Modeling and scheduling of production systems by using max-plus algebra

Al Bermanei, Hazem; Böling, Jari; Högnäs, Göran

Published in:
Flexible Services and Manufacturing Journal

DOI:
https://doi.org/10.1007/s10696-023-09484-z

Published: 06/03/2023

Document Version
Final published version

Flexible Services and Manufacturing Journal
https://doi.org/10.1007/s10696-023-09484-z

Modeling and scheduling of production systems by using max-plus algebra

Hazem Al Bermanei[1] · **Jari M. Böling**[2] · **Göran Högnäs**[3]

Accepted: 29 January 2023

Abstract
Max-plus algebra provides mathematical methods for solving nonlinear problems by using linear equations. These kinds of the problems arise in areas such as manufacturing, transportation, allocation of resources, and information processing technology. In this paper, the scheduling of production systems consisting of many stages and different units is considered, where some of the units can be used for many stages. If a production unit is used for different stages cleaning is needed in between, while no cleaning is needed between stages of the same type. Cleaning of units takes a significant amount of time, which is considered in the scheduling. The goal is to minimize the total production time, and such problems are often solved by using numerical optimization. In this paper a max-plus formalism is used for the modeling and scheduling of such production systems. Structural decisions such as choosing one unit over another proved to be difficult in the latter case, but this can be viewed as a switching max-plus linear system. No switching (and thus no cleaning) is considered as a base case, but for larger production batches the durability constraints will require switches. Switching as seldom as possible is shown to be optimal. Scheduling of a small production system consisting of 6 stages and 6 units is used as a case study.

Keywords Production systems · Scheduling · Discrete event systems · Switching max-plus linear systems

✉ Hazem Al Bermanei
hazem.al-bermanei@turkuamk.fi

[1] Faculty of Business ICT and Life Sciences, Turku University of Applied Sciences, Turku, Finland

[2] Process Control Laboratory, Faculty of Natural Sciences and Engineering, Åbo Akademi University, Turku, Finland

[3] Department of Mathematics, Faculty of Natural Sciences and Engineering, Åbo Akademi University, Turku, Finland

Published online: 06 March 2023

1 Introduction

Scheduling of manufacturing systems is a difficult task, since they consist of many units with complex relationships to and interdependences. In order to deal with this complexity, modeling and scheduling techniques are used to guarantee that the whole production process is executed in a more dynamic and reliable way than producing decisions manually. The questions of production scheduling in manufacturing processes is becoming more important considering the increasing demand of economic and environmental constraints. Therefore, in this paper, we consider industrial production scheduling problems for a manufacturing system consisting of parallel batch processes. Such processes can interact to each other and, therefore, influence the production of different batches in these processes. These kinds of problems can be represented using discrete event systems, which in general lead to a nonlinear description when using conventional algebra. Therefore, it is advantageous to use the max-plus technique which results in systems that are linear in the max-plus algebra. Because of this feature, we can use more effective methods that are available for modeling, simulation, and analysis of such systems. Scheduling of production systems is a common engineering problem, see for example Giffler and Thompson (1960) and Mutsaers et al. (2012). Scheduling using max-plus algebra has also been studied in Baccelli et al. (1992). van den Boom et al. (2006), Boom et al. (2020) show that a large class of scheduling problems can be handled using switching max-plus linear (SMPL) modelling. In this paper, the production scheduling in a deterministic manufacturing process consisting of 6 stages and 6 units done in parallel batches is considered as a case study of a SMPL system. The scheduling problem has earlier been studied by Björkqvist et al. (2002) using MILP, and in this paper much larger batch sizes are scheduled in a much shorter time.

The paper is organized as follows. Sections 2 and 3 present the background information on max-plus algebra and max-plus linear systems. The studied manufacturing system is described in Sect. 4. A max-plus model for the system is presented and solved. Properties of the continuous production case are studied in Sect. 5, and Sect. 6 gives concluding remarks.

2 Max-plus algebra

Max-plus algebra (Heidergott et al. (2006), De Schutter (2008), Cuninghame-Green (2004)) is a class of discrete algebraic systems, also known as an effective tool for modeling and analyzing several types of discrete event systems. In max-plus algebra we work with the max-plus semi-ring which is the $\mathbb{R}_{max} = \mathbb{R} \bigcup \{-\infty\}$ and the two binary operations addition $\oplus$ and multiplication $\otimes$ which are defined by:

$$a \oplus b = \max(a,\ b),\ a \otimes b = a + b \textit{ and } (-\infty) \otimes a = -\infty.$$

Define $\varepsilon = -\infty$ and $e = 0$. The additive and multiplicative identities are thus ε and e respectively and the operations are commutative, associative, and distributive

$$(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$$

as in conventional algebra.

Furthermore, the pair of operations $(\oplus, \otimes)$ can be extended to matrices and vectors similarly as in conventional linear algebra:

- For all $A, B \in \mathbb{R}_{max}^{m\times n}$, $(A \oplus B)_{ij} = a_{ij} \oplus b_{ij} = \max\left(a_{ij}, b_{ij}\right)$
- For $A \in \mathbb{R}_{max}^{m\times n}andB \in \mathbb{R}_{max}^{n\times p}$ define their product by

$$(A \otimes B)_{ij} = \oplus_{k=1}^{n}\left(a_{ik} \otimes b_{kj}\right) = \max_{k\in\{1,2,\ldots,n\}}\left(a_{ik} + b_{kj}\right) 1 \le i \le m,\ 1 \le j \le p$$

- The $n \times n$ identity matrix I_n in max-plus is defined as:

$$I_n = \begin{cases} e if i = j \\ \varepsilon if i \neq j \end{cases} \text{For } A \in \mathbb{R}_{\max}^{m\times n}, I_m \otimes A = A \otimes I_n = A.$$

- For a square matrix A and positive integer n the n^{th} power of A is written as:

$A^{\otimes n}$ and it is defined by $A^{\otimes n} = \underbrace{A \otimes A \otimes \ldots\ldots\ldots\ldots\ldots. \otimes A}_{n\text{times}}.$

- The eigenvalue λ and eigenvector v of the matrix A are defined as in ordinary liner algebra

$$A \otimes v = \lambda \otimes v$$

The power method (Schutter 2000, Braker and Olsder 1993) is usually used for finding the eigenvalue and eigenvector.

See also Heidegrott et al. (2006), De Schutter and van den Boom (2008), and Baccelli et al. (1992).

3 Max-plus linear systems

A discrete event system (DES) with synchronization and no concurrency can be modeled by a max-plus-algebraic model as in De Schutter (1996) and Al-Bermanei (2021):

$$x(k) = A \otimes x(k-1) \oplus B \otimes u(k) \tag{1}$$

$$y(k) = C \otimes x(k) \tag{2}$$

with $A \in \mathbb{R}_{\max}^{n\times n}, B \in \mathbb{R}_{\max}^{n\times m}$, and $C \in \mathbb{R}_{\max}^{p\times n}$ where n is the number of states, m is the number of inputs and p is the number of outputs. The vector x represents the state,

u is the input vector, and y is the output vector of the system. It is important to note that in Eqs. (1) and (2) the components of the input, the output, and the state are event times, and that the counter k in (1) and (2) is an event counter. For a manufacturing system, $u(k)$ would typically represent the time instants at which raw material is fed to the system for the k^{th} time, $x(k)$ the time instants at which the machines start processing the k^{th} batch of intermediate products, and $y(k)$ the time instants at which the kth batch of finished products leaves the system.

