On the applicability of neural networks for production planning under uncertainty
Stehouwer, H.P.; Aarts, E.H.L.; Wessels, J.

Published: 01/01/1994

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):
Memorandum COSOR 94-32

On the Applicability of Neural Networks for Production Planning under Uncertainty

H.P. Stehouwer
E.H.L. Aarts
J. Wessels

Eindhoven, August 1994
The Netherlands
On the Applicability of Neural Networks for Production Planning under Uncertainty

H.P. Stehouwer

E.H.L. Aarts\(^1\),\(^2\)

J. Wessels\(^1\),\(^3\)

\(^1\) Eindhoven University of Technology, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands

\(^2\) Philips Research Laboratories, P.O. Box 80000, NL-5600 JA Eindhoven, The Netherlands

\(^3\) International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

1 Introduction

Recent advances in the design and manufacturing of integrated circuits have brought the construction of parallel computers, consisting of thousands of individual processing units, within our reach. A direct consequence of these technological advances is the growing interest in computational models that support the exploitation of massive parallelism. Connectionist models [Feldman & Ballard, 1982] are computational models that are inspired by an analogy with the neural network of human brains, in which massive parallelism is generally considered to be of great importance. The corresponding parallel computers are called neural networks and the field of research neural computing. The greatest potential of neural computing is in the areas where high computation rates are required and present computer systems perform poorly. However, the potential benefit of neural computing extends beyond purely technical advantages. Many models in neural computing have human-like capabilities such as association and learning, which are essential in areas such as speech and image processing [Kohonen, 1988]. Moreover, these capabilities provide a kind of robustness and fault tolerance, since they compensate for minor variations in input data and damages in the network components.

In general, a neural network consists of a network of elementary nodes that are linked through weighted connections. The nodes represent computational units, which are capable of performing a simple computation that consists of a summation of the weighted inputs of the node, followed by the addition of a constant called the threshold or bias, and the application of a non-linear response function. The result of the computation of a unit constitutes the output of the corresponding node. Subsequently, the output of a node is used as an input for the nodes to which it is linked through an outgoing connection. For a detailed review of the different neural network models, we refer the reader to the textbooks by Aarts & Korst [1989], Hecht-Nielsen [1990], Hertz, Krogh & Palmer [1991], and Kosko [1992].

Production can be viewed as a transformation process in which materials are transformed into end products. These transformations, however, require resources. Resources are nonstorable services including manpower, machines and facilities, while materials are storable goods including purchased parts, raw materials, fuels, etc. Production planning is concerned with the efficient use of production resources in order to satisfy production requirements over time. Of obvious practical importance, it has been the subject of extensive research, and an impressive amount of literature covering an
apparently unlimited number of problem types has been created. Most research has traditionally
been focused on either deterministic or stochastic analysis. A problem is called deterministic if all
information that defines an instance of that problem is known in advance. Deterministic production
planning is part of mathematical programming.

During the past decade a substantial amount of literature emerged in which neural networks
are used in combinatorial optimization, which is part of mathematical programming. The reader is
referred to the survey article by Looi [1992J for an overview of this literature. The usual approach
is to formulate the combinatorial optimization problem in terms of a cost function, which is to be
minimized by the neural network. This approach suggests the use neural networks as an alternative
for classical techniques from combinatorial optimization [Papadimitriou & Steiglitz, 1982] and other
novel approaches like local search [Aarts & Korst, 1989] and constraint satisfaction [Nadel, 1989].
The most important motivation for using neural networks in this area is the potential speed up obtained
by massively parallel computation. However, until so far the results obtained using neural networks
in this area are disappointing compared with these other techniques.

