

Fuzzy Supervisory Control of Manufacturing Systems

Stratos Ioannidis, Nikos Tsourveloudis, *Associate Member, IEEE*, and Kimon Valavanis, *Senior Member, IEEE*

Abstract—A supervisory controller is derived for scheduling (single/multiple-part-type, reentrant) production networks. The supervisory controller is used to tune a set of lower level distributed fuzzy control modules that reduce work-in-process (WIP) and synchronize the production system's operation. The overall production-control system is viewed as a two-level surplus-based system with the overall control objective to keep the WIP and cycle time as low as possible, while maintaining quality of service by keeping backlog to acceptable levels. The production rate in each production stage is controlled to satisfy demand, avoid overloading, and eliminate machine starvation or blocking. The system's improvement is demonstrated using a set of performance measures. Extensive simulation results show that the supervisory controller, when compared with the single-level distributed fuzzy controllers reduces WIP and cycle time while keeping backlog to acceptable levels.

Index Terms—Backlog, fuzzy supervisory control, production networks, work-in-process (WIP).

I. INTRODUCTION

CURRENT ADVANCES in manufacturing have resulted in improving manufacturing processes, but have also led to changes in manufacturing management. Concepts such as throughput, cycle time, work-in-process (WIP), flexibility, and quality are traditionally some of the most important performance measures of manufacturing systems. The increased need for speedy and punctual delivery of products and goods has placed more emphasis on the reduction of product cycle time, backlog, and inventory-related costs.

Production-control policies include, among others, research on simulation studies of certain policies on specific systems; queuing theory-based performance analysis, stability, and optimal control, as well as fluid approximations of discrete systems. According to Gershwin [12], production-control policies may be classified as token based, time based, and surplus based. Token-based systems, including Kanban, Production Authorization Card [5], and Extended Kanban Control Systems [8], involve token movement in the manufacturing system to trigger

events. When an operation is performed or a demand arrives, a token is either created or taken from a specific location. Only when a token exists at the first location (and only if space for tokens exists at the second location), does the operation take place.

Time-based systems operate on a time basis; for example, material requirements planning (MRP) systems attempt to determine the time at which an operation should take place.

In surplus-based systems, decisions are made on the basis of how far cumulative production is ahead of or behind cumulative demand. Hedging-point, two-boundary, and base-stock policies are based on surplus and backlog. Wein [21] found in an extensive series of experiments on simulations of semiconductor fabrication facilities that the most important factor in determining the performance of a factory was the process by which material is released into the system. Decisions made once the material is inside (sequencing, dispatching) are less important. Surplus-based systems may be viewed as a consequence of Wein's observation, since the release policy is the most important factor in factory performance; the best policy is one that views all or many points as release points to the rest of the factory.

When considering simple manufacturing systems, analytical results produced thus far have demonstrated the superiority of surplus-based systems [1], [4]. More specifically, hedging-point policies have been proven optimal in minimizing production cost in single-stage, single-part-type system scheduling [1], [4]. Generalizations to more than one-part-type or production stages have proven to be difficult [18], [20], but obtained solutions [15], [16] may be successfully applied in real manufacturing systems [14], [23].

In summary, it is a common belief among many researchers that for complex production systems, the problem of scheduling, in order to minimize costs due to inventory and nonsatisfaction of demand, cannot be solved analytically. Since neither analytical nor computational solutions are achievable, heuristic policies are suggested to control job flow within production systems [2], [3], [7], [19], [22].

The research reported in this paper is the natural outgrowth of the work published in [19], in which a set of distributed single-level fuzzy controllers has been derived to reduce WIP and synchronize the production system's operation. This paper considers multiple-part-type production networks, and it views the overall production-control system as a surplus-based system. However, it differs from existing approaches and previous work [19], in that the overall system is a two-level control architecture with a supervisory controller at the higher level used to tune the operation of the lower level distributed fuzzy controllers. The overall control objective is to keep the WIP and cycle time as low as possible and, at the same time, maintain quality of service (QoS) by keeping backlog at acceptable levels. The production

Manuscript received April 6, 2003. This paper was recommended for publication by Associate Editor N. Viswanadham and Editor S. Hutchinson upon evaluation of the reviewers' comments. This work was supported in part by the Hellenic General Secretariat for Research and Technology, Joint Research and Technology Programs, R&TD Collaboration between Slovenia and Greece, under Grant Code: EPAN M.4.3 2013555. This paper was presented in part at the 10th IEEE Mediterranean Conference on Control and Automation, Lisbon, Portugal, July 2002.

S. Ioannidis and N. Tsourveloudis are with the Department of Production Engineering and Management, Technical University of Crete, Chania 73100, Greece (e-mail: nikost@dpem.tuc.gr).

K. Valavanis was with the Department of Production Engineering and Management, Technical University of Crete, Chania 73100, Greece. He is now with the Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: kvalavan@csee.usf.edu).

Digital Object Identifier 10.1109/TRA.2004.825518

rate in each production stage is controlled in a way that demand is satisfied, overloading of the production system is avoided, and the production system operation is synchronized to eliminate machine starvation or blocking.

The rest of the paper is organized as follows. Section II summarizes the architecture of the three production modules of the distributed fuzzy logic controllers, already published in [19]. Section III describes the proposed fuzzy-logic-based supervisory controller architecture. Section IV includes simulation results along with comparisons between the single-level fuzzy distributed control approach and the proposed fuzzy supervisory method. A cost analysis of the proposed methodology is also presented in the same section. Contributions are summarized in Section V, and further research is outlined.

II. PRODUCTION-CONTROL MODULES

A production system is usually viewed as a network of machines/workstations and buffers. Items receive an operation at each machine and wait for the next operation in a buffer with finite capacity. Random machine breakdowns disturb the production process, and phenomena such as starvation and/or blocking may occur. Due to a failed machine with operational neighbors, the level of the downstream buffer decreases, while the level of the upstream buffer increases. If the repair time is large enough, the broken machine will either block the next station or starve the previous one. This adverse effect will propagate throughout the system.

The events that may occur in a production network include changes in buffer states and changes in machine states. Buffers may be full or empty while machines may be up (operating) or down (under repair). When a machine is up, it can be starved if one of the preceding buffers is empty. In this case, the machine is forced to stop producing (idle) until a part arrives at the empty buffer. Respectively, if a machine is up, it can be blocked if one of the succeeding buffers is full. Then, the machine is forced to be idle until the full buffer level is decreased. When a machine breaks down, preceding machines keep operating until one of their downstream buffers is filled. Similarly, succeeding machines continue processing until their upstream buffers become empty.

According to the production-floor modeling approach introduced and explained in [19], every manufacturing system may be decomposed into three basic modules: the line [Fig. 1(a)], assembly [Fig. 1(b)], and disassembly [Fig. 1(c)] module. The line module includes a machine M_i , which takes unfinished items from an upstream buffer $B_{j,i}$ and after processing, sends them to a downstream buffer $B_{i,l}$. In the assembly operation, a machine M_i obtains two or more parts or subassemblies, following an assembly factor $a_{j,i}$ (the number of items i needed to form a unit of the assembled item), from more than one upstream buffer $B_{j,i}$, brings them together to form a single unit, and sends it to a downstream buffer $B_{i,l}$, as shown in Fig. 1(b).

The disassembly operation involves a machine M_i taking unfinished single units from one upstream buffer $B_{j,i}$, separates them to two or more parts or subassemblies, following a disassembly factor $\delta_{i,k}$ (the number of items i in which a unit of the disassembled item is decomposed), and sends them to down-

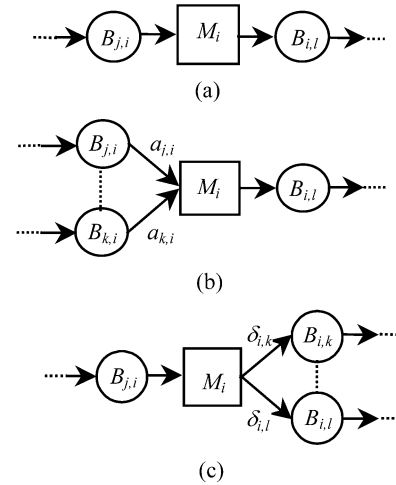


Fig. 1. (a) Line module. (b) Assembly module. (c) Disassembly module.

stream buffers $B_{i,k}$, as shown in Fig. 1(c). These modules, if connected to each other, may represent manufacturing networks of various layouts.