Due to the analogy with conventional linear time-invariant systems, a DES that can be modeled by Eqs. (1) and (2) will be called a max-plus linear time-invariant DES system. Typical examples of systems that can be modeled as max-plus linear DES are production systems, railroad networks, urban traffic networks, and queuing systems. Switching max-plus linear (SMPL) systems have been considered in van den Boom et al. (2006, 2020). Switching means in this case that the system switches between different A and B matrices.

We will now illustrate in detail how the scheduling of a manufacturing system can be done by using a switching max-plus linear model of the form (1) – (2), or actually all that is needed is $x(k) = A \otimes x(k-1)$.

4 The studied production system

Consider the production system shown in Fig. 1, the scheduling of which has been previously studied by Björkqvist et al. (2002) using optimization. Our results confirm and go beyond their results.

This manufacturing system consists of six processing stages A, B, C, D, E and F. Out of these, B and C are performed in parallel, and D and E can overlap by 6 h, otherwise the stages are performed in order. Some of the stages have smaller capacities than the others, which means that they must be performed more often, see Table 1. As in Björkqvist et al. (2002) we only consider batches which are multiples of 300 kg.

For the stages, there are six different units, out of which units 1 and 2 can be used for stages A and D, and units 3 and 4 can be used for stages B and C. The last two units can only be used for E and F respectively. The reactors used for stage A and D are also used for temporary storage, which means that the reactors cannot be used for another task before they are emptied. Half of the output from A goes through B, and the other half goes through C. After stages B and C, temporary storages S1 and

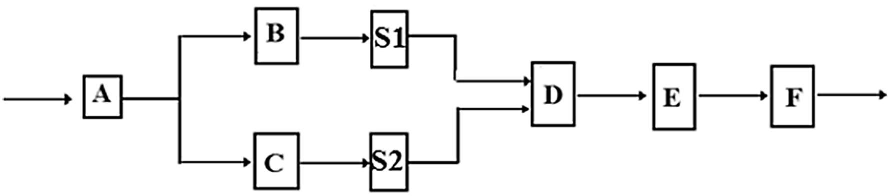

Fig. 1 A manufacturing system

Table 1 Specifications for the production stages in the manufacturing system

Stages	Processing time (h)	Cleaning time (h)	Durability (h)	Amount/batch (kg), repetitions/300 kg in parenthesis
A	7	4	48	300 (1)
B	4	2	60	50 (3)
C	7	2	60	75 (2)
D	10	4	24	150 (2)
E	8	4	24	150 (2)
F	1	0	48	150 (2)

S2 are used, gathering all intermediates needed for stage D. Stage E is performed overlapping with stage D, so that stage D and E are performed for $P_O = 6$ hours in parallel. The final stage F must be performed immediately after stage E has finished. The processing and cleaning times for stage N are denoted P_N and c_N respectively and given in Table 1.

A central limiting factor for production scheduling is the durability of the products produced in stages B and C, which can only be stored for 60 h. This means that the unit producing A should be switched to production of D at times, emptying the temporary storage containers S1 and S2.

4.1 A Max-plus model for the production system

Some simplifications were found necessary and/or appropriate. In particular, the following assumptions were made:

1. Unit 3 is used only for stage B, and unit 4 is used only for stage C.
2. The temporary storage containers after B and C respectively are only limited by the storage time Björkqvist et al. (2002), not by their size.
3. Scheduling of stage D is done based on the preliminary schedule obtained from the max-plus model. The model is constructed so that it gives all the alternatives for stage D, and the final schedule is obtained based on a set of simple rules.
4. Stages E and F are directly dependent on D, so the schedule for these is constructed based on the selected schedule for D, and these are left out from the max-plus model. Normally E and F simply follow D as they are faster than D, but when D is performed using both unit 1 and 2, E becomes the limiting factor. However, this can also be simply taken into account when selecting the schedule for D.
5. Durability of the output from each stage is first ignored, as it can be shown that it does not limit production of batches smaller than 4200 kg. The durability constraints will play a major role later, in the continuous production case studied in Sect. 5.

The goal is to perform production as fast as possible, subject to all the constraints that are present. This will now be formulated using a max-plus model, and for that we need 10 states x, listed and described below (where U_i = Unit i):

1. U_1 doing A
2. U_2 doing D, step 1
3. U_2 doing D, step 2
4. U_3 doing B, step 1
5. U_3 doing B, step 2
6. U_3 doing B, step 3
7. U_4 doing C, step 1
8. U_4 doing C, step 2
9. U_1 doing D, step 1
10. U_1 doing D, step 2

Now we write down the max-plus-algebraic state space model of this production, with all the constraints from the production and cleaning times included, in the equations that follow.

$$x_1(k) = \max(x_1(k-1) + P_A, x_5(k-1) + P_B, x_7(k-1) + P_C)$$

$$x_2(k) = \max(x_3(k-1) + P_D, x_5(k) + P_B, x_7(k) + P_C)$$

$$x_3(k) = \max(x_2(k) + P_D, x_6(k) + P_B, x_8(k) + P_C)$$

$$x_4(k) = \max(x_1(k) + P_A, x_6(k-1) + P_B)$$

$$x_5(k) = \max(x_4(k) + P_B)$$

$$x_6(k) = \max(x_5(k) + P_B)$$

$$x_7(k) = \max(x_8(k-1) + P_C, x_1(k) + P_A)$$

$$x_8(k) = \max(x_7(k) + P_C)$$

$$x_9(k) = \max(x_6(k) + c_A, x_8(k) + c_A)$$

$$x_{10}(k) = \max(x_9(k) + P_D)$$

In order to obtain an equation of type

$$x(k) = A \otimes x(k-1) \tag{3}$$

the right-hand-side expressions containing k or higher indices are substituted with expressions containing index $k-1$ at most. $x(k)$ is the earliest possible schedule given $x(k-1)$. After some straightforward (but tedious) algebraic manipulations and simplifications, we obtain the following equations:

$$x_1(k) = \max(x_1(k-1) + P_A, x_5(k-1) + P_B, x_7(k-1) + P_C)$$

$$\begin{aligned} x_2(k) = \max(&x_1(k-1) + 2P_A + k_1, x_3(k-1) + P_D, \\ &x_5(k-1) + P_A + P_B + k_1, x_6(k-1) + 3P_B, \\ &x_7(k-1) + P_A + P_C + k_1, x_8(k-1) + 2P_C) \end{aligned}$$

$$\begin{aligned} x_3(k) = \max(&x_1(k-1) + 2P_A + k_2, x_3(k-1) \\ &+ 2P_D, x_5(k-1) + P_A + P_B + k_2, x_6(k-1) \\ &+ 3P_B + k_3, x_7(k-1) + P_A + P_C + k_2, x_8(k-1) + 2P_C + k_4) \end{aligned}$$

$$x_4(k) = \max(x_1(k-1) + 2P_A, x_5(k-1) + P_A + P_B, x_6(k-1) + P_B, x_7(k-1) + P_A + P_C)$$