Deterministic models are based on the assumption that all problem data are known in advance.
This assumption is often not justified. For example the production requirements may be subject to
random fluctuations. In stochastic optimization production planning problems are considered from
a probabilistic point of view. A deterministic production planning model may give rise to various
stochastic counterparts, as there is a choice in the parameters that are randomized and the optimality
criterion that is applied. A characteristic feature of these models is that the stochastic parameters are
regarded as independent random variables with a given distribution and that their realization becomes
known only after the planning decision has been made. Modelled as such, several production planning
problems can be solved using Markov and semi-Markov decision theory [Tijms, 1986]. However,
analysis in this area is usually already technically complicated for relatively simple models of the
underlying random processes.

Stochastic optimization models assume that the nature of the underlying random processes are
well understood and are not subject to change. In practice, however, this is often not the case. In such
cases the black-box paradigm might be useful. In this paradigm one accepts that one is not able to
model the underlying processes and one takes a parametrized black-box and tries to fit the parameters
in such a way that the black-box shows a sensible behavior on at least a representative set of examples.
Multi-layered perceptrons [Minsky & Papert, 1969] appear to be interesting candidates for being used
as black-boxes in production planning; for their generalization and interpolation abilities, for their
pattern recognition skills and for their abilities to deal with non-linearities and noisy data. Although
present, in the case of using multi-layered perceptrons, inherent parallelism is of minor importance. In
this report we investigate the use of multi-layered perceptrons as black-boxes in production planning
situations.

This report is organized as follows. In Section 2 we give an introduction in multi-layered
perceptrons and supervised learning. Section 3 discusses the applicability of multi-layered perceptrons
for production planning using a well-known taxonomy of Anthony [1965]. We divide the problem
area in which we see potential possibilities for the use of multi-layered perceptrons into two subareas:
forecasting, which is treated in Section 4 and decision making, which is treated in Section 5. In
Section 6 we give an example of how multi-layered perceptrons can be used in a lot sizing situation
with incomplete demand information. Finally, after the conclusion in Section 7, we give some
references.
2 Multi-layered perceptrons

Multi-layered perceptrons (MLPs) can be viewed as an extension of the single-layered perceptrons designed by Rosenblatt; see [Rosenblatt, 1958; Rosenblatt, 1962]. Rosenblatt showed that perceptrons can be used for adaptive pattern classification, by proving his famous perceptron convergence theorem.

This theorem states that the perceptron convergence procedure finds the weights, called connection strengths, of a one-layered perceptron that solves a given classification problem if such a solution exists. Among others, Minsky & Papert [1969] demonstrated the limitations of one-layered perceptrons by showing that they can only classify sets that are linearly separable. Minsky & Papert suggested the use of multi-layered perceptrons to overcome these difficulties. After the convincing argument of Minsky & Papert and the lack of a convergence procedure for multi-layered perceptrons, interest in the perceptrons dropped to a modest level. Recently, multi-layered perceptrons regained interest due to the discovery of suitable learning algorithms such as the back-propagation algorithm, that can be used to automatically find the connection strengths that correspond to a given input-output behavior [Rumelhart, Hinton & Williams, 1986; Werbos, 1990a].

In a multi-layered perceptron (MLP) the nodes are arranged in layers, and the connections are not allowed to cross a layer, i.e., there are connections between the inputs of the network and the nodes in the first layer and between subsequent layers only. This implies that the inputs of a node in the first layer correspond to the inputs of the network, while the inputs of the nodes in a higher layer are the outputs of the nodes in the preceding layer. The outputs of the nodes in the highest layer form the outputs of the network. The nodes that are not output nodes are called the hidden nodes, and the corresponding layers the hidden layers.

