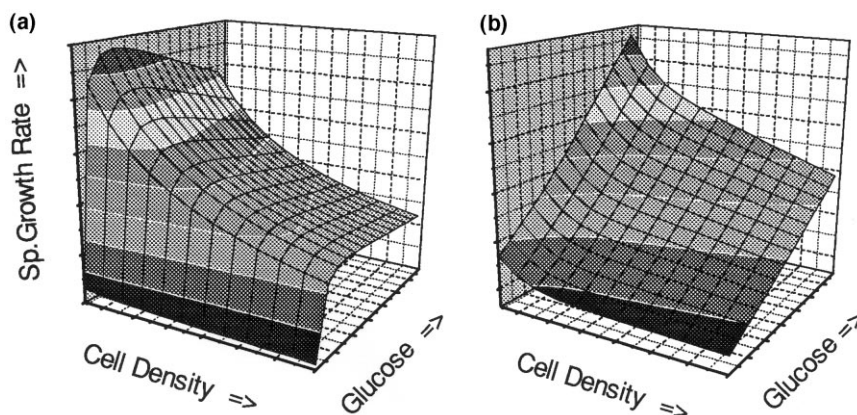
Fig. 4. Specific growth rate prediction in recombinant mammalian cell cultivation. (a) Prediction with Monod-type model. (b) Prediction with ANN. (c) Performance of ANN. (d) Prediction with the combined model.

sufficiently accurate trajectories within development times, which are acceptable for most applications in bioengineering. An essential point from the point of view of hybrid modeling is that these techniques are not restricted to well behaving models. Hence, using random search techniques instead of, e.g. the Pontryagin technique essentially reduces the development cost.

5. Software support

Up to now ideas and basic algorithms have been discussed. However, in order to obtain a benefit from them, the ideas must be implemented and work successfully at a concrete example, in a real plant.

When looking for practical implementation of the ideas, once again the primary aim is to maintain the benefit/cost-ratio within limits acceptable to the industrial applicants.

What we need to reduce the cost, are software tools that allow for a time efficient:
1. Process data acquisition.
2. Data exploration and rating.
3. Model generation and control system design.
4. Front-end control system development.

What can we expect from such software? Within a concrete hybrid model, all these blocks are dependent on each other and must thus be

compatible. In other words, they must be able to work co-operatively.

Another aspect is that each of them must easily be applicable by the biochemical engineer, in the sense that he must be able to apply them with a minimum of formal mechanical work. The software must work reliably, semi-automatically, quickly and be user friendly, so that the user gets enough time to concentrate on biochemical engineering problems.

The solution proposed here to that software problem is the development of two software packages:
1. The first is our process data management system, RISP, which covers the two items (1) and (4).
2. This works together with the second package, HYBNET, which focuses on hybrid modeling and thus, covers the points (2) and (3).

Here, we mention the HYBNET development only.