$$\begin{aligned} x_5(k) = \max(&x_1(k-1) + 2P_A + P_B, x_5(k-1) + P_A \\ &+ 2P_B, x_6(k-1) + 2P_B, x_7(k-1) + P_A + P_B + P_C) \end{aligned}$$

$$\begin{aligned} x_6(k) = \max(&x_1(k-1) + 2P_A + 2P_B, x_5(k-1) + P_A + 3P_B, \\ &x_6(k-1) + 3P_B, x_7(k-1) + P_A + 2P_B + P_C) \end{aligned}$$

$$x_7(k) = \max(x_1(k-1) + 2P_A, x_5(k-1) + P_A + P_B, x_7(k-1) + P_A + P_C, x_8(k-1) + P_C)$$

$$\begin{aligned} x_8(k) = \max(&x_1(k-1) + 2P_A + P_C, x_5(k-1) + P_A + P_B + P_C, \\ &x_7(k-1) + P_A + 2P_C, x_8(k-1) + 2P_C) \end{aligned}$$

$$\begin{aligned} x_9(k) = \max(&x_1(k-1) + 2P_A + c_A + k_1, x_5(k-1) + P_A + P_B + c_A + k_1, \\ &x_6(k-1) + 3P_B + c_A, x_7(k-1) + P_A + P_C + c_A + k_1, x_8(k-1) + 2P_C + c_A) \end{aligned}$$

$$\begin{aligned} x_{10}(k) = \max(&x_1(k-1) + 2P_A + P_D + c_A + k_1,\ x_5(k-1) + P_A + P_B + P_D + c_A + k_1, \\ &x_6(k-1) + 3P_B + P_D + c_A,\ x_7(k-1) + P_A + P_C + P_D + c_A + k_1, \\ &x_8(k-1) + 2P_C + P_D + c_A). \end{aligned}$$

For simplicity, the following constants have been introduced in the above equations: $k_1 = \max\left(2P_B, P_C\right), k_2 = \max\left(2P_B + P_D, 3P_B, P_C + P_D, 2P_C\right), k_3 = \max(P_B, P_D)$ and $k_4 = \max(P_C, P_D)$. After introduction of numerical values from Table 1, the A-matrix of the system becomes

$$A = \begin{pmatrix} 7 & \varepsilon & \varepsilon & \varepsilon & 4 & \varepsilon & 7 & \varepsilon & \varepsilon & \varepsilon \\ 22 & \varepsilon & 10 & \varepsilon & 12 & 12 & 22 & 14 & \varepsilon & \varepsilon \\ 32 & \varepsilon & 20 & \varepsilon & 22 & 22 & 32 & 24 & \varepsilon & \varepsilon \\ 14 & \varepsilon & \varepsilon & \varepsilon & 4 & 4 & 14 & \varepsilon & \varepsilon & \varepsilon \\ 18 & \varepsilon & \varepsilon & \varepsilon & 8 & 8 & 18 & \varepsilon & \varepsilon & \varepsilon \\ 22 & \varepsilon & \varepsilon & \varepsilon & 12 & 12 & 22 & \varepsilon & \varepsilon & \varepsilon \\ 14 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 14 & 7 & \varepsilon & \varepsilon \\ 21 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 21 & 14 & \varepsilon & \varepsilon \\ 26 & \varepsilon & \varepsilon & \varepsilon & 16 & 16 & 26 & 18 & \varepsilon & \varepsilon \\ 36 & \varepsilon & \varepsilon & \varepsilon & 26 & 26 & 36 & 28 & \varepsilon & \varepsilon \end{pmatrix} \tag{4}$$

The matrix A is not irreducible which means that the eigenvalue is not unique. The matrix restricted to D-stages has an eigenvalue of 20, while the matrix restricted to the other stages has an eigenvalue of 15. This is due to the fact that there are no limits on the storage between production stages B/C and D, so there is nothing synchronizing them. One can see this in Eq. (4), columns 2, 3, 9 and 10: The states related to stages A, B and C, that is 1 and 4–8, are not constrained by the other states related to stage D. Moreover, as the states related to D have a longer period, they will eventually not be constrained by the faster states.

4.2 Production schedule from iteration of the state equation

The max-plus model is still useful for production scheduling by iteration of the state Eq. (3). The initial state $x(0)$ is set to 0, and $x(k)$ is given by Eq. (3) using A from Eq. (4), and with no B or u. As mentioned in Sect. 4.1, the results need to be interpreted in order to obtain an optimal schedule:

1. The production is of batch type, so a certain number of repetitions of each stage are needed every time. For all stages but D and E, this means that only the necessary number of steps is used from the start of the schedule.
2. The schedule includes all possible D stages, most of them need to be discarded, based on the following:

   (a) Unit 1 should finish all A-stages before switching to doing stage D. Thus, all scheduled D-events based on states 9 and 10 prior to this should be discarded.
   (b) Out of the remaining D-stages, only the fastest up to the necessary number should be chosen.

3. The model does not include the limitation of the fact that the storage containers of B and C cannot be negative. This is possible in the beginning when the storage containers are empty when A is started. In that case, the first start of D cannot be earlier than $P_A + \max(2P_B, P_C)$ (= 15 h in this case), and the second start of D cannot be earlier than $P_A + \max(3P_B, 2P_C)$ (= 21 h). In practice, the latter is automatically taken care of by the limitations of stage E as in 4. below.

Table 2 Preliminary production schedule for stage D. Numbers in parenthesis and red are discarded as unit 1 is still needed for stage A, and numbers in blue and with a * are discarded as they are excessive

Batch size	300kg	600kg		900kg			1200kg			
Unit 2, step 1	15	15	35	15	35	55	15	35	55	75*
Unit 2, step 2	25*	25	45*	25	45	65*	25	45	65	85*
Unit 1, step 1	19	(19)	34	(19)	(34)	49	(19)	(34)	(49)	64
Unit 1, step 2	29*	(29)	44*	(29)	(44)	59*	(29)	(44)	(59)	74

Table 3 Final production schedule for stage D including the constraint for E, meaning in this case that there must be 8 h between each start of D. Unit 2 in black, unit 1 in red and marked with a *

Batch size	stage 1	stage 2	stage 3	stage 4	stage 5	stage 6	stage 7	stage 8
300kg	15	23*						
600kg	15	25	34*	42				
900kg	15	25	35	45	53*	61		
1200kg	15	25	35	45	55	64*	72	80*

4. The model does not contain limitations of stage E. E is only related to stage D, and it is necessary that the previous E-stage needs to finish before contents of the D-stage can be moved to E. The easy way to handle this constraint is to require that there is P_E between each start of D. This constraint becomes active after both unit 1 and 2 start doing stage D.
5. The schedule of E (unit 5) is always $P_D - P_O$ after the start of the corresponding stage D.

In Björkqvist et al. (2002), four different batch sizes were considered: 300, 600, 900 and 1200 kg. The schedule for the stages that are not related to stage D is obtained from the eigenvalue and the eigenvector of the matrix where the states related to D (2, 3, 9 and 10) are left out.