Each of the three modules may be implemented in terms of a fuzzy controller with input variables:

- buffer levels b_{ji} and b_{ik} of the upstream and downstream buffers;
- state s_i of machine M_i ;
- production surplus x_i of M_i , which is the difference between actual production and demand.

The output variable of every controller (or production module) is the processing rate r_i of each machine M_i . The buffer levels, surplus, and the processing rate of each machine take linguistic variations with certain membership functions. The machine state s_i is crisp and can be 1 (up) or 0 (down). The control objective in all cases is to meet demand and, at the same time, to keep WIP as low as possible. This is achieved by regulating the processing rate at every time instant, according to the following general rules.

- If the surplus level is satisfactory, then try to prevent starving or blocking by increasing or decreasing the production rate accordingly.
- If the surplus level is not satisfactory, meaning that is either too low or too high, then produce at maximum or zero rate, respectively.

A buffer tends to be empty when the upstream machine is either under repair or producing at a slower rate than the downstream machine. Similarly, a buffer tends to fill when the downstream machine is either under repair or producing at a slower rate than the upstream machine. The controllers keep buffers neither full nor empty regulating the machine rates. When a buffer tends to be full, the controller is increasing the rate of the downstream machine and decreasing the rate of the upstream machine. In the same way, when a buffer tends to be empty, the controller is increasing the rate of the upstream machine and decreasing the rate of the downstream machine. The information needed to synchronize the operation of the production network is transferred to each control module by the level change of each buffer. Every event occurring in the production network is affecting level of buffers close to the area of the event. In that way,

the production system is operating at satisfactory rates while the WIP is kept at low levels.

In fuzzy controllers, the control policy is described by linguistic IF–THEN rules with appropriate mathematical meaning [10] (see also Appendix A). A rule antecedent (IF part) describes conditions under which the rule is applicable and forms the composition of the inputs. The consequent (THEN part) gives the response or conclusion that should be taken under these conditions. A two-input (antecedent) rule of the Mamdani type has the form [10], [17]

IF X is A AND Y is B , THEN Z is C

where X, Y are the input and Z is the output variable, and A, B , and C their linguistic variations, respectively, that are fuzzy sets with certain membership functions.

The inference procedure may be briefly described as follows. Let x^*, y^* be the numerical values of the input variables that are converted into fuzzy sets, with membership functions denoted by $\mu_A(x^*)$ and $\mu_B(y^*)$. These functions are compared with fuzzy sets A and B (of the antecedent part) and determine the output value of the fuzzy set C . The outputs of the activated rules are aggregated, forming the value of the overall control output, which is then defuzzified into a crisp number C^* .

The *line-control module* contains rules of the following form:

IF $b_{j,i}$ is $LB^{(k)}$ AND $b_{i,l}$ is $LB^{(k)}$ AND s_i is $LS_i^{(k)}$
AND x_i is $LX^{(k)}$, THEN r_i is $LR_i^{(k)}$ (1)

where k is the rule number ($k = 1, \dots, K$), i is the number of machine or workstation, LB is a linguistic value of the variable *buffer level* b with term set $B = \{\text{Empty, Almost Empty, OK, Almost Full, Full}\}$, s_i denotes the state of machine i , which can be either 1 (operative) or 0 (stopped), and consequently, $S = \{0, 1\}$. LX represents the value that surplus x takes, and it is chosen from the term set $X = \{\text{Negative, OK, Positive}\}$. The *production rate* r takes linguistic values LR from the term set $R = \{\text{Zero, Low, Almost Low, Normal, Almost High, High}\}$.

Consider now that a machine is not stopped, and the actual buffer levels of the upstream and downstream buffers can be represented as $b_{j,i}^*$ and $b_{i,l}^*$ with membership functions $\mu_B^*(b_{j,i})$ and $\mu_B^*(b_{i,l})$, respectively. The production surplus at a given time instant is denoted as x_i^* with membership function $\mu_X^*(x_i)$. The production rate r_i^* , the control action, at every time instant is given by

$$r_i^* = \frac{\sum r_i \mu_R^*(r_i)}{\sum \mu_R^*(r_i)} \quad (2)$$

where $\mu_R^*(r_i)$ is the membership function of the aggregated production rate, which is given by

$$\mu_R^*(r_i) = \max_{b_{j,i}, b_{i,l}, x_i} \min[\mu_{\text{AND}}^*(b_{j,i}, b_{i,l}, x_i), \mu_{FR^{(k)}}^*(b_{j,i}, b_{i,l}, x_i, r_i)] \quad (3)$$

where $\mu_{\text{AND}}^*(b_{j,i}, b_{i,l}, x_i)$ is the membership function of the inputs, and $\mu_{FR^{(k)}}^*(b_{j,i}, b_{i,l}, x_i, r_i)$ is the membership function of

the k th activated rule. \max / \min is denoted the selected inference procedure [10]. Similarly, the generic rules of the *assembly* and *disassembly control modules* may be written as follows:

IF $b_{j,i}$ is $LB^{(k)}$ AND \dots AND $b_{k,i}$ is $LB^{(k)}$ AND
 $b_{i,l}$ is $LB^{(k)}$ AND s_i is $LS_i^{(k)}$ AND x_i is $LX^{(k)}$,
THEN r_i is $LR_i^{(k)}$ (4)

IF $b_{j,i}$ is $LB^{(k)}$ AND $b_{i,k}$ is $LB^{(k)}$ AND \dots AND
 $b_{i,l}$ is $LB^{(k)}$ AND s_i is $LS_i^{(k)}$ AND x_i is $LX^{(k)}$,
THEN r_i is $LR_i^{(k)}$. (5)

A very brief summary regarding the involved terms is given in Appendix A.

This formulation is expanded to multiple-part-type production systems without major modifications. In multiple-part-type systems, machines are divided into submachines producing only one part type. Thus, a multiple-part-type system may be decomposed into as many single-part-type systems as the number of parts produced. The structure of the fuzzy controller remains the same.

This production-control approach has given very good results, compared with other production-control approaches, in reducing WIP and cycle time, while satisfying demand [19]. One of its major advantages is that due to its modularity and distributivity (one controller at each production module), it is easily implemented in production systems of almost any geometry and magnitude. This advantage can be seen, from another point of view, as a drawback. The distributed controllers use mainly local information concerning the neighborhood of the controlled machine. The basic assumption is that information is propagated through the production system by the changes of buffer levels. This is a rather slow way to transmit information through complex production networks. Further, the information concerning the overall production-system performance, which is necessary for the optimization of the system's operation, cannot be obtained in this way. On the other hand, production surplus is giving a more precise picture of the system's state. If it is negative, customers are not satisfied. If it is positive and large, WIP is high. Still, the surplus-based approach is not using any information about the overall system's performance as total WIP and backlog. This kind of information could be valuable in an effort to improve the system's performance and its ability to adapt to demand changes. To achieve this goal, a supervisory controller is used.

III. SUPERVISORY CONTROLLER

In control-systems literature, a supervisor is a controller (supervisory controller) that uses available data to characterize the overall system's current behavior, potentially modifying the lower level controllers to ultimately achieve desired specifications. In addition, the supervisor may be used to integrate other information into the control decision-making process. It may incorporate certain user inputs, or inputs from other subsystems. Given information of this type, the supervisor can seek to tune the supervised controller to achieve better performance [17].

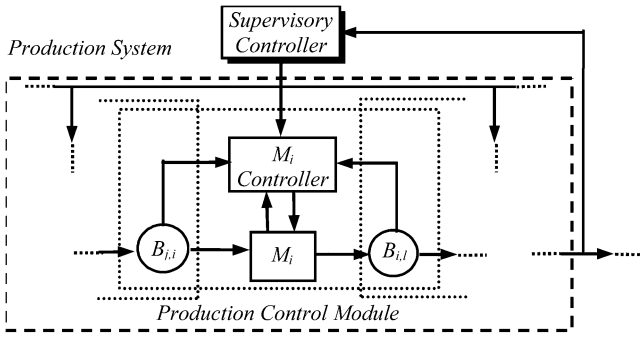


Fig. 2. Supervisory control architecture.

The supervisory controller in this paper is tuning the distributed fuzzy controllers in a way that some performance measures are improved without a dramatic change in the structure of the control architecture. The proposed control approach remains modular, since the production-control modules are not modified, but simply tuned by the additional supervisory controller. The proposed supervisory control architecture is shown in Fig. 2.