The HYBNET software package is a development system for hybrid structure definition and optimization (Oliveira et al., 1996). It is able to combine:
1. Differential equation systems,
2. Analytical mathematical relations,
3. Artificial neural networks,
4. Fuzzy expert systems/fuzzy neural networks
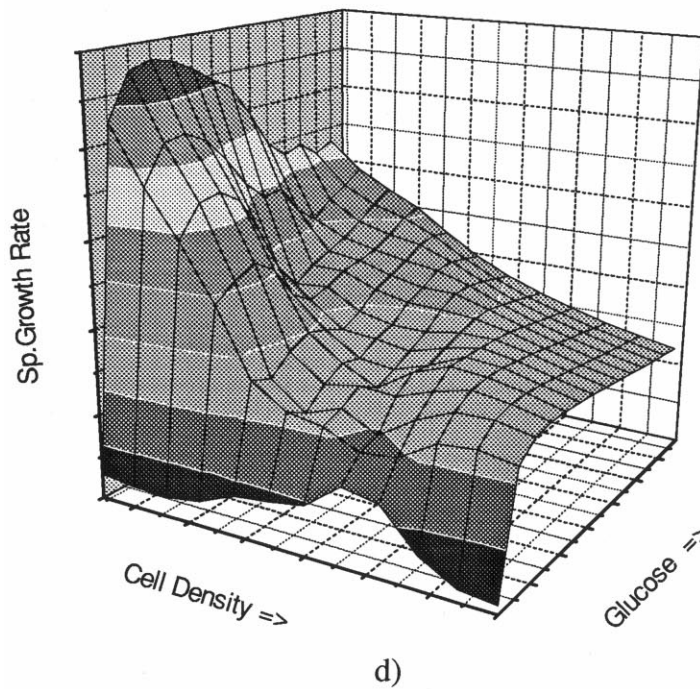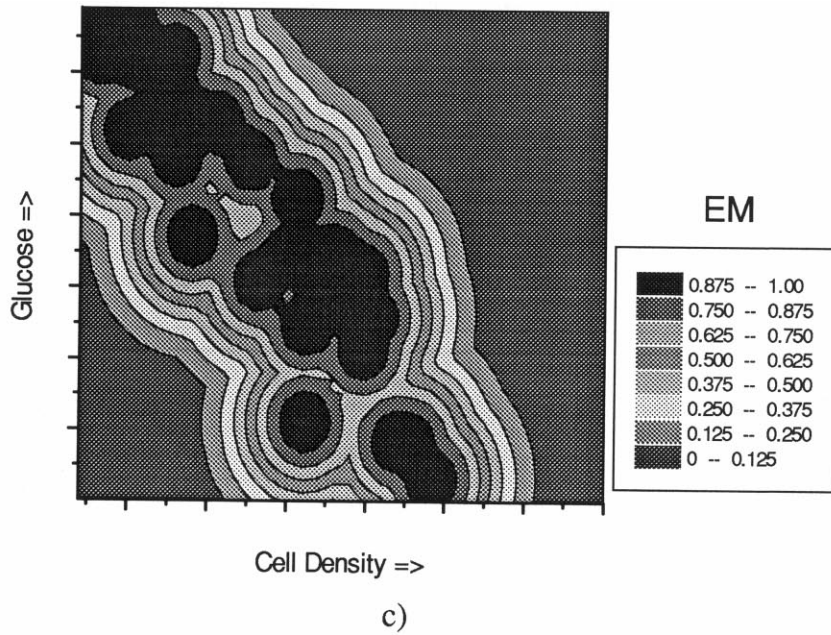and variants of them as individual components in

c)

d)

Fig. 4. (*Continued*)