This results in the eigenvalue of 15 and eigenvector of $[0\ \ 7\ \ 11\ \ 15\ \ 7\ \ 14]^T$. This means that each unit should be started according to the eigenvector plus a multiple of 15. The schedule for the D-stages is obtained from the preliminary schedule using point 2a and 2b. The capacity of stage D is 150 kg, so 2, 4, 6 and 8 stages of D respectively are needed in the four considered cases. The relevant states are 2, 3, 9 and 10, and the schedule for these are shown in Table 2.

On top of that the constraint related to stage E, that is the time between the start of a unit producing D should not be less than 8, is enforced. This affects all the events starting from the first-time producing D using unit 1, which can be seen in Table 3, where the final schedule for D is given.

Production times and rates for the different batches are given in Table 4.

Table 4 Production rate for different batch sizes

Batch size	300	600	900	1200	1500	1800	2100	2400	2700	3000	3300	3600	3900	4200
Prod. time	36	55	74	93	112	131	150	169	188	207	226	245	264	283
Prod. rate	8.33	10.91	12.16	12.90	13.39	13.74	14.00	14.20	14.36	14.49	14.60	14.69	14.77	14.84

In Björkqvist (2002) the first four scheduling problems were studied using mixed integer linear programming (MILP). They proved that the schedules for batch sizes 300 and 600 are optimal, and we got the same schedules in our study. The reported CPU times were 4.4 and 26.6 s, respectively. For batch sizes 900 and 1200 they reported CPU times 59.9 and 192 s, respectively. For 900 kg the optimization had stopped in a suboptimal schedule of 76 h. Our calculations were in practice instantaneous.

As can be seen from Table 4, the production time is increased with 19 h for each addition of 300 kg of production. In the absence of the durability constraints, it is shown below in the next section that the production rate will converge towards $\frac{300}{19}\frac{\text{kg}}{\text{h}} \approx 15.79\text{kg/h}$.

5 Continuous production case

With the continuous production case we mean scheduling of very large (unlimited) batches so that the durability constraints need to be considered. Production consists of two modes related to the storage of B and C: the filling mode and the emptying mode. The production cases are introduced below with a graph of each case. A typical process involves the A, B, C, and D stages, as already mentioned in Sect. 4.1.

In the filling mode, the initial production schedule, presented in Sect. 4.2, accumulates B and C in storage. Accumulation is due to the fact that A is produced at a rate of 300 kg/15 h, and D at a minimum production rate of 300kg/20 hours. A simple but suboptimal strategy is to delay the production of A, so that it also produces at a rate $300\text{kg}/20\text{h} = 15\text{kg/h}$ and, thus, avoid the accumulation of the B and C storages. However, it is beneficial to clean the unit used for production of A and produce D using two units, resulting in a production rate of $300\text{kg}/16\text{h}$. Thus, the scheduling of production can be seen as switching between the mode where only one unit is used for production of D, and the mode where two units are used for production of D. In the first mode, the total production is limited by the production of D, and the production rate is $r_1 = 150/10$. In the second mode, the total production will be limited by the production of E, when the production rate is $r_2 = 150/8$. The actual production rate r_p will be with production times that are a weighted average of r_1 and r_2, that is

$$r_p = \left(t_{d1} \cdot r_1 + t_{d2} \cdot r_2\right) \cdot \frac{1}{p},$$

where $p = t_{d1} + t_{d2} + 2t_c$ is the period of production, and t_{d1} is the time for production using only one unit for D, t_{d2} is the time for production using two units for D, and t_c is the cleaning time. Now $t_{d2} = \frac{t_{d1} \cdot r_a}{r_2}$ asymptotically, where r_a is the accumulation rate.

The production rate will be

$$r_p = \frac{t_{d1} \cdot r_1 + \left(\frac{t_{d1} \cdot r_a}{r_2}\right) \cdot r_2}{p} = \frac{t_{d1} \cdot r_1 + t_{d1} \cdot r_a}{p},$$

which can be differentiated with respect to t_{d1}, and

$$\frac{\partial r_p}{\partial t_{d1}} = \frac{\left(r_1 + r_a\right) \cdot t_{d2} + \left(r_1 + r_a\right) \cdot 2t_c}{p^2} > 0.$$

The last inequality follows as all the times and rates are positive, which means that t_{d1} should be chosen as large as possible, that is at the durability constraint. Both B and C have a durability of 60 h, and this is the durability constraint that should be the target for the schedule.

Production schedules can be obtained using max-plus for both the filling and the emptying phase. Switching between these phases cannot be done using max-plus, but it is easily done by keeping track of the B and C storages. Switching can be carried out after a certain number of repetitions of A in the filling mode and after a certain number of repetitions of D in the emptying mode. The schedule after the switch can be initialized based on the previous schedule.

Long-term production consists of two different modes:

1. One of the units 1 or 2 is used for production of A, and the other is used for production of D. In this mode B and C are accumulated in storage. The length of this filling mode is characterized by the number of repetitions of A, denoted n_A.
2. Both units 1 and 2 are used for production of D, when the storage containers for B and C are emptied. The length of this emptying mode is characterized by the number of repetitions of D, denoted n_D.

Scheduling of long-term production essentially consists of the choice of these repetitions n_A and n_D. The choice of n_A affects the storage times, the more repetitions the more accumulation in the storage areas, and the longer time is needed for emptying them. The choice of n_D is bound to the choice of n_A, as the storage containers need to be sufficiently emptied before restart of production of A. It will be shown that n_A should be about twice n_D. The storage containers will build up during production of A at a rate

$$\frac{300 \text{ kg}}{15 \text{ h}} - \frac{300 \text{ kg}}{20 \text{ h}} = 5 \text{ kg/h}$$

The storage containers will be emptied at a rate $150/8$ kg/h during usage of the two D-units. An A-repetition takes 15 h, and a D-repetition takes 8 h, and in the long run buildup and emptying must be the same, that is $n_A \cdot 15\,\mathrm{h} \cdot \frac{5\mathrm{kg}}{\mathrm{h}} - n_D \cdot 8\,\mathrm{h} \cdot \frac{150}{8}\mathrm{kg/h}$ is bounded by 150 kg, and thus

$$\lim_{n_A \to \infty} \frac{n_D}{n_A} = \frac{1}{2}$$

If the storage time is unlimited, one can let n_A be unlimited, and consider it continuous production. The total production is given by $n_A \cdot 300\,\mathrm{kg}$. The time it takes to complete the production is $T_{\mathrm{startup}} + n_A \cdot 15\,\mathrm{h} + T_{\mathrm{switch}} + n_D \cdot 8\,\mathrm{h} + P_F$, where T_{startup} is the time it takes to start up the system with empty storages, T_{switch} is the time it takes to clean the unit producing A and start up the production of D, and P_F is the time it takes to do the final step of production of F. This gives

$$\lim_{n_A \to \infty} \frac{n_A \cdot 300\mathrm{kg}}{n_A \cdot 15\mathrm{h} + n_D \cdot 8\mathrm{h} + T_{\mathrm{startup}} + T_{\mathrm{switch}} + P_F} = \frac{300}{19} \frac{\mathrm{kg}}{\mathrm{h}}, \tag{5}$$

which is the production rate for continuous production.

One can choose to repeat the D-production one time more than necessary, as this reduces the storage times, with a cost of having the D-production unit idle for a while at the beginning of the next filling mode. It might still increase the production rate in the case of limited storage time. This is tested in Table 8, and it was found to reduce the overall production rate in all cases.