The input variables of the supervisor are:

- the mean surplus of the end product mx_e ;
- the error of the end-product surplus e_x , which is the difference between the end-product surplus x_e and the initial lower bound of surplus I_l ;
- the relative WIP error e_w , which is

$$e_w = \frac{\text{WIP}(t) - \overline{\text{WIP}(t)}}{\overline{\text{WIP}(t)}} \quad (6)$$

where $\overline{\text{WIP}(t)}$ is the mean WIP (including the end-product buffer level) of the production system until time t , interpreted and used as a target value. Relative WIP error e_w is used as a measure of WIP performance, since an analytical measure of the optimal mean WIP cannot be assessed. This is based on the assumption that WIP needed for the smooth operation of the production system is approximately equal to its mean value, and large deviations from it should be avoided.

The supervisory controller output variables are:

- the production surplus upper-bound correction factor u_c ;
- the surplus lower-bound correction factor l_c , where $-1 \leq l_c, u_c \leq 1$.

These correction factors express the percentage by which surplus bounds are altered. Production surplus is divided into three areas. If the surplus is lower than a lower surplus bound l_b , then the machine is producing at maximum rate. If the surplus is above the upper bound u_b , then production is stopped. When surplus is between lower and upper bound, the production rate is decided in relation with the adjacent buffer levels and machine state. The production surplus bounds are modified according to the following mechanism:

$$\begin{aligned} u_b &= I_u + u_c n_u + \min(x_e, 0) \\ l_b &= \min[(I_l + l_c n_l), u_b] \end{aligned} \quad (7)$$

where I_u and I_l are the initial upper and lower surplus bounds, respectively, and n_u, n_l are constants chosen in such a way that l_b will never exceed u_b when x_e is positive.

Adapting surplus bounds may improve the production system's performance, especially when some of its parameters are dynamically modified, as they would be in the case of time-dependent demand.

The first machine surplus x_1 provides valuable information about the amount of material released in the production system. By controlling x_1 , one may control material released in the system and, consequently, WIP. Thus, u_b is, in fact, an upper bound of x_1 . If N is the total amount of parts in process, then $x_1 = N + x_e$. When the end-product surplus x_e is negative, it is necessary to add it in the upper surplus bound u_b , because x_1 can then become very small, even negative. In that case, u_b without x_e added would not be an effective WIP control tool. The lower surplus bound l_b represents the desired level of x_e . If x_e is very high, in normal operation conditions, it is unlikely to fall under zero when a perturbation, as a machine failure or starvation, occurs, and thus, backlog is minimized. Unfortunately, this is achieved at the cost of increased WIP, so there must be a tradeoff between WIP and backlog. If x_e is far below zero, it is possible for u_b to become smaller than l_b ; in order to prevent that, l_b is chosen as the minimum of u_b and $I_l + l_c n_l$.

The rule base of the supervisory controller contains rules of the following form (see also Appendix B):

$$\begin{aligned} \text{IF } mx_e \text{ is } LMX^{(k)} \text{ AND } e_x \text{ is } LE_x^{(k)} \text{ AND } e_w \text{ is } LE_w^{(k)}, \\ \text{THEN } u_c \text{ is } LU_c^{(k)} \text{ AND } l_c \text{ is } LL_c^{(k)} \end{aligned} \quad (8)$$

where k is the rule number ($k = 1, \dots, K$), i is the number of machine or workstation; LMX is a linguistic value of the variable mean end-product surplus with term set $MX = \{\text{Negative Big, Negative Small, Zero, Positive Small, Positive Big}\}$; e_x denotes the error of end-product surplus, which is the difference between surplus x_e and the lower bound of surplus; and the term set of the corresponding linguistic value is $E_x = \{\text{Negative, Zero, Positive}\}$. LE_w represents the relative deviation of WIP from its mean value, and it is chosen from the term set $E_w = \{\text{Negative, Zero, Positive}\}$. The upper surplus bound correction factor takes linguistic values LU_c from the term set $U_c = \{\text{Negative, Negative Zero, Zero, Positive Zero, Positive}\}$, and the lower surplus bound correction factor takes linguistic values LL_c from the term set $L_c = \{\text{Negative, Negative Zero, Zero, Positive Zero, Positive}\}$. The rule given in (8) is equivalent to the two rules

$$\begin{aligned} \text{IF } mx_e \text{ is } LMX^{(k)} \text{ AND } e_x \text{ is } LE_x^{(k)} \text{ AND} \\ e_w \text{ is } LE_w^{(k)}, \text{ THEN } u_c \text{ is } LU_c^{(k)} \end{aligned} \quad (9)$$

$$\begin{aligned} \text{IF } mx_e \text{ is } LMX^{(k)} \text{ AND } e_x \text{ is } LE_x^{(k)} \text{ AND} \\ e_w \text{ is } LE_w^{(k)}, \text{ THEN } l_c \text{ is } LL_c^{(k)}. \end{aligned} \quad (10)$$

This means that one may use two single-output fuzzy controllers to imitate the behavior of a single controller with the same two outputs. This is what is followed in the sequel. First, the case where the controller output is the upper surplus bound

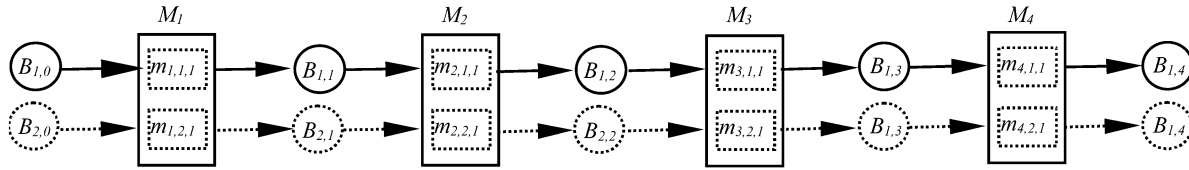


Fig. 3. Test Case 1 production line: multiple part types.

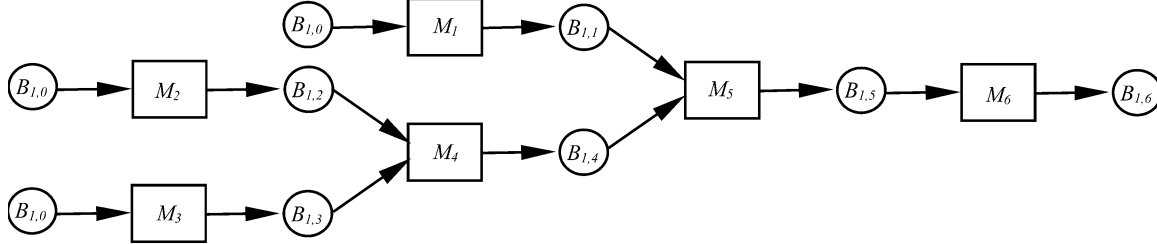


Fig. 4. Test Case 2 production network: single part type.

correction factor is examined. The mathematical meaning of the k th rule can be given as a fuzzy relation $FR^{(k)}$ on $MX \times E_x \times E_w \times U_c$, which in the membership function domain is

$$\mu_{FR^{(k)}}(mx_e, e_x, e_w, u_c) = f_{\rightarrow} \left[\mu_{LMX^{(k)}}(mx_e), \mu_{LE_x^{(k)}}(e_x), \mu_{LE_w^{(k)}}(e_w), \mu_{LU_c}(u_c) \right] \quad (11)$$

where f_{\rightarrow} denotes the implication operator, which is the min operator for rules of the Mamdani type [10].

Consider now that the actual mean surplus of the end product can be represented as mx_e^* , the error of surplus as e_x^* , and the relative error of WIP as e_w^* with membership functions $\mu_{MX}^*(mx_e)$, $\mu_{E_x}^*(e_x)$ and $\mu_{E_w}^*(e_w)$, respectively. The membership function of the conjunction of the three inputs, for AND = min, is

$$\mu_{AND}^*(mx_e, e_x, e_w) = \mu_{MX}^*(mx_e) \wedge \mu_{E_x}^*(e_x) \wedge \mu_{E_w}^*(e_w). \quad (12)$$

The upper bound correction factor u_c^* , e.g., the control action at every time instant, is given by

$$u_c^* = \frac{\sum u_c \cdot \mu_{U_c}^*(u_c)}{\sum \mu_{U_c}^*(u_c)} \quad (13)$$

where $\mu_{U_c}^*(u_c)$ is the membership function of the aggregated upper bound correction factor, which is computed by applying the max/min composition on the outcome of (11) and (12). That is

$$\mu_{U_c}^*(u_c) = \max_{mx_e, e_x, e_w} \min [\mu_{AND}^*(mx_e, e_x, e_w), \mu_{FR^{(k)}}(mx_e, e_x, e_w, u_c)]. \quad (14)$$

The fuzzy system representing the second output of the fuzzy controller is formulated similarly.