a modular model which is organized in a network structure. The modules can describe different parts of a process in a comprehensive model or different models of the same partial component of the process, e.g. bioprocess kinetics as a part of the entire process model can be described by a
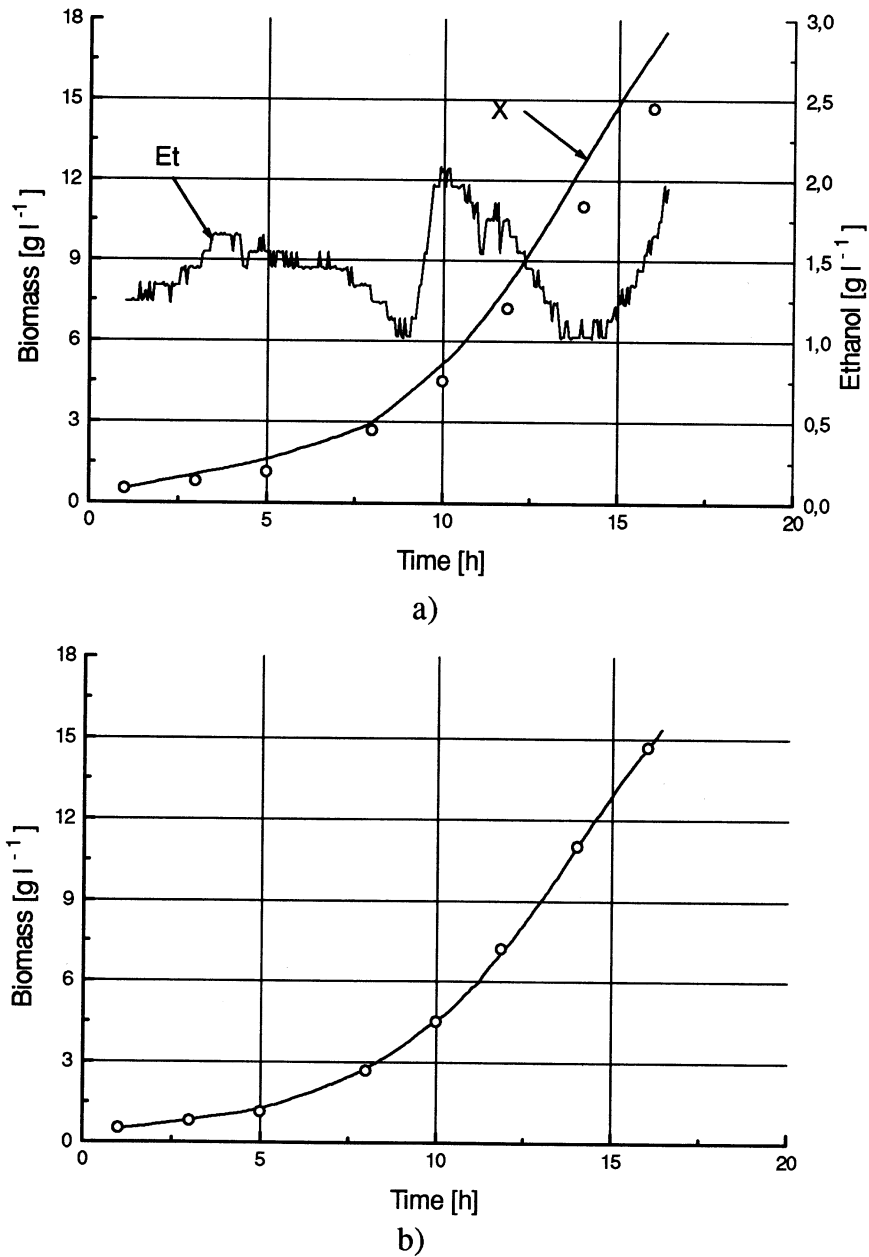
Fig. 5. Comparison of biomass estimation quality using. (a) Simple stoichiometry based estimator. (b) Hybrid modeling aproach in fed-batch yeast fermentation process.

classical Monod model and/or by a fuzzy neural network.

The essential new idea behind HYBNET is to allow the entire model assume a network struc-