The $n_A - n_D$ cycle must be eventually periodic. It is uniquely determined by the storage in B and C at the beginning of the filling mode, and there are only a finite number of possibilities for those. For example, if the storage containers are empty when the filling mode starts, the process is periodic from the beginning (as when n_A = 15 and n_D= 8, cf. Fig. 6).

5.1 Continuous production case using max-plus

The continuous production case consists of two modes, one filling mode and one emptying mode. Three different max-plus models are formulated, describing different parts of the system:

M1. The production of A, B, and C, during the filling mode
M2. The production of D and E using only one of the units for D, during the filling mode
M3. The production of D and E using both units for D, during the emptying mode

Between the filling and the emptying mode there is a switch, where production of A, B, and C is ended, and the unit producing A is cleaned and used for production of D. This means that the production of D is switched from M2 to M3, and that the state of M3 is initialized using M1 and/or M2 (depending on if the cleaned A-producing unit from M1 or the D-producing unit from M2 is the first available to start M3). Between the emptying and filling mode there is also a switch, where one of the units used for production of D is cleaned, and production of A, that is M1, is started. Furthermore, M3 is switched to M2. The states of M1 and M2 are both initialized using states of M3.

5.2 Production of A, B, and C during the filling mode, model M1

This system consists of three processing stages A, B, and C. Out of these B and C are performed in parallel. For the stages, there are three different units, out of which unit 1 can be used for stage A and unit 3 can be used for stages B with three steps. Unit 4 can only be used for C with two steps. This will now be formulated using a max-plus model, and for that, we need 6 states x_i, listed and described below:

1. U_1 doing A
2. U_3 doing B, step 1
3. U_3 doing B, step 2
4. U_3 doing B, step 3
5. U_4 doing C, step 1
6. U_4 doing C, step 2

Now we write down the max-plus-algebraic state space model of this DES.

$$x_1(k) = \max(x_1(k-1) + P_A, x_3(k-1) + P_B, x_5(k-1) + P_C)$$

$$x_2(k) = \max(x_1(k-1) + 2P_A, x_3(k-1) + P_A + P_B, x_4(k-1) + P_B, x_5(k-1) + P_A + P_C)$$

$$\begin{aligned} x_3(k) = &\max(x_1(k-1) + 2P_A + P_B, x_3(k-1) + P_A + 2P_B, \\ &x_4(k-1) + 2P_B, x_5(k-1) + P_A + P_B + P_C) \end{aligned}$$

$$\begin{aligned} x_4(k) = &\max(x_1(k-1) + 2P_A + 2P_B, x_3(k-1) + P_A + 3P_B, \\ &x_4(k-1) + 3P_B, x_5(k-1) + P_A + 2P_B + P_C) \end{aligned}$$

$$x_5(k) = \max(x_1(k-1) + 2P_A, x_3(k-1) + P_A + P_B, x_5(k-1) + P_A + P_C, x_6(k-1) + P_C)$$

$$\begin{aligned} x_6(k) = &\max(x_1(k-1) + 2P_A + P_C, x_3(k-1) + P_A + P_B + P_C, \\ &x_5(k-1) + P_A + 2P_C, x_6(k-1) + 2P_C) \end{aligned}$$

After introduction of numerical values from Table 1, the A-matrix of the system becomes as follows:

$$A = \begin{pmatrix} 7 & \varepsilon & 4 & \varepsilon & 7 & \varepsilon \\ 14 & \varepsilon & 11 & 4 & 14 & \varepsilon \\ 18 & \varepsilon & 15 & 8 & 18 & \varepsilon \\ 22 & \varepsilon & 19 & 12 & 22 & \varepsilon \\ 14 & \varepsilon & 11 & \varepsilon & 14 & 7 \\ 21 & \varepsilon & 18 & \varepsilon & 21 & 14 \end{pmatrix} \tag{6}$$

The eigenvalue of the A-matrix is 15 and the eigenvector is $\begin{bmatrix} 0 & 7 & 11 & 15 & 7 & 14 \end{bmatrix}^T$.

5.3 Production of D and E during the filling mode using only one of the units for D, model M2

This system consists of two processing stages D and E. For the stages, there are two different units, out of which unit 1 can be used for stage D, and unit 5 can be used for stage E. In this case, unit 2 is used for production of A. This will now be formulated using a max-plus model, and for that, we need 2 states x_i, listed and described below.

1. U_1 doing D
2. U_5 doing E

Now we write down the max-plus-algebraic state space model of this DES.

$$x_1(k) = \max(x_1(k-1) + P_D, x_2(k-1) + P_E - P_D + P_o)$$

$$x_2(k) = \max\left(x_1(k-1) + 2P_D - P_o\right), x_2(k-1) + P_E)$$

After introduction of numerical values from Table 1, the A-matrix of the system becomes

$$A = \begin{pmatrix} 10 & 4 \\ 14 & 8 \end{pmatrix}. \tag{7}$$

The eigenvalue of the A-matrix is 10 and the eigenvector is $\begin{bmatrix} 0 & 4 \end{bmatrix}^T$

5.4 Production of D and E during the emptying mode using both units for D, model M3

This system consists of two processing stages D, E, and F (see Fig. 2). Stage F is not modeled, as it follows directly after E, and is so fast that it does not constrain anything. For the other stages, there are three different units, out of which units 1 and 2 can be used for stage D, and unit 5 can be used for stages E with two steps.

This will now be formulated using a max-plus model, and for that, we need 4 states x_i, listed and described below:

1. U_1 doing D
2. U_2 doing D
3. U_5 doing E, step 1
4. U_5 doing E, step 2

The production of E is modeled using two states, modeling which D production unit is currently served.

Now we write down the max-plus-algebraic state space model of this DES.

$$x_1(k) = \max(x_1(k-1) + P_D, x_4(k-1) + P_E - P_D + P_o)$$

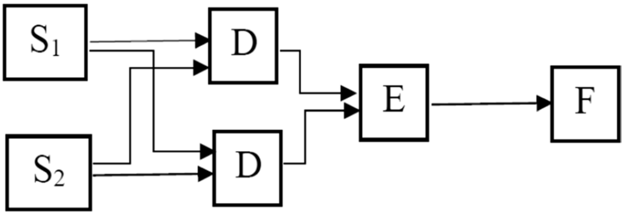

Fig. 2 System in emptying mode

$$x_2(k) = \max(x_1(k) + P_D - P_o)$$

$$x_3(k) = \max(x_3(k-1) + P_D, x_2(k) + P_E - P_D + P_o)$$

$$x_4(k) = \max(x_3(k) + P_D - P_o)$$

In order to obtain an equation of type $x(k) = A \otimes x(k-1)$, the right-hand-side expressions containing k or higher indices are substituted with expressions containing index $k-1$ at most. After some straightforward algebraic manipulations and simplifications, we obtain the following equations:

$$x_1(k) = \max(x_1(k-1) + P_D, x_4(k-1) + P_E - P_D + P_o)$$

$$x_2(k) = \max(x_1(k-1) + 2P_D - P_o, x_4(k-1) + P_E)$$

$$x_3(k) = \max(x_1(k-1) + P_D + P_E, x_3(k-1) + P_D, x_4(k-1) + 2P_E - P_D + P_o)$$

$$x_4(k) = \max(x_1(k-1) + 2P_D + P_E - P_o, x_3(k-1) + 2P_D - P_o, x_4(k-1) + 2P_E)$$

After introduction of numerical values from Table 1, the A-matrix of the system becomes

$$A = \begin{pmatrix} 10 & \varepsilon & \varepsilon & 4 \\ 14 & \varepsilon & \varepsilon & 8 \\ 18 & \varepsilon & 10 & 12 \\ 22 & \varepsilon & 14 & 16 \end{pmatrix}. \quad (8)$$

The eigenvalue of the A-matrix is 16 where the eigenvector is $\begin{bmatrix} 0 & 4 & 8 & 12 \end{bmatrix}^T$.