Since a multiple-part-type production system may be decomposed into single-part-type systems, as many supervisors as the number of part types may be used.

IV. SIMULATION TESTING AND RESULTS

The proposed supervisory-control approach is tested and compared with the unsupervised/distributed production-control approach introduced in [19]. Three test cases are considered. A multiple-part-type production line and two production networks, as shown in Figs. 3–5. The assumptions made for all simulations are the following.

- Machines fail randomly with a failure rate p_i .
- Machines are repaired randomly with rate rr_i . Unlimited repair personnel is assumed. There is always somebody to start working on a failed machine.
- Time to failure and time to repair are exponentially distributed.
- Demand is stochastic with rate d_i and follows the Poisson distribution.
- All machines operate at known, but not necessarily equal rates.
- The initial buffers are infinite sources of raw material, and consequently, the initial machines are never starved.
- Buffers between adjacent machines M_i, M_j have finite capacities.
- Set-up times or transportation times are negligible or are included in the processing times.

Each machine i performs k operations on the j part type. If the production network under consideration is a reentrant one, there may be more than one operation of the same part type performed in each machine. Each machine is “virtually” divided into as many submachines as the number of operations to be performed. Thus, submachine $m_{i,j,k}$ represents the controller of machine M_i regulating the execution of operation k on parts of type j . Submachines are shown with dotted-line squares in the corresponding figures (Figs. 3 and 5). Each part type is divided into l items according to the stage of process. Items represent the operations performed through the production network to the parts of a specific type. There is a buffer for every item of a part type, thus $B_{j,l}$ is the buffer of item l of part type j . Each submachine and the adjacent buffers form a production-control module, as described in a previous section. By this formulation, a multiple-part-type production network is decomposed into as many single-part-type networks as the number of part types produced.

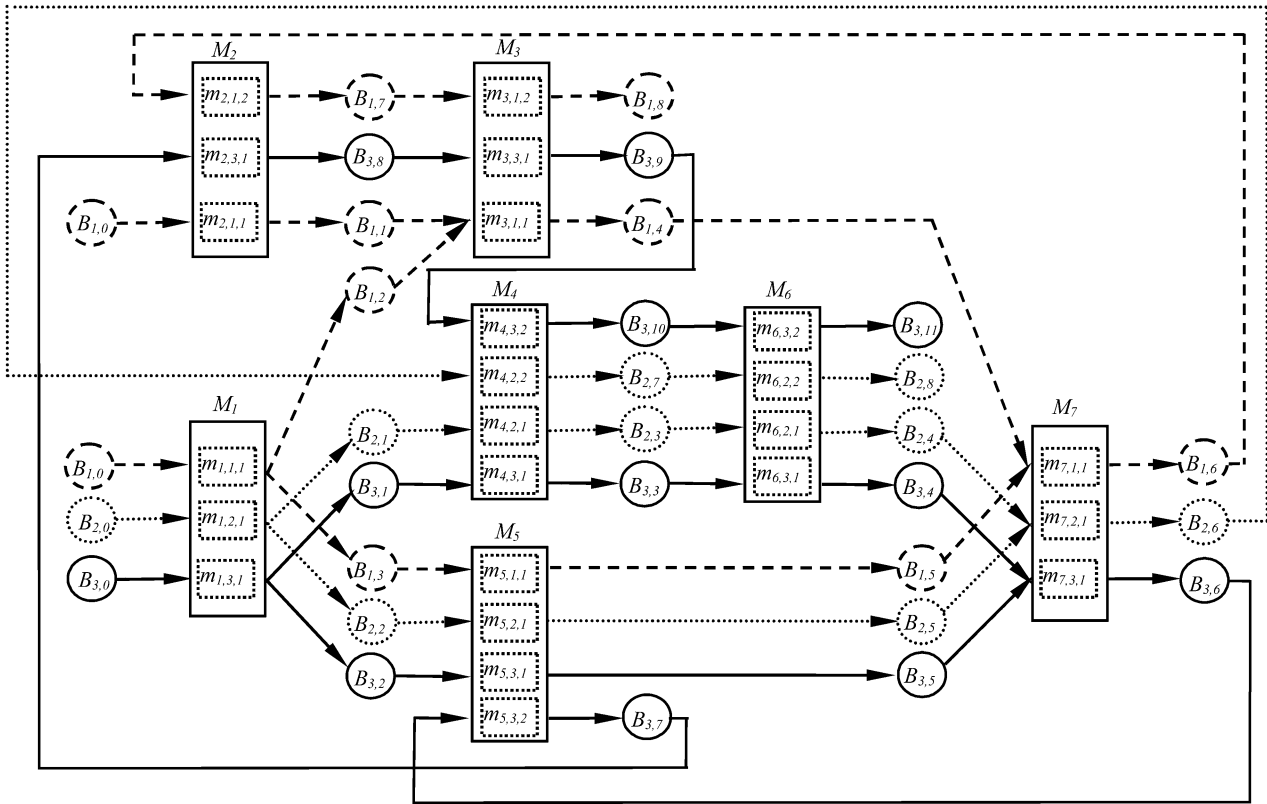


Fig. 5. Test Case 3 production network: multiple part types and reentrant flow.

A problem arising in multiple-part-type systems is the distribution of machine operating times to the different part types, and consequently, the decision of the production rate of every submachine. The production rate $pr_{i,j,k}$ of submachine $m_{i,j,k}$ is decided as follows:

$$pr_{i,j,k} = \left[A_{i,j,k} r_{i,j,k} + \left(1 - \sum_J \sum_{K_j} A_{i,j,k} r_{i,j,k} \right) G_{i,j,k} \right] \frac{1}{\tau_{i,j,k}} \quad (15)$$

where

$\tau_{i,j,k}$ process time of operation k of part type j in machine M_i ;

$r_{i,j,k}$ a number in $[0, 1]$ representing the percentage of the maximum possible production rate of part's j operation k in machine i .

d_j mean demand rate of j part type until time t ;

J set of part types processed in machine M_i ;

K_j set of operations of part type j performed in M_i ;

$A_{i,j,k}$ is given by

$$A_{i,j,k} = \frac{d_j \tau_{i,j,k}}{\sum_J \sum_{K_j} d_j \tau_{i,j,k}} \quad (16)$$

$G_{i,j,k}$ is given by

$$G_{i,j,k} = \begin{cases} \frac{d_j \tau_{i,j,k} U_{i,j,k}}{\sum_J \sum_{K_j} d_j \tau_{i,j,k} U_{i,j,k}}, & \text{for } U_{i,j,k} = 1 \\ 0, & \text{for } U_{i,j,k} = 0 \end{cases} \quad (17)$$

$U_{i,j,k}$ is given by

$$U_{i,j,k} = \begin{cases} 1, & \text{for } r_{i,j,k} = 1 \\ 0, & \text{for } r_{i,j,k} < 1 \end{cases} \quad (18)$$

where $r_{i,j,k} \in [0, 1]$.

The machine's total operating time is distributed to individual operations according to a fixed fraction of production capacity, which is $A_{i,j,k}$. The percentage of time machine i should devote in order to meet demand d_j is equal to $d_j \tau_{i,j,k}$. The sum $\sum d_j \tau_{i,j,k}$ gives the percentage of time machine M_i should be working in order to meet demand. This percentage should be less than or equal to the percentage of time M_i is working.

If $r_{i,j,k} < 1$, we get $pr_{i,j,k}$ by multiplying $A_{i,j,k}$ with the outcome of fuzzy controller $r_{i,j,k}$.

The remaining operating time of the machine i , if any, is distributed to the operations with the highest priority. The highest priority is given to the operations having fuzzy rate $r_{i,j,k}$ equal to one. This means that the specific operation should be processed in the maximum feasible rate. Distribution of the remaining operating time is done with the use of $G_{i,j,k}$. This is a fraction similar to $A_{i,j,k}$. The difference between them is the index $U_{i,j,k}$, which identifies the operations with the highest priority.

