

PERFORMANCE EVALUATION AND LUMPED PARAMETER MODELLING OF SINGLE SERVER FLOWLINES SUBJECT TO BLOCKING: AN EFFECTIVE PROCESS TIME APPROACH

A.A.A. Kock¹, F.J.J. Wullems³, L.F.P. Etman¹, I.J.B.F. Adan², J.E. Rooda¹

¹Department of Mechanical Engineering, ²Department of Mathematics and Computer Science
Eindhoven University of Technology, WHOog -1.15, PO Box 513, 5600 MB, Eindhoven, The Netherlands

³Reliability and Maintainability Department, Steelweld B.V. (currently at Ormit B.V.)

{i.j.b.f.adan,l.f.p.etman,a.a.a.kock,j.e.rooda}@tue.nl

phone: +31-40-247-3578

Abstract: The present paper extends the so-called Effective Process Time (EPT) approach to single server flowlines with finite buffers and blocking. The power of the EPT approach is that it quantifies variability in workstation process times without the need to identify each of the contributing disturbances, and that it directly provides an algorithm for the actual computation of EPTs. It is shown that EPT realizations can be simply obtained from arrival and departure times of lots, by using sample path equations. The measured EPTs can be used for bottleneck analysis and for lumped parameter modeling. Simulation experiments show that for lumped parameter modeling of flowlines with finite buffers, in addition to the mean and variance, offset is also a relevant parameter of the process time distribution. A case from the automotive industry illustrates the approach.

keywords: Blocking, Discrete event simulation, Effective process time, Finite buffer capacity, Flow-line optimisation, Lumped parameter modelling, Performance analysis

1 Introduction

Single server workstations with finite buffer sizes in a tandem flowline represent an important class of manufacturing systems. Examples of such flowlines are semi-synchronous lines and assembly lines, as e.g. encountered in the automotive industry.

The performance of a flowline is commonly expressed in terms of throughput and flow time. Both performance indicators are influenced by blocking. The finite capacity of the buffers in the single server flowlines considered in this paper introduces blocking in the line.

Blocking causes suspension of service to a lot (which implies loss of production capacity) since a finished lot cannot be send on due to a saturated downstream buffer. Starvation refers to the situation where processing of the next lot is suspended due to an empty upstream buffer.

Variability in process times is the main reason that blocking and starvation occur. The variability of process times can be traced to several common sources. First of all, natural process times are variable due to differences in product types, machine states at product entry, operator behaviour, etcetera. Furthermore, disturbances such as setups, preventive maintenance, machine failures and absence of operators occur. These disturbances cause loss of production capacity effectively available at the workstation, which in turn reduces the throughput and increases the variability of process times. Subsequent workstations affect one another more prominently as the variability of process times increases. Variability of process times on workstation j can cause starvation on workstation $j + 1$. Furthermore, in a flowline with finite buffers, variability of process times of workstation j can cause blocking on workstation $j - 1$.

Obviously, for performance analysis of a finitely buffered flowline, an analysis tool that quantifies both the production losses and the level of variability of process times is required. A commonly applied performance analysis metric is the overall equipment efficiency, OEE. However, OEE can only be used for quantifying production losses. Therefore an alternative analysis tool will be used in this paper.

Hopp and Spearman (2001) introduced this alternative concept to account for irregularities in process times of workstations. The new concept, effective process time (EPT), is defined as the total time seen by a lot at a workstation from a logistical point of view. Here, total time indicates the total time that the lot has effectively consumed production capacity of the workstation. EPT is based on the notion that,

from a logistical perspective, a workstation does not care whether production capacity is claimed since the server is processing the lot or whether production capacity is claimed by other influences. These other influences are included in the EPT of the workstation.

Hopp and Spearman’s notion of including disturbances on processing in the effective process times is not new, see e.g. the work of Chen and Chen (1990), Dallery and Gershwin (1992), Buzacott and Shanthikumar (1993). The aforementioned authors all assume, or measure, distributions for the disturbances on processing and combine these into one single distribution. However, from industrial practice, it is often hard to identify and quantify all individual disturbances. For this purpose, Jacobs *et al.* (2001) adapted the EPT concept for infinitely buffered, isolated single server workstations. They presented a method based on sample path analysis to translate lot arrivals and departures into an EPT distribution.