The best schedule, under the durability constraint 60 h, using $n_A = 14$ and $n_D = 7$ is presented in Tables 5, 6, 7. The period length is 268 h, and the production rate is 15.67.

Table 5 shows that the production, also seen in Figure 4, is periodic from the start. The period length is 268 h, which is the difference between the numbers in the orange cells, which indicates the restarting times for production of A. Table 6 gives the schedule for the production of D and E, and the same period 268 h can be seen between the switching time instances highlighted with green cells. Table 7

Table 5 Schedule for units producing A, B and C according to model M1. In the last round, the start of B3 (highlighted with yellow) liberates A for cleaning (which takes 4 h) and gives the starting time of D1 in Table 7. The cells highlighted with orange are the ones initialized by the orange cells in Table 7

Step index	1	2	3	…	14	15	16	…	28	29	30	…	42	43	44	…
A	0	15	30	…	195	268	283	…	463	536	551	…	731	804	819	…
B1	7	22	37	…	202	275	290	…	470	543	558	…	738	811	826	…
B2	11	26	41	…	206	279	294	…	474	547	562	…	742	815	830	…
B3	15	30	45	…	210	283	298	…	478	551	566	…	746	819	834	…
C1	7	22	37	…	202	275	290	…	470	543	558	…	738	811	826	…
C2	14	29	44	…	209	282	297	…	477	550	565	…	745	818	833	…

Table 6 Schedule for units producing D and E according to model M2. The cells highlighted with green are the ones initialized by the green cells in Table 7

Step index	1	2	3	4	…	20	21	22	…	41	42	43	…
D	15	25	35	45	…	205	272	282	…	472	540	550	…
E	19	29	39	49	…	209	276	286	…	476	544	554	…

Table 7 Schedule for units producing D and E according to model M3. The cells highlighted with yellow are the ones initialized by the yellow cells in Table 5. The start of the last round of D2, highlighted with orange, is after completion (10 h) and cleaning (4 h) used for the starting time of A in Table 5. The start of the last round of D1, highlighted with green, is after completion (10 h) used for the starting time of D in Table 6

Step index	1	2	3	4	5	6	7	8	9	10	11	12	13	…
D1	214	230	246	262	482	498	514	530	750	766	782	798	1018	…
E1	218	234	250	266	486	502	518	534	754	770	786	802	1022	…
D2	222	238	254		490	506	522		758	774	790		1026	…
E2	226	242	258		494	510	526		762	778	794		1030	…

gives the schedule for the production of D and E, using two units for D. The schedule becomes periodic starting from iteration 21, and the period length of 268 h can again be seen between the switching time instances highlighted with yellow and green respectively.

As can be seen from Table 6, the production schedule becomes periodic after the initial phase, where the empty storage containers introduce constraints that distort the schedule. For example, in the first round A starts at time 0 and D starts at time

15, while in the second round A starts at 268 and D at 272. This faster start of D is possible because one can use raw material from the storage containers of B and C. It is clear that the schedule will continue as periodic, as the storage containers are also changing using the same period, and the system starts at the same initial condition at each start of a new period. Tables 5 and 7 are periodic from the start.

5.5 Optimal continuous production schedules

Using the models M1-M3 we can calculate the production schedules using different choices of.

n_A and n_D presented in Table 8. Based on the schedules we can calculate the periods and production rates, also presented in Table 8.

As can be seen in Table 8, the number of D-steps is the same in odd numbered A-steps as in the previous even numbered case. In addition, production is larger during the D-steps (when both units are used for production of D, with a rate 150/8 = 18.75 kg/h) than during A-steps (when only one unit is used for D, with a rate of 15 kg/h). In our case, the limit for the storage times is 60 h, and the most relevant schedules are illustrated in Figs. 3, 4, 5.

Figure 6 results in a worse production rate than in Fig. 4, and it is not possible to improve on this by increasing the A-steps. As can be seen from Table 8, it will result in a violation of the durability constraint.

Increasing the storage time the production rate converges to 300/19 as shown in Eq. (5), which can be seen in Table 9 and Fig. 7.

6 Conclusion

This paper described how a max-plus model for a manufacturing system can be constructed, and an optimal schedule can be found without optimization. The scheduling of production systems consisting of many stages and different units was considered, where some of the units were used for multiple production stages.

Table 8 Periods, storage times and production rates for long-term strategies for different number of A- and D-repetitions

n_A	n_D	Period (h)	Maximum storage times for B and C (h)	Production rate (kg/h)	Figure
12	6	230	50	15.65	
12	7	238	48	15.13	
13	6	249	54	15.66	
13	7	257	52	15.18	
14	7	268	58	15.67	3 and 4
14	8	276	56	15.22	
15	7	287	62	15.68	5
15	8	295	60	15.25	6
16	8	306	66	15.69	
16	9	314	64	15.29	

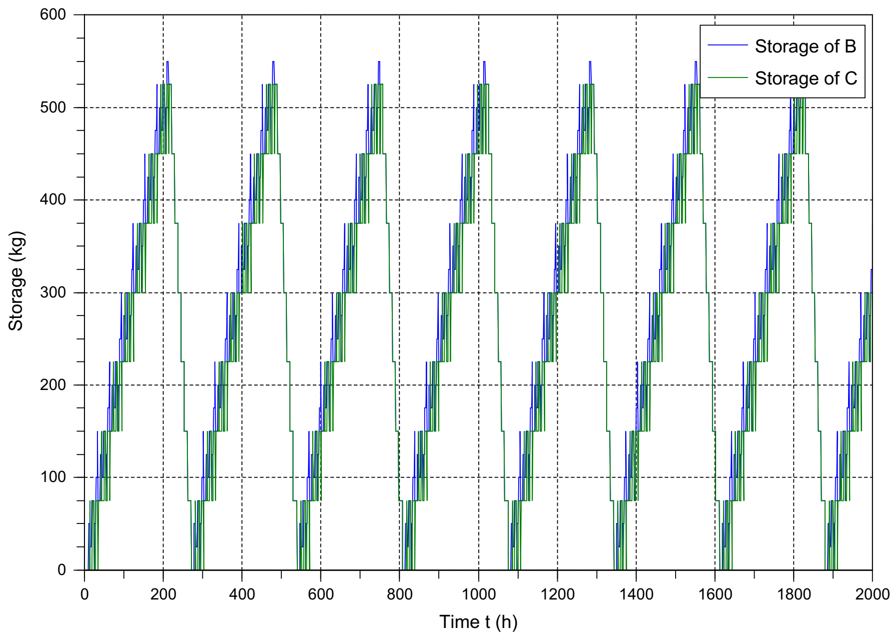

Fig. 3 Optimal schedule with maximum storage time 58 h from the start. The period is 268 h, and the periodicity starts at $t = 3h$. Production rate 15.67 kg/h