The loading times for each machine are determined by a heuristic policy known as the staircase strategy [11]. If the machine is available, load the part having the maximum positive difference between the integral of production rate and the actual cumulative production.

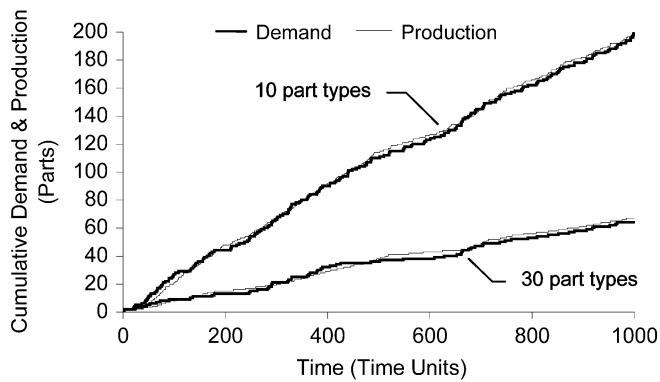


Fig. 6. Test Case 1: Cumulative production and demand.

MATLAB’s Fuzzy Logic Toolbox [13] and Simulink have been the software tools for building and testing all simulations. The performance of the fuzzy supervisory-control approach is evaluated through a series of test cases. The supervised fuzzy control approach is compared with the distributed fuzzy approach, which has given very good results, compared with other production-control approaches [19].

A. Test Case 1: Multiple-Part-Type Production Lines

The developed fuzzy supervisory-control system is first tested for the case of a multiple-part production line, presented in Fig. 3. The production line under consideration consists of four machines producing two product types. The failure and repair rates are equal for all machines. The repair rates are $rr_i = 0.5$. The processing times are also equal for all machines and product types ($\tau_{i,j,k} = 0.325, i = 1, \dots, 4, j = 1, 2, k = 1$).

Fig. 6 presents the evolution of cumulative demand and production of part type 1 for the supervised-control scheme for 10 ($j = 1, \dots, 10, d_j = 0, 2$) and 30 ($j = 1, \dots, 30, d_j = 0, 066$) different product types. Results for the other part types are similar.

Comparative results for the WIP and backlog of part type 1 in relation to demand for various buffer capacities (BC) are shown in Fig. 7(a)–(b). Results for WIP and backlog of part type 1 in relation to demand for various failure rates are shown in Fig. 7(c)–(d). Results for the other part types are similar (and, thus, omitted).

B. Test Case 2: Single-Part-Type Production Networks

The second test case is the single-part production network presented in Fig. 4. The production system under consideration consists of six machines producing one part type. The failure and repair rates of all machines are equal. The repair rates are $rr_i = 0.5$. All assembly and disassembly factors are assumed equal to one. The processing times are $\tau_{i,j,k} = 0.3$ (here $i = 1, \dots, 6, j = 1, k = 1$) equal for all machines. Comparative results for the WIP and backlog in relation to demand for various BC are shown in Fig. 8(a)–(b). Results for WIP and backlog in relation to demand for various failure rates are shown in Fig. 8(c)–(d).

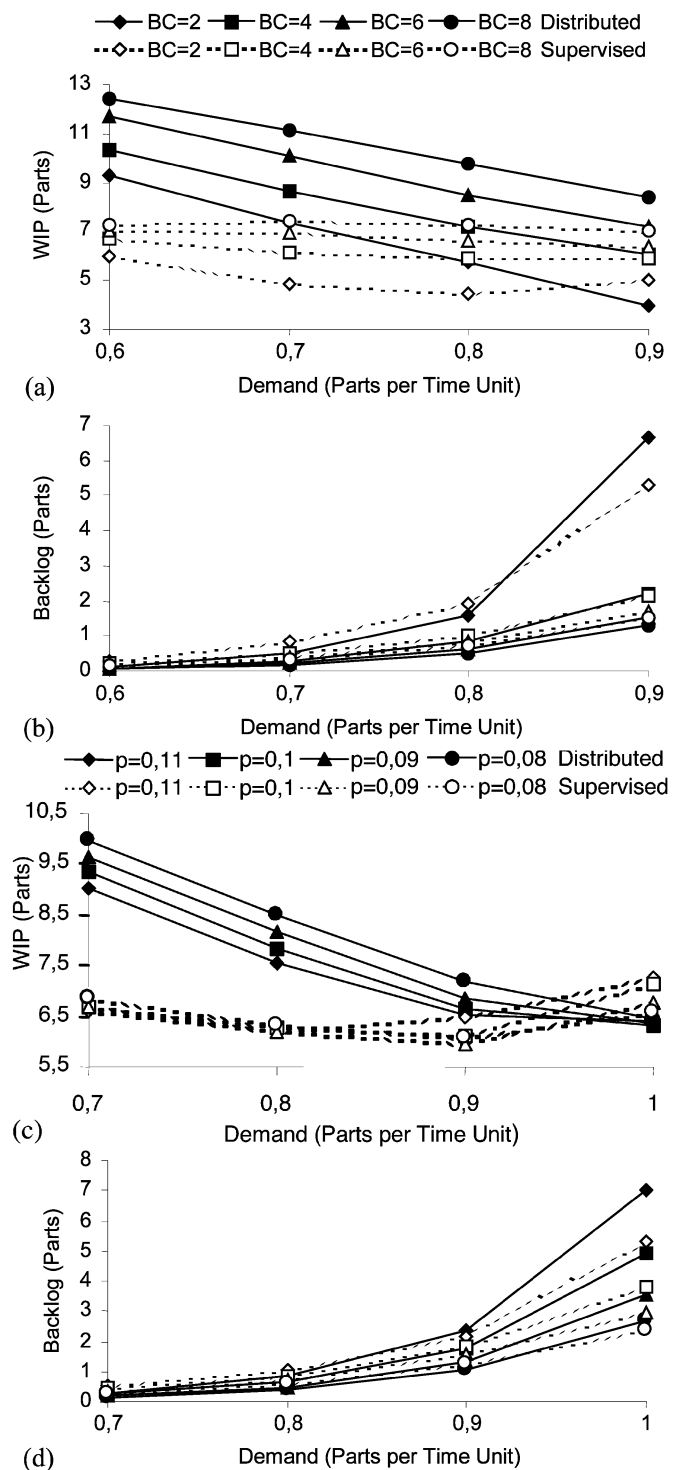


Fig. 7. Test Case 1. (a) WIP versus demand for various BC. (b) Backlog versus demand for various BC. (c) WIP versus demand for various failure rates (p). (d) Backlog versus demand for various p.

C. Test Case 3: Multiple-Part-Type Reentrant Production Networks

The developed fuzzy supervisory control system is also tested for the case of a multiple-product reentrant production network presented in Fig. 5. The production system under consideration consists of seven machines and produces three product types.