The inputs and the connection strengths of a multi-layered perceptron are in general real valued. The outputs of the nodes can also be real valued, depending on the choice of the response function that is used. A well-studied example of such a function is the logistic function \( \sigma : \mathbb{R} \to [0, 1] \) defined by

\[
\sigma(x) = \frac{1}{1 + \exp(-x)}.
\]

2.1 Learning and generalization

When applying MLPs to a certain task, besides choosing the number of layers and the number of units per layer, one has to choose the weights such that the network performs the task accurately. These are the parameters of the black-box to be fitted, which in general cannot be determined beforehand. The number of layers together with the number of units per layer determine the architecture of an MLP. In choosing an appropriate architecture there are two options. The first option is to take the architecture fixed while fitting the weights as good as possible. In this way a number of different architectures can be developed and compared. The second option is to adapt the architecture while fitting the weights; see for example [Fahlman & Lebiere, 1990]. The common terms for fitting parameters in the context of neural networks are learning or training, and when the learning is done on the basis of direct comparison of the output of the network with known correct answers, one speaks of supervised learning.

Given an MLP with \( N \) inputs and \( M \) outputs and a finite set \( S = \{s_1, \ldots, s_T\} \) of examples, the supervised learning problem can be described as follows. \( S \) is called the learning set and each example \( s_k \in S \) consists of a pair \( s_k = (x_k, t_k) \), where \( x_k \in \mathbb{R}^N \) represents an input vector for the MLP and \( t_k \in \mathbb{R}^M \) represents the corresponding desired output or target vector. The problem is to find weights such that the difference between the output vector of the MLP on input of a particular example and the target vector is minimized for the entire learning set. The traditional measure how well the network
has memorized the training examples is the sum of squared errors defined by

$$E(w) = \sum_{k=1}^{T} \| f(w; x_k) - t_k \|^2,$$

where $w$ denotes the weight vector and $f(w; x_k)$ denotes the output vector of the MLP with weights $w$ after processing input vector $x_k$. This makes the supervised learning problem equivalent to the task of searching weight space for a minimum of $E(w)$. In the case when $E(w)$ is differentiable, there are several powerful methods for descending the $E$ surface. The best-known among these is the generalized delta method, which can be implemented efficiently on an MLP by the backpropagation algorithm [Rumelhart, McClelland & Williams, 1986; Werbos, 1990b]. For an extensive overview of the literature on learning methods we refer the reader to [Xu, Klasa & Yuille, 1992].

Although the ability of a MLP to memorize and recall data in the abovementioned way is impressive, its functionality is nothing more than that of a lookup table. It becomes interesting when the network could extend this behavior to similar data it had never seen. This is called generalization or rule extraction.

If a MLP has been trained on a particular data set, its generalization ability is measured on examples which are not in this set. For this purpose often a so called test set is constructed. During training the MLP does not 'see' the examples in the test set. We define the network’s generalization ability to be the accuracy with which the network’s input-output relation agrees with the test set. In order for generalization to make sense, it is required that both the learning set and the test set are representative for the underlying task. In general it suffices to construct these sets by the same random process. For a detailed discussion of generalization we refer the reader to [Denker, Schwartz, Solla, Howard, Jackel & Hopfield, 1987].

### 2.2 Possibilities and limitations

In the neural network literature it had been demonstrated that MLPs are able to generalize in a broad scala of tasks; see [Maren, Harston & Pap, 1990] for an overview on neural network applications. A special well-documented class of supervised learning problems are those for which the target vectors $t_k$ are one of a finite number of possibilities (classes). These are called classification problems or pattern recognition problems. For any classification problem MLPs offer a possible solution, as an alternative for conventional techniques from classification; see the book by Duda & Hart [1973] for an introduction in this area. Huang & Lippmann [1988] studied the performance of neural network classifiers compared to conventional classifiers and observed a comparable generalization ability.

In the application of trainable neural networks in general it is important to distinguish between the following two types of applications [Barnard & Wessels, 1992]. In the first type of application the learning set is expected to be fully representative of the inputs to the net during applications. Specialized preprocessing may be used to ensure that only such “predictable” inputs arise, or the application may be such that predictability is inherent to the environment in which the system is used. In the second type of application the learning set can only represent certain features of the application at hand, and it is expected that inputs from time to time will differ distinctly from those in the learning set. In the prior case, the problem can therefore appropriately described as interpolation between the training samples, whereas the latter case also requires extrapolation to regions where no samples have occurred. It is commonly believed that MLPs are good in interpolation and bad in extrapolation [Barnard & Wessels, 1992; Haley & Soloway, 1992]. Therefore it is important that a problem to be attacked using MLPs is or can be transformed into an interpolation problem.