ture. This has the decisive advantage, that the ideas of network training developed for simple artificial neural networks can be applied on a higher level. It can be shown that in many practi-
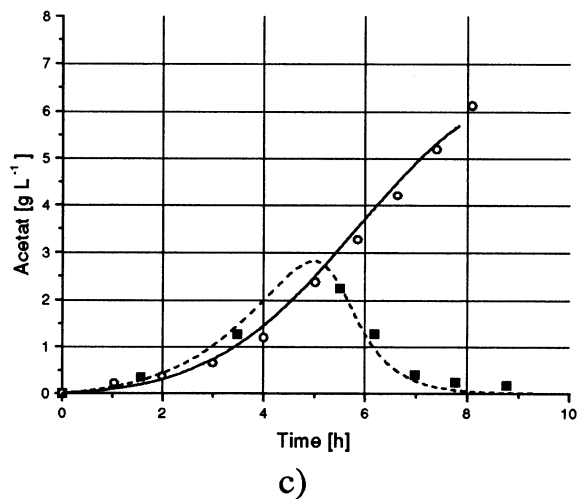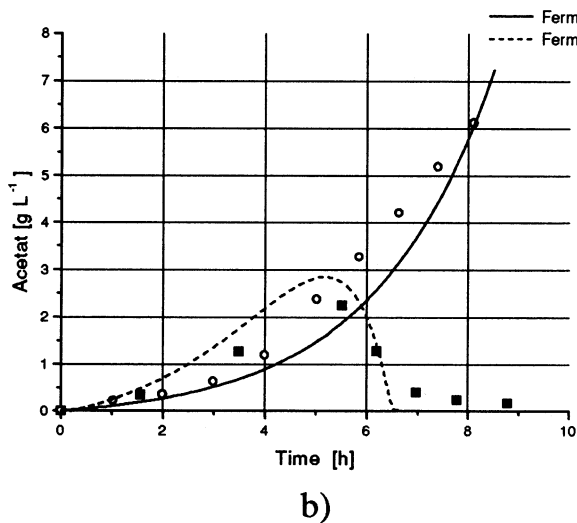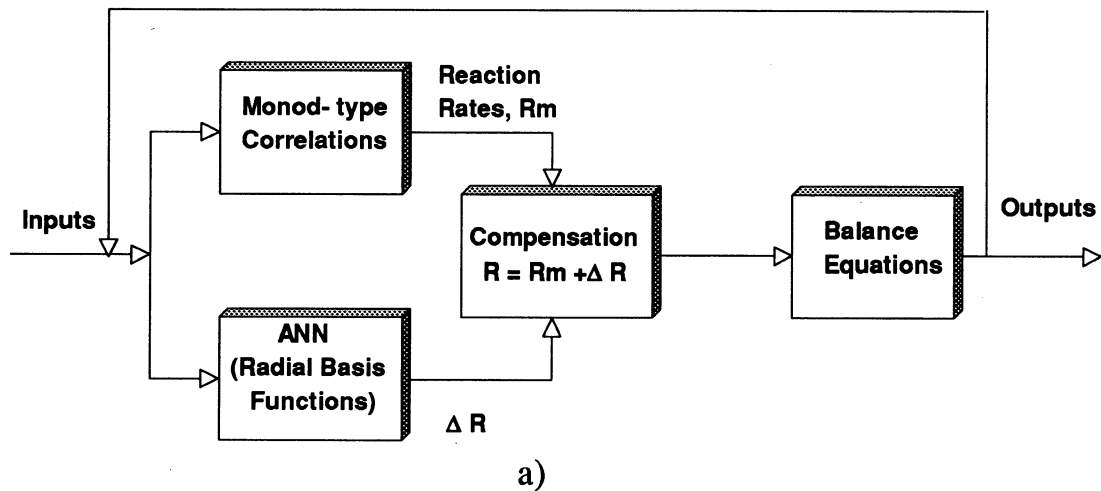
Fig. 6. Example for a neural network compensating the classical model's deficiencies in a recombinant protein production process. (a) Modeling scheme. (b,c) Quality of acetate predictions, on two validation set. (b) Results using Monod-type kinetic only. (c) Results with using hybrid model.

cal cases error back-propagation schemes can be used with advantage for tuning the entire process model (Oliveira et al., 1996).