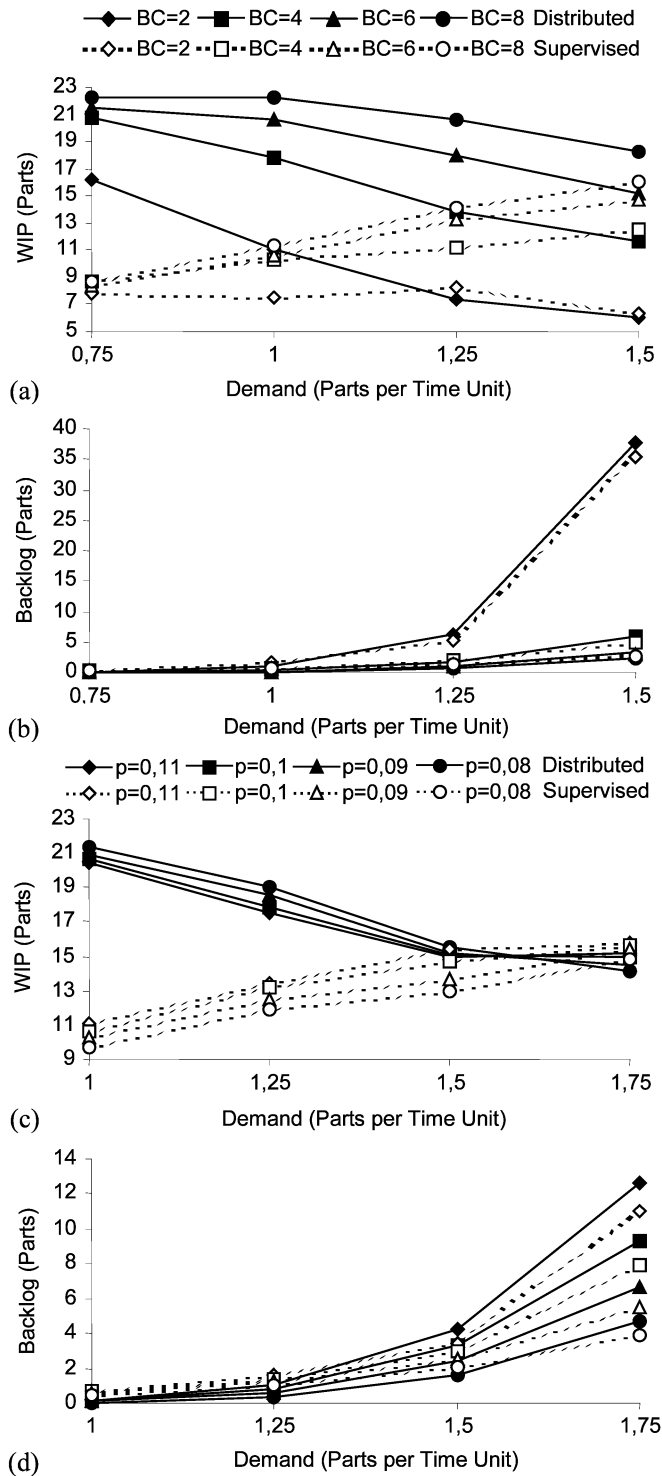


Fig. 8. Test Case 2. (a) WIP versus demand for various BC. (b) Backlog versus demand for various BC. (c) WIP versus demand for various failure rates (p). (d) Backlog versus demand for various p .

All intermediate BC are equal to $BC_{j,l} = 10$, and the end-product buffers have capacities equal to 20. The failure and repair rates of all machines are $p_i = 0.1$ and $rr_i = 0.5$, respectively (here $i = 1, \dots, 7, j = 1, 2$). All the assembly and disassembly factors are assumed equal to one. The machine processing times are shown in Table I.

Comparative results for the WIP, cycle time, and backlog in relation to demand are shown in Fig. 9(a)–(c).

D. Production Cost Analysis

To get a better picture of the significance of the results, a cost analysis is carried out. The production cost associated with the proposed control architecture consists of inventory and backlog costs. Inventory costs are due to the capital invested for the purchase of raw material and the added value from the parts processing. Thus, the cost of each unit of inventory per time unit, noted as c_I , should be different for each part item. It is assumed that inventory cost is independent from the stage of process. Further, it is assumed that WIP is not suitable for cost assessment, because the production system examined is an assembly/disassembly network, and thus, the various items of a part do not correspond to the same amount of raw material. To assess inventory cost, we need to use the number of i -type parts in system N_i which is the number of end parts that can be produced by the material in the system, or one may say, the number of parts released in the system. Thus, the mean production cost C is given by

$$C = c_I N + c_b BL \quad (19)$$

where c_I, c_b are the unit costs of inventory and backlog, respectively, and N is the mean number of parts in the system, which is the number of final products that will be produced by the material in the system. BL is the mean backlog.

The cost analysis results for part type 1 of the production network examined in Test Case 3 are presented in Fig. 10. The supervisory-control relative cost is presented, which is the supervisory-control cost C_s , divided by the distributed-control cost C_d , in relation to the inventory cost ratio, which is $c_I/(c_I + c_b)$. Cost analysis of other part types gives similar results.

E. Statistical Significance of Results

Ten simulation runs of 10 000 time units each have been carried out. Table II presents the maximum relative errors of mean WIP estimates, with level of significance $\alpha = 0.05$, for each test case.

F. Discussion

Based on the obtained results, the following observations are made.

The supervised approach achieves a substantial reduction of WIP when demand rate is low. As demand increases, the difference between the WIP of the two approaches is reduced, and finally, there is a small increase of WIP attained by the supervised approach when demand reaches system productivity [see Figs. 7(a) and (c), 8(a) and (c), and 9(b)].

These results are due to the fact that WIP and backlog are competitive measures of the production system's performance. In order to reduce backlog, one has to increase the system's throughput, and thus, increase WIP. The use of supervisory control achieves a clear improvement for the most important performance measures. When demand is very high, one may consider that service rate, and thus, backlog, is more important than WIP.

When demand can be easily satisfied and backlog is at low levels, a substantial reduction of WIP may be more important than a small increase in backlog.

TABLE I
MACHINE PROCESSING TIMES

Part type	Operation	Machine						
		1	2	3	4	5	6	7
1	1	0.2	0.2	0.2	-	0.15	-	0.2
	2	-	-	-	-	-	-	-
2	1	0.2	0.2	0.2	0.15	0.15	0.15	0.2
	2	-	-	-	0.15	-	0.15	-
3	1	0.2	0.2	0.2	0.15	0.15	0.15	0.2
	2	-	-	-	0.15	0.15	0.15	-

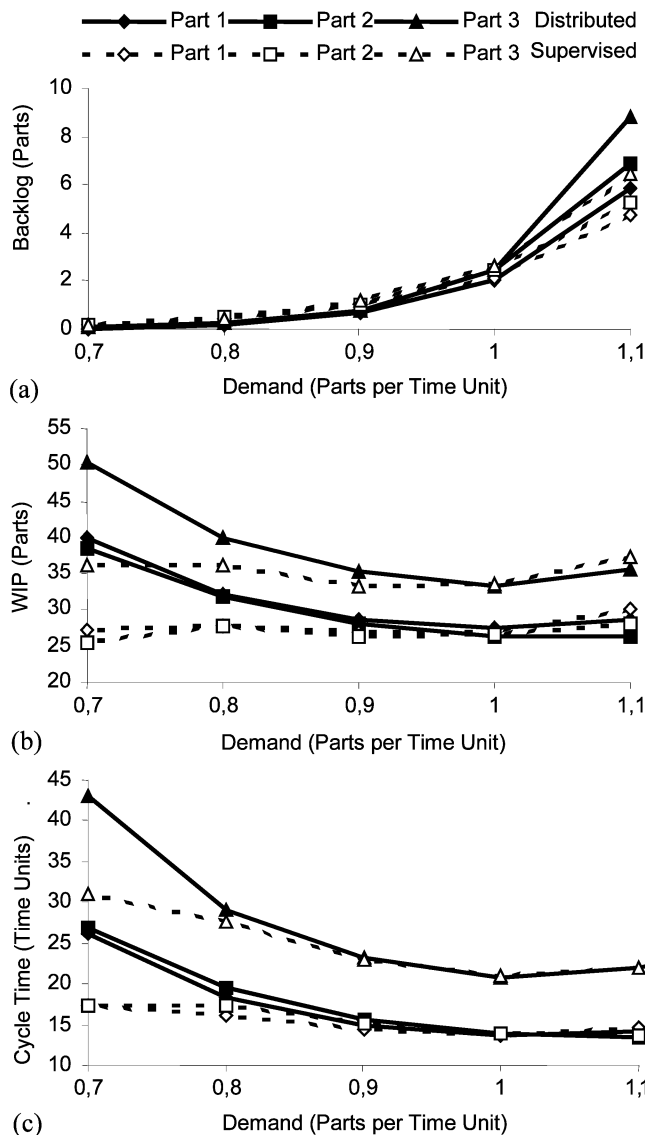


Fig. 9. Test Case 3. (a) Backlog versus demand. (b) WIP versus demand. (c) Cycle time versus demand.