In general, historical data is used for the construction of learning examples. However it is not absolutely necessary for such data to be available. It is even possible that, when for example decisions
are taken only once and a while, historical data is no longer up to date or relevant. In such cases it may be possible to construct learning examples from old data (extrapolate) or without any data. One option then would be to model all relevant situations and construct learning examples for them.

3 Production planning

In planning production in any production organization, we can roughly distinguish two types of approaches. The first type of approach develops a single model that solves the problem. This could be for example a single mixed-integer mathematical programming problem solved on a rolling horizon basis. However, such an approach requires detailed information such as the forecasted demand of every end product for a relatively large planning horizon. Such forecasts usually contain large forecast errors and therefore the resulting planning is often unreliable. Even if long term forecasts are provided with sufficient accuracy, this approach is undesirable because of the size of the model and the associated data requirements. Notice also [Gelders & Van Wassenhove, 1981] that such an approach does not conform to industrial practice, where functional decision units with separate responsibilities and hierarchical decision levels exist.

A standard strategy for solving a large complex problem is to break it down into smaller parts, which are more manageable. This leads us to the second type of approach. This strategy of ‘divide and conquer’ typically leads to solving an overall problem by sequentially solving a number of subproblems. Using such a type of approach, there is need for integration between the different subproblems. An example of a production planning concept in which subproblems are treated in an integrated way is Hierarchical Production Planning, propagated by Hax & Meal [1975], who integrated decisions belonging to the area of aggregate production planning and detailed scheduling.

3.1 Anthony's hierarchy of decisions

This report concentrates on the second type of approach and in the now following we aim at identifying and classifying these subproblems into a framework of classes. This framework has to be such that it is possible to reach conclusions with respect to the applicability of MLPs for each of the classes. For this purpose we adopt an often cited taxonomy described in a book by Anthony [1965], which classifies topics that fall within the general subject labeled planning and control systems. This taxonomy applies directly to the planning of production in any production organization. Below, we briefly describe this taxonomy.

The following three processes are identified as the main topics of the taxonomy:

- Strategic Planning - "The process of deciding on the objectives of the organization, on changes in these objectives, on the resources used to attain these objectives, and on the policies that govern the acquisition, use, and disposition of these resources"
- Management Control - "The process by which managers assure that resources are obtained and used effectively and efficiently in the accomplishment of the organization's objectives"
- Operational Control - "The process of assuring that specific tasks are carried out effectively and efficiently".

Over the years the names of these topics have been somewhat modified to become strategic planning, tactical planning, and operational control. One of the conclusions drawn by Anthony [1965] is that decisions in strategic planning, tactical planning and operational control tend to correspond to a hierarchy in any of the following dimensions: as to the time span of the consequences (long-range, medium-range, day-to-day); as to the level in the organization (top management, top and operating management, supervision); as to the importance of a single action (major importance,
medium importance, little importance); as to the degree of uncertainty (high, medium, low); as to the level of detail (high, medium, low); and so on. We will refer to this hierarchy as Anthony's hierarchy of decisions.

A for our purposes important dimension is the kind of decisions that have to be made. Moving along this dimension from tactical planning towards strategical planning, decisions become more and more unique. A typical approach used in strategic planning is scenario analysis: different alternatives are being examined, after which one of them is being chosen. Characteristic of this type of decisions is that besides on clear economical grounds, they are often based on rather vague grounds (political, nostalgic). We are interested in decision situations in which we see potential possibilities for the use of neural networks that learn from examples. Our expectations towards the use of such networks lie in situations which are of repetitive nature. By this we mean situations in which the possibility exists to learn from situations in the past about future (re)actions in similar situations. When we move along the abovementioned dimension from tactical planning towards operational control, situations become more and more repetitive. Consequently in the rest of this report we concentrate on the part of the taxonomy covering tactical planning and operational control.