From the practical point of view, the advantage of HYBNET is to provide as much flexibility in arranging partial process models in such a way that they can work together in solving a particular problem. In order to support its application, several ready-to-use components, which originally were developed for practical applications, are available as building block in a module library. They can easily be put together in a user friendly way supported by a user interface. Experienced users can develop their own modules and put them into the module library for future use. Several numerical facilities are implemented in order to liberate the user from most of the efforts in identifying model parameters on the available process data. Suitable algorithms are recommended by the system, however, alternative procedures can be chosen.
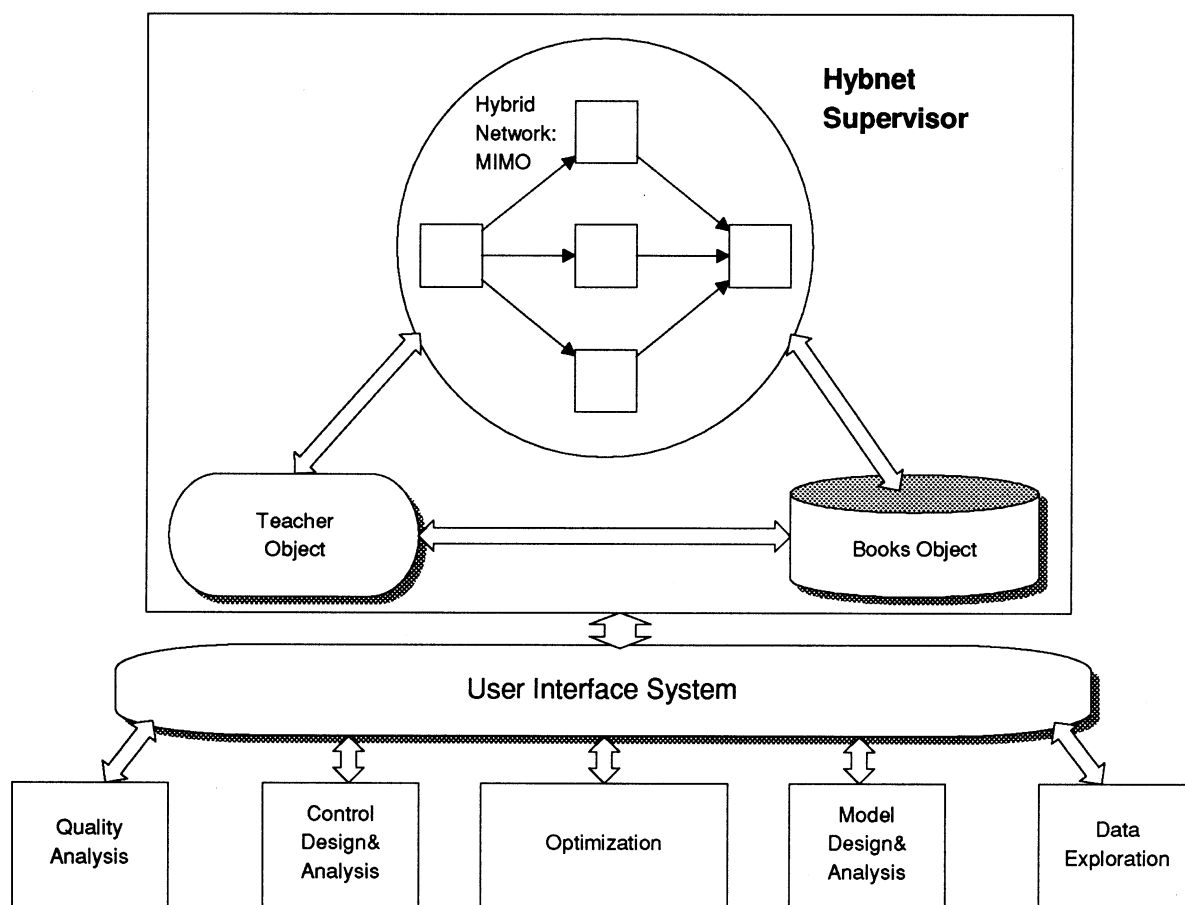
Fig. 7. The construction of the HYBNET software system.

The HYBNET system was developed according to the principles of object-oriented programming (Young, 1995). HYBNET's core is built up by three main objects:

1. The multi-input-multi-output system (MIMO) object.
2. The Teacher object.
3. The Books object.

The interactions between these three elements are controlled by the HYBNET's Supervisor (Fig. 7). The system's functionality is symbolized by the blocks that communicate with the supervisor through its user-interface system.