This may be seen more clearly in the results of the production cost analysis. Observing Fig. 9, for Test Case 3, when demand is low (less than one), the distributed approach gives better results only when the inventory cost ratio is very small. This means that the distributed approach gives a better cost only when the inventory unit cost c_I is much smaller than the backlog unit cost c_b . On the other hand, when demand is close to the system's

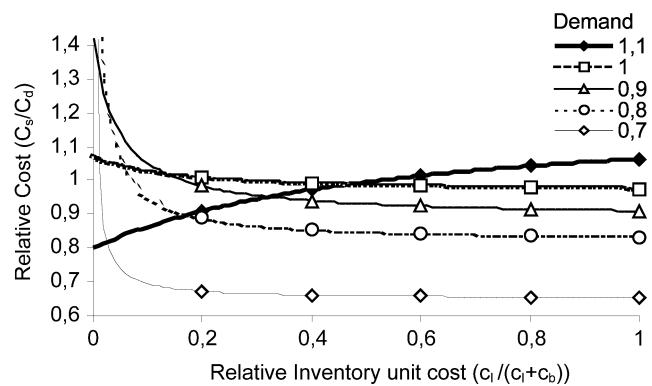


Fig. 10. Cost analysis for part type 1 of Test Case 3.

TABLE II
MAXIMUM RELATIVE ERRORS OF MEAN WIP ESTIMATES

	Distributed (%)	Supervised (%)
Test Case 1	2.95	2.05
Test Case 2	1.8	1.25
Test Case 3	3	2.45

capacity (demand is equal to 1.1), the supervised approach gives lower cost when c_I is smaller than c_b .

Another important observation is that the supervised-control approach is not affected by changes of important system parameters, such as BC and failure rates. In Test Cases 1 and 2, we perturbed the production system configuration by changing BC and failure rates (see Figs. 7 and 8). In all cases, the supervised approach reduced WIP when demand was low and increased backlog when demand was high. This is due to the fact that a change in BC and/or in failure rates results in a change of system productivity.

Even when the number of part types produced is high, the production-control scheme can satisfy the demand, as one may observe in Fig. 6.

In Table II, one may see that the supervised approach always gives smaller relative error in the estimation of WIP than the distributed approach. This indicates that the supervised approach reduces the variation of important system measures. This behavior was expected, since the goals of the supervisory controller are to keep the mean end-product surplus close to zero and WIP close to its mean value.

The performance improvement described above is not achieved at the cost of a significant increase in control-system complexity. The control approach remains modular and distributive. As we have mentioned before, we need as many

supervisory controllers as the number of parts produced by the production system. In the production network of Test Case 3, for example, we need only three extra supervisory controllers. Insignificant increase of system's computational complexity is also demonstrated by the small difference in time needed to simulate the operation of the examined production network in a PC. It took about 22 hours to simulate the supervised production network of Test Case 3, for 10 000 time units on a Pentium III PC operating at 600 MHz, while the simulation of the unsupervised system required approximately 20 hours on the same machine. All experiments have been carried out using MATLAB's Simulink.

V. CONCLUSION

A supervisory fuzzy control approach which is used to tune a distributed fuzzy control architecture (presented in previous work) has been presented. Tuning of the distributed fuzzy controllers is made in a way that performance measures which are more important in every case are improved. More specifically, WIP and cycle times are substantially reduced, while backlog is kept to low levels, when demand is low. On the other hand, when demand is close to the system's capacity, backlog is reduced without a dramatic increase of WIP. Simulation results for a series of production systems with stochastic demand have shown noticeable improvement of performance and production-related costs, in most cases. These results are achieved while modularity and distributivity of the control architecture is maintained.

In the future, it would be very interesting to consider the case of time-varying demand and to include production costs due to WIP and backlog in the supervisory-control scheme in an effort to minimize them. Another interesting extension would be the integration of the proposed approach with a design mechanism, in order to optimize the selection of important system parameters which are now selected by tuning and seeking further improvement of the supervisor's performance. Some of these parameters could be BC and initial surplus bounds. This could be done with the use of reinforcement-learning control methods.

APPENDIX A

Let X be a collection of objects, called the universal set, whose elements are denoted by x . Then, a fuzzy set A in X is defined as

$$A = \{(x, \mu_A(x))\}, \quad x \in X, \quad \mu_A(x) \in [0, 1]$$

where $\mu_A(x)$ is the membership function of x in A . The membership function denotes the degree to which x belongs in A . The closer the value of $\mu_A(x)$ is to one, the more x belongs to A . Membership functions are not unique, as different people might define various membership functions for the same fuzzy set.

A fuzzy conditional statement, or a fuzzy IF-THEN rule, is an expression of the type "IF x is A THEN y is B ," denoted $A \rightarrow B$, where A and B are values of the linguistic variables x and y . A linguistic variable is mainly characterized by: 1) its linguistic values (or linguistic variations); 2) the physical domain over which the variables, e.g., x, y , take their quantitative values; and 3) the membership function of the linguistic values,

TABLE III
RULEBASE OF FUZZY SUPERVISORY CONTROLLER

Rule	Inputs			Outputs	
	LMX	LE_x	LE_w	LU_c	LL_c
1	Negative Big	Any	Any	Positive	Positive
2	Negative Small	Negative	Negative	Zero	Positive
3	Negative Small	Negative	Zero	Positive Zero	Positive
4	Negative Small	Negative	Positive	Negative Zero	Positive
5	Negative Small	Zero	Negative	Zero	Positive Zero
6	Negative Small	Zero	Zero	Zero	Positive Zero
7	Negative Small	Zero	Positive	Negative Zero	Positive Zero
8	Negative Small	Positive	Negative	Zero	Positive Zero
9	Negative Small	Positive	Zero	Zero	Positive Zero
10	Negative Small	Positive	Positive	Negative Zero	Positive Zero
11	Zero	Negative	Negative	Zero	Zero
12	Zero	Negative	Zero	Zero	Zero
13	Zero	Negative	Positive	Negative	Zero
14	Zero	Zero	Negative	Zero	Zero
15	Zero	Zero	Zero	Zero	Zero
16	Zero	Zero	Positive	Negative	Zero
17	Zero	Positive	Negative	Negative Zero	Negative Zero
18	Zero	Positive	Zero	Negative Zero	Negative Zero
19	Zero	Positive	Positive	Negative	Negative Zero
20	Positive Small	Negative	Negative	Zero	Negative Zero
21	Positive Small	Negative	Zero	Zero	Negative Zero
22	Positive Small	Negative	Positive	Negative	Negative Zero
23	Positive Small	Zero	Negative	Negative Zero	Negative Zero
24	Positive Small	Zero	Zero	Negative Zero	Negative Zero
25	Positive Small	Zero	Positive	Negative	Negative Zero
26	Positive Small	Positive	Negative	Negative Zero	Negative
27	Positive Small	Positive	Zero	Negative Zero	Negative
28	Positive Small	Positive	Positive	Negative	Negative
29	Positive Big	Any	Any	Negative	Negative

e.g., $\mu_A(x), \mu_B(y)$. For example, *buffer level* can be regarded as a linguistic variable taking linguistic values such as *low*, *about low*, *average*, *high*, and so on. The physical domain of the variable buffer level is the set $\{0, 1, \dots, BC\}$, where BC denotes buffer capacity. Thus, $\mu_{average}(x)$ is the degree to which the level x of buffer belongs to the linguistic value (fuzzy set) *average*.

Fuzzy or *approximate reasoning* is an inference procedure that derives conclusions from a set of fuzzy IF-THEN rules and known facts (observations). Let A, A^* be fuzzy sets on X and B, B^* fuzzy sets on Y . The concept of fuzzy inference is illustrated as follows:

Premise 1 (Fact):	x is A^*	(Observation)
Premise 2 (Rule):	IF x is A , THEN y is B	(Rule)
Consequence:	y is B^*	(Conclusion)

The conclusion B^* induced by the observation " x is A^* ," and the fuzzy rule $A \rightarrow B$ is defined by

$$B^* = A^* \circ (A \rightarrow B)$$

where " \circ " is the max-min composition, which in the membership function domain is

$$\begin{aligned} \mu_{B^*}(y) &= \max_x \min[\mu_{A^*}(x), \mu_{A \rightarrow B}(x, y)] \\ &= \vee [\mu_{A^*}(x) \wedge \mu_{A \rightarrow B}(x, y)] \end{aligned}$$

with $\mu_{B^*}(y)$ the membership function of B^* , $\mu_{A^*}(x)$ the membership function of A^* , $\mu_{A \rightarrow B}(x_i, y_j)$ the membership function of the rule $A \rightarrow B$, and \vee, \wedge denote maximum and minimum, respectively.

For details on fuzzy control, the reader is referred to [10] and [17].