The literature on production planning in this area describes a number of production planning concepts, each with its own philosophy and designed for a specific production situation. Examples of such concepts are Hierarchical Production Planning [Hax & Meal, 1975], Material Requirements Planning [Orlicky, 1975], and Just-in-Time Manufacturing [O'Connor, 1982; Hall, 1983]. In this report we adopt a more general framework described in [Silver & Peterson, 1985]. This framework embraces Anthony’s hierarchy of decisions and identifies a number of managerial activities related to production planning decision making. In Figure 1 the for our purposes relevant part of this framework is shown and the hierarchy of decisions is indicated by a dashed arrow. For a detailed description of the different activities the reader is referred to [Silver & Peterson, 1985]. One of the dimensions of this hierarchy is the degree of uncertainty. Moving along this dimension from tactical planning towards operational control the degree of uncertainty decreases and decision making becomes more and more deterministic.

Most of the reported investigations in this area are either at the level of operational control or involve the application of neural networks in forecasting. Examples of applications at the level of operational control can be found in [Dagli, 1994; Maren, Harston & Pap, 1990] and include applications in process monitoring, fault detection, and quality control. Although interesting, we will not go into this subject and will concentrate on that part of the framework in which uncertainty about the future influences the decision process. In Figure 1 these are the activities which take as input forecasted information. In Section 4 we briefly discuss forecasting using MLPs and in Sections 5 and 6 we investigate the possibilities for using MLPs in decision making under uncertainty.

4 Forecasting

Forecasting is used to predict changeable circumstances so that planning can take place to meet coming conditions. It can be used to predict revenues, costs, profits, prices, and a lot of other variables. In a production organization forecasting is used to predict or estimate future demand. Therefore we focus on demand forecasting. In the now following we briefly describe two commonly used forecasting techniques, time series analysis and the use of economic indicators. A time series is a set of time ordered observations on a variable during successive and equal time periods. By studying how demand changes over time, a relationship between demand and time can be formulated and used to predict future demand levels. In time series analysis historical data are analyzed and decomposed to identify the relevant components which influence the variable being forecasted. We refer the reader to the
textbook of Tersine [1988] for a detailed description of the various time series analysis techniques.

Economic indicators are frequently used to predict future demand. The decision maker searches for one or more economic indicators (gross national product, bank deposits, etc.) that have a relationship with the variable to be forecasted. Regression techniques are used to estimate the relationship between the variables. An important feature of economic indicators is that they can also be used to predict turning points in demand, based on the changed values of known indicators. MLPs have been for example applied to time series analysis by Lapedes & Farber [1987], Weigend, Huberman & Rumelhart [1990], Tang, De Almeida & Fishwick [1991], Chakraborty, Mehrotra, Mohan & Ranka [1992] and Foster, Collopy & Ungar [1992], to prediction by using economic indicators by Kimoto & Asakawa [1990] and Sri rengan & Looi [1991], and to the prediction of turning points in the gross national product using economic indicators by Hoptroff, Bramson & Hall [1991]. From this literature we may conclude that the use MLPs in forecasting, although still in its infancy, has the potential to become a serious alternative for the traditional techniques.
5 Decision making