MIMO defines the relevant input/output variable, it determines the structure of its relationships and defines the parameters which must be optimized later on during the model identification procedure. The process data are managed in the Books object. Books provides the data, for both learning and validation.

Learning and validation procedures themselves are controlled by the Teacher object, which essentially performs the parameter estimation. It should be stressed that Teacher orients both learning and testing at the learning targets i.e. the actual objective function of the task to be solved.

An example of a HYBNET application in industrial biochemical process control is shown in Fig. 8 (Preusting et al., 1996). The model is used for the state prediction and optimization of the penicillin cultivation process. As input variables, the different feeds (carbon source, precursor, ammonia) as well as air flow and stirrer speed were
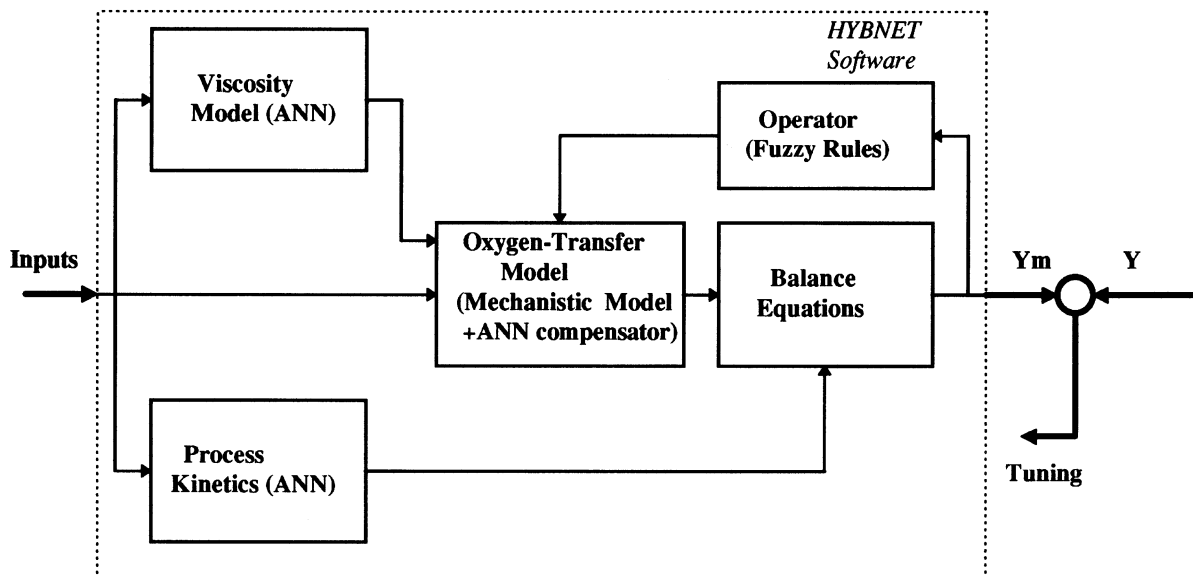
Fig. 8. Simple example of a Hybrid-Net structure for an industrial process.

taken. Output variables are biomass concentration, penicillin titer, oxygen uptake rate, pO_2, precursor and ammonia concentration. The HYBNET software allowed to integrate various type of process sub-models and to make their adaptation on the basis of the experimental data easily.

Currently, the software runs on DEC VAX/α Stations under the OPEN VMS and UNIX operating systems. It is being extended to WINDOWS NT system.

A final point we would like to mention regards the benefit that can be expected from an application of a model-supported optimization and control in an industrial environment. The benefit usually depends on two aspects. The first is the advantage, in terms of the process objective, which can be optained by applying the software. The second and often more important aspect is the time period over which the software can successfully be applied.