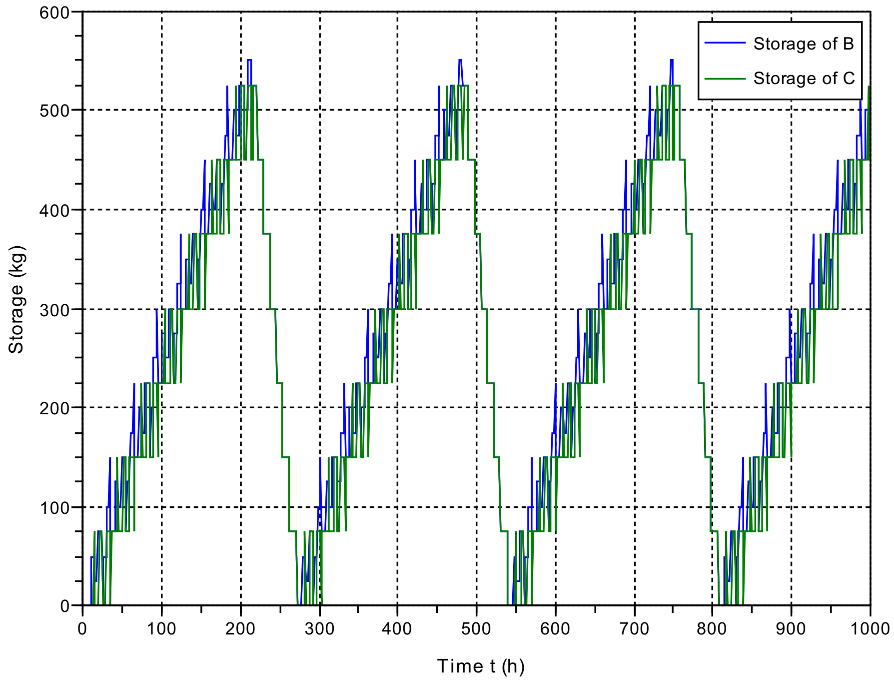

Fig. 4 The first 1000 h of the schedule in Fig. 2

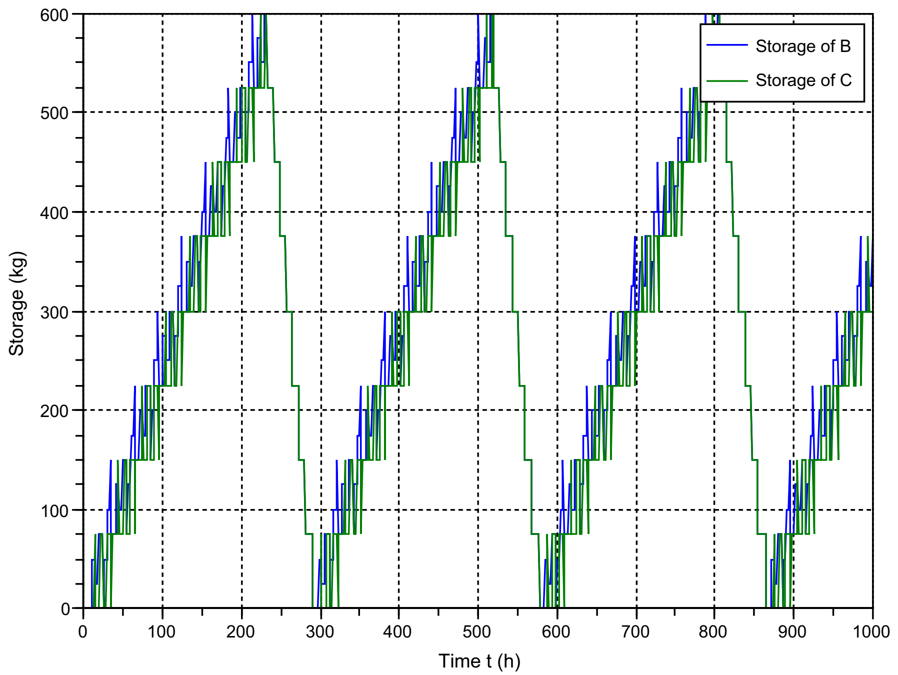

Fig. 5 Optimal schedule with maximum storage time 62 h from the start. The period is 287 h, and the periodicity starts at $t = 3$ h. Production rate 15.68 kg/h

If a production unit is used for different stages, cleaning is needed in between, while no cleaning is needed between stages of the same type. The obtained state update equation was in this case also rewritten in the form $x(k) = A \otimes x(k-1)$ using several cross-substitutions, and extension of the state space with delayed states. Structural decisions, such as using a unit for different tasks, were found to be difficult to formulate in max-plus algebra. Three possible operation modes with the structure fixed was identified and modeled separately using max plus. The resulting model is a switching max-plus linear system. The central driving factor for structural switches was durability constraints, which were present in the production. Thus, only a part of the schedule was obtained by solving eigenvalue problems of the max-plus model, the structural decisions were made on the basis of a few alternative schedules obtained using max-plus. This was based on the finding that structural switches should be postponed as late as possible, so the criterion used was to do the switch one step before the step that was the first that violated at least one durability constraint. By using such switches between two max-plus models, an optimal continuous production schedule was obtained.

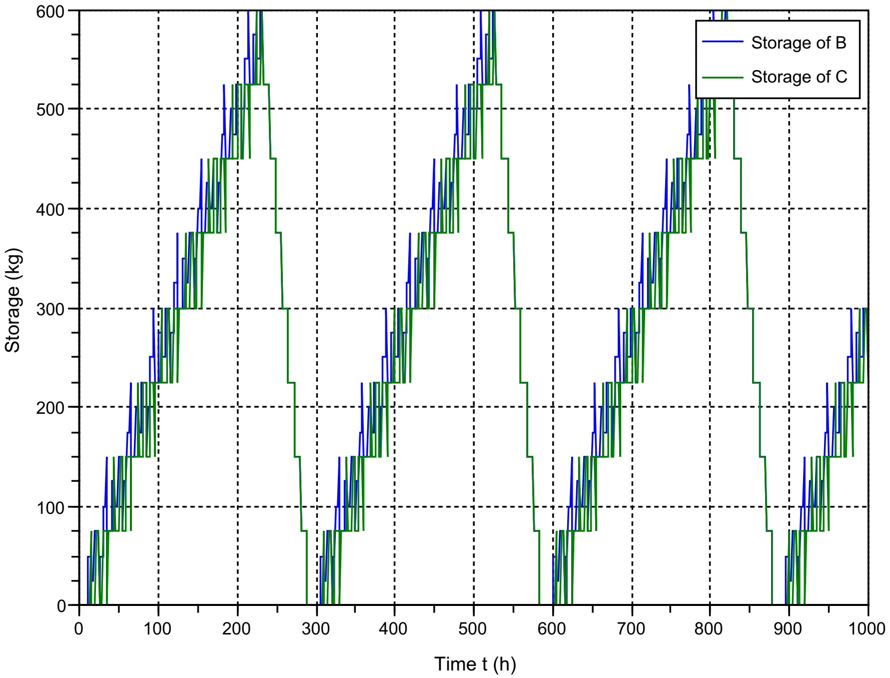

Fig. 6 Schedule with delayed switch from D to A, resulting in eleven hours waiting for the other D producing unit at each cycle. The period is the 295 h, and the periodicity starts in the beginning. Production rate 15.25 kg/h

Table 9 Numerical test of upper limit on production rate

Storage time	n_A	n_D	Production rate
58	14	7	15.67
114	28	15	15.73
230	57	29	15.76
462	115	58	15.775
⋮			⋮
∞			300/19 ≈ 15.789474

It was also shown that the such obtained schedule converges towards the theoretical maximum of the production rate.