5.1 Modeling

Production planning models introduced in the literature differ in their orientation, scope and methodology. Two important criteria when considering production planning models are how well a model describes the processes by which production plans are determined in practice and how well a model prescribes the decision maker what to do in a certain production situation. Following Saad [1982], we call these criteria descriptive behavior and normative model design respectively. According to these criteria production planning models can be classified. A typical example of a descriptive model is the management co-efficient model [Bowman, 1963; Kunreuther, 1969]. This model assumes that managers behave efficiently on average, but suffer from inconsistency and bias to recent events. Linear regression is used to develop decision rules for actual production and work force decisions utilizing variables such as past sales, scheduled production, inventory and work force. The resulting regression model provides a descriptive model of past decision making. Note the resemblance with forecasting using economic indicators as discussed in Section 4. See for example [Dutta, Shekhar & Wong, 1994], [Hill & Remus, 1994], [Madey, Weinroth & Shah, 1992], [Murtaza & Fisher, 1991] for examples of applications of MLPs in this area. In the operations research literature, most of the time problems are formulated as the determination of optimal values for decision variables over time with respect to some optimality criterion. Such models are typically normative. Examples are models from mathematical programming and stochastic optimization.

In most models of decision making situations, time is modeled by the nonnegative part of the real line. Time can be modeled either continuously or discretely. While continuous time models speak for themselves, discrete time models subdivision the real line into discrete time intervals of units length called periods and labeled $t = 1, 2, \ldots$. The time encompassed by a model is called the planning horizon, which can either be finite or infinite. We concentrate on discrete time models with finite planning horizons. The reason for this is that the type of neural networks we considered has a finite number of inputs and outputs, which naturally leads to discrete time finite horizon decision models.

For practical applications a finite planning horizon implies that the world is supposed to stop beyond that horizon. Even though the system being modeled is sure to exist beyond the time encompassed by the model, there are strong practical reasons for preferring this formulation. First, information about the distant future tends to be less reliable than for the near future. Second, since problems encompassing more time periods tend to be more difficult to solve, models with shorter planning horizons are attractive for computational reasons. In order to deal with the ‘infinite horizon’ situation one faces in practice, often a so-called rolling horizon is used [Baker & Peterson, 1979]. When using a rolling horizon, the finite horizon problem encompassing the first say $m$ periods, is solved and only the first period’s decisions are implemented. One period later the planning horizon is shifted and the process can start again. In this way the solution is revised when new information about future periods comes in. An interesting subject of research using a rolling horizon is the influence of the length of the planning horizon on the quality of the overall solution and the existence of so-called forecast horizons, i.e. the minimal length of the planning horizon such that subsequently adding more information about future periods will not change the first period’s decision; see [Kunreuther & Morton, 1973; Kunreuther & Morton, 1974; Lundin & Morton, 1975] for a detailed treatment of this subject.

5.2 Uncertainty

In the real world, many forms of uncertainty effect the production process. One can categorize these different types of uncertainty into environmental uncertainty and system uncertainty [Murthy & Ma,
Environmental uncertainty consists of uncertainties that are beyond the scope of the production organization. This includes demand uncertainty due to uncertainty in customer orders and supply uncertainty due to unreliable vendors. Other examples are uncertainties in reactions of customers on stock out and selling prices. System uncertainty consists of uncertainties within the scope of the production system. These include uncertainties in yield, production lead time, quality, production rates, capacities and the information processing system. As an example in the now following we look at models in which the only possible uncertainty is in the demand process.

In practice one often faces the situation in which for a certain small number of periods into the future demand can be considered deterministic and is known with certainty. Beyond this time span nothing is known with certainty and the further into the future one looks, the less certain one is about demand. An approach, usually followed in practice, is to try to forecast the demand for a number of periods into the future using for example time series analysis [Tersine, 1988]. Using these forecasts, the decision problem can be solved on a rolling horizon basis as mentioned above. Inherent to this approach is that forecasts of future demand are fixed, and considered as the "real" demand. What remains is the problem of optimally planning the production, given this demand. In this way uncertainties in demand are not taken into account. One of the tricks to overcome problems in situation where the actual demand differs from the demand previously forecasted for that period (forecast errors) is to use a so called safety stock. Another remedy is to consider the production capacities lower than they actually are, so there is some space to be able to anticipate on extra, not forecasted demand; see [Silver & Peterson, 1985].