An active company is steadily looking for process improvements. Significant changes in the organism, the process control technology and the equipment and finally changes on the raw material market all require changes in the model or in the objective function. Then, the software cannot be used any longer in its original form. Thus, the benefit of the software essentially depends on the ease by which the software, including the underlying model and optimization procedures, can be adapted to the changing conditions and requirements in the industrial production plant.

Hence, flexibility of the software and its underlying techniques and algorithms is of high importance with respect to its benefit. Only when the software can steadily be adapted to the changing needs at reasonable maintenance costs, the overall benefit/cost-ratio can be kept below the limit, which determines its acceptance.

From this point of view, it is important in hybrid modeling concept to assign the neural network component a high value. Since it can automatically learn from process data, which become more extended during the application of the systems, its accuracy becomes better with time. This fact can be considered by moving the weights from the classical model to the neural network component, e.g. in parallel descriptions of the kinetics. In this way the entire model can automatically be improved.

6. Conclusions

The essential message of this paper is that the benefit/cost-ratio is the quantity that rules the acceptance of model-supported process supervision, optimization and control in industrial scale production processes. The performance of models can be increased by combining artificial neural nets, fuzzy rule systems and classical model approaches in form of hybrid models. This proved to be better than using anyone of these techniques separately. Hybrid modeling is particularly suited to production processes, since there we usually have enough data to make use of data driven approaches in combination with classical mathematical modeling.

The benefit is much dependent on the performance of the entire process model. However, it also sensitively depends on the length of the period it can really be used with advantage. The latter is essentially determined by the flexibility by which the software can be adapted to the changing requirements at the plant.

With HYBNET we presented a practical example of a flexible software covering the model development and process control tasks. It has been successfully used in several industrial and laboratory environments ranging from beer production via production of antibiotics to production of complex recombinant proteins.

Acknowledgements

References

Arkun, Y., Stephanopoulos, G., 1980. Optimizing control of industrial chemical processes: state of the art, a review. In: Proceedings of the Joint Automatic Control Conference, 1, WP5-A.

Bohachevsky, I.O., Johson, M.E., Stein, M.L., 1986. Generalized simulated annealing for function optimization. Technometrics 28, 209–217.

Bremermann, H.J., Anderson, R.W., 1989. An alternative to back-propagation: a simple rule for synaptic modification for neural net training and memory. Internal Report, Department of Mathematics, University of California, Berkeley.

Cardoco, M.F., Salcedo, R.L., Feyo de Azevedo, S., 1996. The simplex–simulated annealing approach to continuous nonlinear optimization. Comput. Chem. Eng. 20, 1065–1080.

Fogel, D.J., 1995. Evolutionary computation: toward a new philosophy of machine intelligence. IEEE, New York.

Goergen, J.L., Marc, A., Engasser, J.M., 1994. Biochemistry of cells in culture: modeling. In: Doyle, A., Griffiths, J.B., Newel, D.G. (Eds.), Cell and Tissue Culture: Laboratory Procedures. Wiley, Chichester.

Haykin, S., 1994. Neural networks. A Comprehensive Foundation. MacMillan, New York.

Kell, D.B., Sonnleitner, B., 1996. Good modelling practice (GMP): an essential component of good manufacturing practice. Trends Biotechnol. 13, 481–492.

Leonard, J.A., Kramer, M.A., Ungar, L.H., 1992. A neural network architecture that computes its own reliability. Comp. Chem. Eng. 16, 819–835.

Lim, H.C., Lee, K.S., 1991. Control of bioreactor systems. In: Rehm, H.J., Reeds, G. (Eds.), Biotechnology, vol. 4. Verlag Chemie, Weinheim, pp. 509–560.

Lübbert, A., Simutis, R., 1994. Adequate use of measuring data in bioprocess modeling and control. Trends Biotechnol. 12, 304–311.

Manikowski, M., 1996. Modellgestützte Überwachung und Steuerung des Gärprozesses in einer Bierbrauerei, Dissertation, Universität Hannover.