As future work, one could consider different generalizations, what are the limitations, what kind of systems can actually be handled using the studied methodologies.

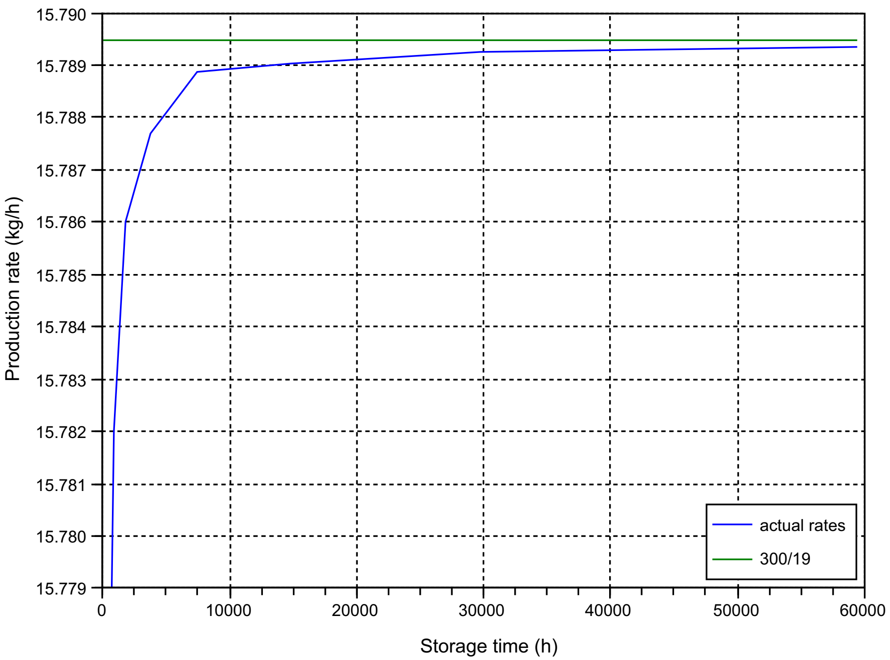

Fig. 7 Actual production rates as a function of allowed storage times, and the presumed limit 300/19 for the production rate

Funding Open access funding provided by Abo Akademi University (ABO).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Al-Bermanei, H. (2021), Applications of max-plus algebra to scheduling, Ph.D. thesis, Åbo Akademi University Press, https://urn.fi/URN:ISBN:978-951-765-983-3 .

Baccelli F, Cohen G, Olsder GL, Quadrat JP (1992) Synchronization and Linearity. Wiley, New York

Björkqvist J., Roslöf J., Rönnback R., and Westerlund T. (2002). Evaluation of an iterative mathematical optimization approach for industrial size batch scheduling. AIDIC Conference Series, Vol. 5, 33–40, Elsevier Science.

van den Boom TJ, Muijsenberg MVD, De Schutter B (2020) Model predictive scheduling of semi-cyclic discrete-event systems using switching max-plus linear models and dynamic graphs. Discrete Event Dynamic Systems 30(4):635–669

van den Boom TJ, De Schutter B (2006) Modelling and control of discrete event systems using switching max-plus-linear systems. Control Eng Pract 14(10):1199–1211

Braker JG, Olsder GJ (1993) The power algorithm in the max algebra. Linear Algebra Appl 182:67–89

Cuninghame-Green RA, Butkovic P (2004) Bases in max-algebra. Linear Algebra Appl 389:107–120

De Schutter, B. (1996), Max-Algebraic System Theory for Discrete Event Systems, Ph.D. dissertation, Faculty of Applied Sciences, K.U. Leuven, Leuven, Belgium.

De Schutter, B, van den Boom, T. (2008). Max-plus algebra and max-plus linear discrete event systems: An introduction. In Proceedings of WODES 2008, 9th International Workshop on Discrete Event Systems, IEEE.

Giffler B, Thompson GL (1960) Algorithms for solving production-scheduling problems. Oper Res 8:487–503

Heidergott B, Olsder G, van der Woude J (2006) Max Plus at Work. Princeton University Press, Princeton, New Jersey

Mutsaers M., Özkan, L., Backx, T. (2012). Scheduling of energy flows for parallel batch processes using max-plus systems. In 8th IFAC Symposium on Advanced Control of Chemical Processes, IFAC Proceedings Volumes, Volume 45, Issue 15, 614–619.

De Schutter B (2000) On the ultimate behavior of the sequence of consecutive powers of a matrix in the max-plus algebra. Linear Algebra Appl 307:103–117

Hazem Al Bermanei is a senior lecturer of Mathematics and applied Mathematics at IT program, Faculty of Business ICT and Life Sciences, Turku University of Applied Sciences, Finland. He has received the M.Sc. degree in Mathematics from South-Bank University - London in1997 and the Licentiate degree and the Ph.D. degree in Applied Mathematics from Åbo Akademi University, Turku, Finland in 2012 and 2021, respectively which was about Applications of Max-Plus Algebra to Scheduling. His research areas are application of max-plus algebra, practice-oriented pedagogy, and pedagogical approaches to learning.

Jari M. Böling received the M.Sc. degree in chemical engineering and the Ph.D. degree in control engineering from Åbo Akademi University, Turku, Finland, in 1994 and 2001, respectively. From 2003 to 2004, he was a Postdoctoral researcher with the University of California Santa Barbara, USA. Since 2005, he has been a Senior Lecturer of control engineering with the Faculty of Science and Engineering, Åbo Akademi University. His research interests include system identification, adaptive control, machine learning, modeling and simulation of dynamical systems.

Göran Högnäs is Professor emeritus of Applied Mathematics at Åbo Akademi University (ÅAU) in Finland. He received his PhD in Applied Mathematics at ÅAU in 1974 and did postdoctoral work at Université de Strasbourg (France), Aarhus University (Denmark), and the University of South Florida in Tampa. He served as Associate Professor at ÅAU from 1980 to 1991 and as Full Professor from 1991 until his retirement in 2013. His research areas are probability theory, stochastic dynamical systems, and mathematical models for population processes. In 1997 - 2009 he was chairman of the board of the national Finnish Graduate School in Stochastics and Statistics. Together with Arunava Mukherjea of the University of South Florida he authored the monograph Probability Measures on Semigroups. Convolution Products, Random Walks and Random Matrices (2nd edition 2011 by Springer Science + Business Media). He serves as Associate Editor for the Journal of Theoretical Probability since 1988 and the Journal of Difference Equations and Applications since 2008.