As mentioned in the introduction another way of dealing with this kind of uncertainty in demand would be to suppose that, say from period \( n \) on, the demands are independent realizations of a random process, possibly with unknown parameter values. If unknown, these parameters can be estimated from historical data. In this way, the realization of the demand in period \( i \), for \( i > n \) becomes known at the beginning of period \( i - n \) and demand would always be known for a time-horizon of \( n \) periods. With such a modelling of the demand process, decision problems can be formulated as Markov decision problems [Tijms, 1986] and solved as such. Disadvantage of this type of approach is that it only works for very simple models of the demand process, whereas, in reality, the demand process is frequently rather complex and not well-understood. In such cases a black box approach such as the use of MLPs that learn from examples might be helpful. Major advantage of such an approach would be that uncertainties are taken into account implicitly. No explicit model of the demand process has to be provided and the examples, with which the network is trained, are taken from situations from the past of which after all it is relatively easy to calculate what had been the optimal decision. In the next section we go into this subject and discuss the possibilities for using MLPs in a production planning situation with incomplete demand information.

6 An example: lot sizing with incomplete demand information

Consider the following situation. Production has to be planned for a single end product on a single resource with an unlimited amount of production capacity per period. The planning horizon is discrete and infinite, with periods labeled \( t = 1, 2, \ldots \). There is incomplete information with respect to future demand. The demand for end product in period \( t \) becomes known at the beginning of period \( t - n \), i.e., demand is always known for a finite horizon of \( n \) periods. Production is on order, so assume that \( n \geq 1 \). There can be several costs involved. Examples are production costs, inventory holding costs, stockout costs, back-logging costs and smoothing costs. The problem is how to plan production such that the resource is used efficiently with respect to costs and such that the demand for end product is satisfied. Such models are known in the literature under the name lot sizing models; see [Bahl,
procedure ROLLING_HORIZON
begin
    \( t := 1; \ I_0 = 0; \)
    while true do
        begin
            while \( d_t \leq I_{t-1} \) do
                begin
                    \( X_t := 0; \)
                    \( I_t := I_{t-1} + X_t - d_t; \)
                    \( t := t + 1; \)
                end; \{d_t \gt I_{t-1}\}
                COLLECT(\( d_t, d_{t+1}, \ldots, d_{t+n-1} \));
                FORECAST(\( d_{t+n}, d_{t+n-1}, \ldots, d_{t+n-m-1} \));
                DETERMINE(\( X_t, X_{t+1}, \ldots, X_{t+n+m-1} \));
                IMPLEMENT(\( X_t \));
                \( I_t := I_{t-1} + X_t - d_t; \)
                \( t := t + 1; \)
            end;
        end;
    end;
end;

Figure 2: Pseudocode of a rolling horizon strategy.

Ritzman & Gupta, 1987] for an extensive overview of this literature.

An approach that suits this situation is a rolling horizon strategy. Using this strategy a so called rolling schedule is formed by solving a finite horizon problem and implementing only the first period's decisions. One period later, the finite horizon problem is updated and the process repeated. In principle, extending the planning horizon by including forecasts of periods beyond the horizon will provide some incremental benefit simply because a greater amount of relevant information is brought into the analysis; see for example [Baker & Peterson, 1979; Blackburn & Millen, 1980]. In the now following we will describe a rolling horizon strategy in which the horizon is enlarged by adding forecasts of \( m \) extra periods. Let \( d_t \) denote the demand of end product in period \( t \), let \( I_t \) denote the amount of inventory at the end of period \( t \) and let \( X_t \) denote the amount of production in period \( t \).

The rolling horizon strategy can be described in pseudocode as in Figure 2.

The procedure COLLECT involves the gathering of the demand information for the next \( n \) periods. The procedures FORECAST represents a forecast procedure using time series analysis. The procedure DETERMINE is an algorithm that finds a solution to the associated finite horizon problem, this can be either an optimization algorithm or some heuristic. It depends on the specific cost structure, the demand characteristics and the length of the planning horizon what will give the best results; see for example [Bitran, Magnanti & Vanasse, 1984; Blackburn & Millen, 1980] for a comparison of different lot sizing heuristics in a rolling horizon environment. The procedure IMPLEMENT involves the actual production.