Montague, G.A., Ward, A.C., 1994. A sub-optimal solution to the optimization of bioreactors using the chemotaxis algorithm. Process Biochem. 29, 489–496.

Moody, J., Darken, C., 1989. Fast learning in networks of locally-tuned processing units. Neural Comput. 1, 281–294.

Morris, A.J., Montague, G.A., Willis, M.J., 1994. Artificial neural networks: studies in process modeling and control. Trans. IChemE 72A, 3–19.

Munack, A., 1991. Optimization of sampling. In: Rehm , H.J., Reeds, G., (Eds.), Biotechnology, vol. 4. Verlag Chemie, Weinheim, pp. 252–264.

Narendra, K.S., Parthasarathy, K., 1990. Identification and control of dynamical systems using neural networks. IEEE Trans. Neur. Netw. 1, 4–27.

Oliveira, R., Simutis, R., Azevedo, F., Lübbert, A., 1996. HYBNET, a new tool for advanced process modeling. Proceedings of the 1st European Symposium on Biochemical Engineering Science, Dublin, Ireland, pp. 182–183.

Pollard, J.F., Broussard, M.R., Garrison, D.B., San, K.Y., 1992. Process identification using neural networks. Comp. Chem. Eng. 16, 253–270.

Preusting, H., Noordover, H., Simutis, R., Lübbert, A., 1996. The use of hybrid modeling for the optimization of the penicillin fermentation process. Chimia 50, 416–417.

Psichogios, D.C., Ungar, L.H., 1992. A hybrid neural network, first principles approach to process modeling. AIChE J. 38, 1499–1511.

Royce, P.N., 1993. A discussion of recent developments in fermentation monitoring and control from a practical perspective. Crit. Rev. Biotechnol. 13 (2), 117–149.

Schubert, J., Simutis, R., Dors, M., Havlik, I., Lübbert, A., 1994. Bioprocess optimization and control: application of hybrid modeling. J. Biotechnol. 35, 51–68.

Sjöberg, J., Zhang, O., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P., Hjalmarsson, H., Juditsky, A., 1995. Nonlinear black-box modeling in system identification: a unified overview. Automatica 31 (12), 1691–1724.

Simutis, R., Dors, M., Lübbert, A., 1995a. Artificial neural networks of improved reliabiliity for industrial process supervision. In: Munack, A., Schügerl, K. (Eds.), Preprints of 6th International Conference on Computer Applications in Biotechnology, pp. 59–65.

Simutis, R., Dors, M., Lübbert, A., 1995b. Bioprocess optimization and control: application of hybrid modelling.

Authors's reply to comments by G. F. Andrews, J. Biotechnol. 42, 285–290.

Simutis, R., Lübbert, A., 1997a. Exploratory analysis of bioprocesses: using artificial neural network-based methods. Biotechnol. Prog. 13, 479–487.

Simutis, R., Lübbert, A., 1997b. A comparative study on random search algorithms for biotechnical process optimization. J. Biotechnol. 52, 245–256.

Sonnleitner, B., Käppeli, O., 1986. Growth of *Saccharomyces cerevisiae* is controlled by the 1st limited respiratory capacity: formulation and verification of a hypothesis. Biotechnol. Bioeng. 28, 927–937.

Su, H.-T., Bath, N., Minderman, P.A., McAvoy, T.J., 1992. Integrating neural networks with first principles models for dynamic modeling. IFAC Dynamics and Control of Chemical Reactors, Maryland, USA, pp. 327–332.

Tukey, J., 1977. Exploratory Data Analysis, Addison-Wesley, Reading, MA.

Thomson, M.L., Kramer, M.A., 1994. Modeling chemical processes using prior knowledge and neural networks. AIChE J. 40, 1328–1340.

Young, D.A., 1995. Object-oriented Programming with C++ and OSF™/Motif. Prentice-Hall, Englewood Cliffs, NJ.