APPENDIX B

Table III presents the rulebase of the supervisory controller. The objective is to keep the mean final product surplus as close to zero as possible, and at the same time, to prevent large deviation of WIP from its mean value. When mean surplus is negative, the lower surplus bound is increased. The opposite happens when the mean surplus is positive. An increase of WIP level leads to a tighter upper surplus bound.

ACKNOWLEDGMENT

The authors would also like to thank the reviewers for their helpful comments.

REFERENCES

- [1] R. Akella and P. R. Kumar, "Optimal control of production rate in a failure-prone manufacturing system," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 116–126, Feb. 1986.
- [2] S. X. Bai and S. B. Gershwin, "Scheduling manufacturing systems with work-in-process inventory control: Formulation of single-part-type systems," in *Proc. 29th Conf. Decision and Control*, vol. 2, Dec. 1990, pp. 557–564.
- [3] —, "Scheduling manufacturing systems with work-in-process inventory control: Multiple-part-type systems," *Int. J. Prod. Res.*, vol. 32, no. 2, pp. 365–386, 1994.
- [4] T. Bielecki and P. R. Kumar, "Optimality of zero-inventory policies for unreliable manufacturing systems," *Oper. Res.*, vol. 36, no. 4, pp. 532–541, 1988.
- [5] J. A. Buzacott and J. G. Shanthikumar, *Stochastic Models of Manufacturing Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [6] R. Conway, W. Maxwell, J. O. McClain, and L. J. Thomas, "The role of work-in-process inventory in serial production lines," *Oper. Res.*, vol. 36, no. 2, pp. 229–241, 1988.
- [7] L. Custodio, J. Senteiro, and C. Bispo, "Production planning and scheduling using a fuzzy decision system," *IEEE Trans. Robot. Automat.*, vol. 10, pp. 160–168, Apr. 1994.
- [8] Y. Dallery and G. Liberopoulos, "Extended Kanban control system: Combining Kanban and base stock," *IIE Trans. Des. Manufact.*, vol. 32, no. 4, pp. 369–386, 2000.
- [9] E. Dretoulakis, S. Ioannidis, and N. C. Tsourveloudis, "Distributed intelligent control of unreliable manufacturing systems," in *Computational Intelligence and Applications*. Athens, Greece: World Sci. and Eng. Soc. Press, July 1999, pp. 267–271.
- [10] D. Driankov, H. Hellendoorn, and M. Reinfrank, *An Introduction to Fuzzy Control*. New York: Springer-Verlag, 1993.
- [11] S. B. Gershwin, *Manufacturing Systems Engineering*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [12] —, "Design and operation of manufacturing systems: The control-point policy," *IIE Trans.*, vol. 32, pp. 891–906, 2000.
- [13] N. Gully and J.-S. R. Jang, *Fuzzy Logic Toolbox User's Guide*. Natick, MA: The MathWorks, 1995.
- [14] S. X. C. Lou and P. Kager, "A robust production control policy for VLSI wafer fabrication," *IEEE Trans. Semiconduct. Manufact.*, vol. 2, pp. 159–164, Nov. 1989.
- [15] S. X. C. Lou, S. P. Sethi, and Q. Zhang, "Optimal feedback production planning in a stochastic two-machine flowshop," *Eur. J. Oper. Res.*, vol. 73, pp. 331–345, 1994.
- [16] S. X. C. Lou and G. Van Ryzin, "Optimal control rules for scheduling job shops," *Ann. Oper. Res.*, vol. 17, pp. 233–248, 1989.
- [17] K. M. Passino and S. Yurkovich, *Fuzzy Control*. Reading, MA: Addison-Wesley, 1998.
- [18] N. Srivatsan, "Synthesis of optimal policies for stochastic manufacturing systems," Ph.D. dissertation, Oper. Res. Center, Mass. Inst. Technol., Cambridge, MA, 1993.
- [19] N. C. Tsourveloudis, E. Dretoulakis, and S. Ioannidis, "Fuzzy work-in-process inventory control of unreliable manufacturing systems," *Information Sci.*, vol. 127, pp. 69–83, 2000.
- [20] G. J. Van Ryzin, S. X. C. Lou, and S. B. Gershwin, "Production control for a tandem two-machine system," *IIE Trans.*, vol. 25, no. 5, pp. 5–20, 1993.
- [21] L. Wein, "Scheduling semiconductor wafer fabrication," *IEEE Trans. Semiconduct. Manufact.*, vol. 1, pp. 115–130, Aug. 1988.
- [22] H. P. Wiendahl, *Load-Oriented Manufacturing Control*. Hanover, NH: Springer-Verlag, 1995.
- [23] H. Yan, S. X. C. Lou, S. P. Sethi, A. Gardel, and P. Deosthali, "Testing the robustness of two-boundary control policies in semiconductor manufacturing," *IEEE Trans. Semiconduct. Manufact.*, vol. 9, pp. 285–288, May 1996.



Stratos Ioannidis received the diploma and the M.S. degree in 1995 and 1997, respectively, from the Department of Production Engineering and Management, Technical University of Crete, Chania, Greece, where he is currently working toward the Ph.D. degree.

His research interests are in production systems design and control.



Nikos C. Tsourveloudis (A'90) received the B.S. and Ph.D. degrees in production engineering and management from the Technical University of Crete, Chania, Greece, in 1990 and 1995, respectively.

He was President of the Hellenic Association of Industrial Engineers from 1992 to 1995. During the academic years 1997 and 1998, he was a Visiting Research Scientist at the Center for Advanced Computer Studies, CACS, and the Apparel-Computer Integrated Manufacturing Center (A-CIM) of the University of Louisiana, Lafayette, LA. Currently, he

is an Assistant Professor with the Department of Production Engineering and Management and Director of the Machine Tools Laboratory at the Technical University of Crete. His research interests are in the areas of intelligent control, modeling and scheduling of flexible and computer-integrated manufacturing systems, autonomous operation/navigation of uninhabitant vehicles, virtual reality, and fuzzy-logic applications.

Dr. Tsourveloudis is a member of numerous professional and scientific organizations.



Kimon Valavanis (S'80–M'83–SM'91) received the Diploma in electrical and electronic engineering from the National Technical University of Athens, Athens, Greece in 1981, and the M.Sc. degree in electrical engineering, and the Ph.D. degree in computer and systems engineering from Rensselaer Polytechnic Institute (RPI), Troy, NY, in 1984 and 1986, respectively.

From 1986 to 1990 he held the Analog Devices Career Development Chair for Assistant Professors at the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA,

where he was also Director of the Robotics Laboratory. From 1991 to 1999, he was with The Center for Advanced Computer Studies (CACS) at the University of Louisiana, Lafayette, LA; he served as Associate Director for Research at the Apparel Computer Integrated Manufacturing (A-CIM) Center, Director of the Robotics and Automation Laboratory, and held the A-CIM/[TC]2/Board Of Regents Support Fund Regents Professorship in Manufacturing. In 1999, he joined the Department of Production Engineering and Management, Technical University of Crete, Chania, Greece, serving as Director of the Laboratory of Intelligent Systems and Robotics, Graduate Program Coordinator, and Chair of the University Industrial Advisory Board. He is currently a Professor with the Department of CSEE, University of South Florida, Tampa, FL, and with the Center for Robot-Assisted Search and Rescue. He holds Visiting/Adjunct positions with the Faculty of Electrical Engineering and Computing, Department of Telecommunications, University of Zagreb, Zagreb, Croatia, and at CACS at the University of Louisiana at Lafayette. His research interests are in robotic systems, automation, and distributed intelligence systems. He has published over 220 book chapters, technical journal/transactions and conference papers, and he has coauthored three books. He is also the Book Review Editor of the *Journal of Intelligent and Robotic Systems*.

Dr. Valavanis has served as the IEEE Robotics and Automation Society "Discrete Event Dynamic Systems Technical Committee" Cochair for two years and as an Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION. Since January 1996, he has been the Editor-in-Chief of the *IEEE Robotics and Automation Magazine*. He has been on the organizing committee of many IEEE conferences, serving as General, Program, Registration, and Local Arrangements Chair. In 1998, he was elected as Vice President, Administration of the IEEE Mediterranean Control Association (MCA). He was the General Cochair of the 2003 IEEE Mediterranean Conference in Control and Automation and the 2004 IEEE ICRA Program Chair. He is a former Fulbright Scholar (Senior Lecturing/Research Award) and a Distinguished Speaker in the IEEE Robotics and Automation Society.