Using this strategy as a starting point, there are several possibilities for employing MLPs. A trivial option would be to use an MLP to implement the procedure FORECAST, that is as a forecasting device. This subject has been treated in section 4. Another option would be to use an MLP to implement the procedure DETERMINE. However, being a deterministic problem, there is no need for using MLPs and classical techniques like dynamic and linear programming are more suitable. Another option is to
replace the procedures FORECAST and DETERMINE as a whole by an MLP, which takes as inputs the known demands $d_1, \ldots, d_{i-1}$ and the inventory level $I_{t-1}$. The output of the MLP could then be the production plan $(X_t, X_{t+1}, \ldots, X_{t+n+m-1})$, but by only implementing the first decision, the task for the MLP can easily be simplified in that the network only has to determine $X_t$.

In the original formulation, $X_t$ is based on the known demands $d_1, \ldots, d_{i-1}$ and the explicitly forecasted demands $d_{i+n}, \ldots, d_{i+n+m-1}$. Using an MLP such explicit forecasts need not to be made. The MLP is trained with historical demand information, and this information does not suffer from incompleteness. A learning pair can be constructed as follows. Take a demand sequence of length $n + m$ from the demand history. Use procedure DETERMINE to obtain the corresponding production plan. Let $X$ be the corresponding first period’s production level. The learning pair is constructed by taking the first $n$ demands as input vector and $X$ as desired output vector.

Using this technique in our group at Eindhoven University of Technology, several experiments were done for the so-called Wagner-Whitin model [Wagner & Whitin, 1958]. In that case the only costs are variable production costs, setup costs and inventory holding costs. We have compared this technique with several known heuristics for a number of different demand patterns and the results are very encouraging; for details see [Van Kraaij, 1991; Zwietering, Van Kraaij, Aarts & Wessels, 1991].

Currently we are investigating extensions of the Wagner-Within model, such as situations with more products, more resources, capacitated resources, other cost structures, and so on. It seems that for more complex situations the best approach is to make a decomposition of the problem at hand into a number of subtasks. Some of these subtask are to be attacked by traditional techniques and some by using MLPs. The decomposition must be such that the subtasks to be solved by an MLP must be a classification task and as simple as possible.

7 Conclusion

In this report we have investigated the use of MLPs as black-boxes in production planning situations. We have used Anthony’s hierarchy of decision’s together with a framework described in [Silver & Peterson, 1985] to identify the area in production planning in which MLPs have potential possibilities. Here two criteria appeared important: the kind of decisions that have to be made and the degree of uncertainty. We have discussed forecasting and decision making in production planning and the use of MLPs therein. Finally we have described how MLPs can be integrated in a rolling schedule environment where the problem is to determine the lot sizes and there is incomplete demand information.

References

ANThony, R.N. [1965], Planning and Control Systems: A Framework for Analysis, Studies in management control, Graduate School of Business Administration, Harvard University, Boston.


KUNREUTHER, H. [1969], Extensions of Bowman’s theory on managerial decision making, *Management Science* 15, 415–.


MADEY, G.R., J. WEINROTH, AND V. SHAH [1992], Integration of neurocomputing and system simulation for
modeling continuous improvement systems in manufacturing, Journal of Intelligent Manufacturing 3, 193–204.


O'CONNER, B.J. [1982], How do the Japanese get higher productivity then we do?, Proceedings of the 25th Annual International Conference of the American Production and Inventory Control Society, 477–481.


ROSENBLATT, R. [1958], The perceptron: A probabilistic model for information storage and organization in the brain, Psychological Review 65, 386–408.


TERSINE, R.J. [1988], Principles of Inventory and Materials Management (3 ed.), North-Holland.