The obtained EPT distributions can be used for performance optimization. Based on the characteristic parameters of the EPT distributions, i.e. the mean effective process time t_e and the squared coefficient of variation c_e^2 , a bottleneck analysis can be performed, after which an approximating model can be used to predict the changes in system performance. Hopp *et al.* (2002) also use t_e and c_e^2 as workstation parameters in an open queueing network model for flowline optimisation. However, they compute t_e and c_e^2 from the theoretical process time values by assuming that outages are adequately represented by exponentially distributed failures and repairs. The EPT framework presented in this paper follows the concept of Jacobs *et al.* (2003), which does not require the characterisation of the various contributing disturbances. This is a clear advantage of the EPT–approach since, as mentioned above, it is in practice often hard or impossible to quantify all individual sources of disturbances.

This paper aims to generalize the EPT–approach for application to single server flowlines subject to blocking. That is, the paper considers finite buffers rather than infinite ones. Workstations can no longer be analysed in isolation due to the dependencies introduced by blocking. Therefore, an EPT–algorithm for the blocking case is presented. Furthermore, the effect of the distribution shape on the accuracy of the EPT lumped parameter (ELP) model is investigated. A case from automotive industry is used to illustrate the EPT–approach. Note that throughout the paper, mainly the effects of blocking are discussed since starvation also occurs in infinitely buffered workstations.

The paper is organized as follows. In Section 2, an outline of the EPT–approach is presented. Subsequently, computation of EPT–realisations for single server workstations with finite buffers is considered in Section 3. EPT–based lumped parameter modelling in the context of finitely buffered flowlines is discussed in Section 4. The concepts discussed throughout this paper are illustrated using a case from automotive industry in Section 5. Finally, Section 6 concludes the paper.

2 A framework for implementing EPT

The EPT–approach, based on the concept of Jacobs *et al.* (2003), consists of four stages, as visualized in figure 1.

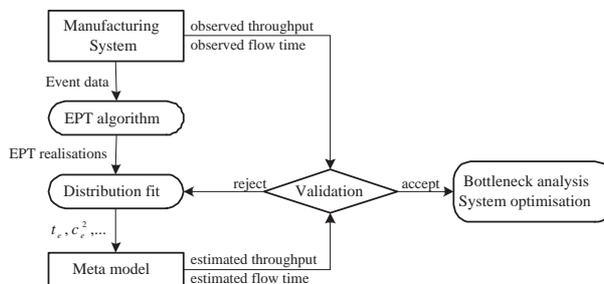


Figure 1: Schematic representation of the EPT–approach

First, EPT–realisations are obtained from the discrete manufacturing system. An EPT realisation is defined by Jacobs *et al.* as: ‘the time a lot was in process plus the time a lot (not necessarily the same lot) could have been in process’. EPT–realisations can be computed from event data, such as

arrivals and departures of lots on workstations. The EPT-realizations are computed by means of an EPT-algorithm. Concepts similar to the EPT are the completion time, the cycle time, the operation time, the processing time and the service time (Chen and Chen 1990, Dallery and Gershwin 1992, Buzacott and Shanthikumar 1993, Rossetti and Clark 2003). Completion times, or similar concepts, are used in sample path-like analyses of queueing systems. However, in sample path analysis, it is common to determine lot departures from corresponding lot arrivals and the completion time. The EPT-concept presented in this paper uses the sample path equations reversibly, that is, effective process times are determined from arrival and departure data. The sample path equations are thus a means to obtain EPT-realizations from an operating production system. The operation time as defined by Rossetti and Clark (2003) is very similar to EPT; however, Rossetti and Clark do not use it to quantify the level of variability.

Next, the EPT-realizations are fitted to distributions. Here, distributions are fitted based on relevant workstation properties, such as the mean EPT t_e and the squared coefficient of variation c_e^2 . Parameter t_e quantifies the mean effective capacity used for a lot by the workstation, c_e^2 quantifies the variability.

Subsequently, a so-called EPT lumped parameter (ELP) model can be built using the fitted distributions. This ELP model can be used for performance prediction and optimisation. The structure of the ELP model follows the original system in terms of the number of servers on each workstation, the buffer sizes of workstations, the flow of materials between workstations, etcetera. In this model, detailed modelling of shop-floor realities such as failures, repairs, setups, operators and lot sizes is avoided. The various sources of variability are aggregated into the EPT-distributions of the workstations. Jacobs *et al.* (2003) used the term 'meta model' rather than 'lumped parameter model'. However, the phrase 'meta model' may suggest that a simplified model is derived from another model. Since this is certainly not the case, the terminology 'lumped parameter model' is used in this paper. Here, the lumped parameters refer to the distribution parameters of the EPT-distributions.

Before the ELP model is accepted, it is validated by comparing the throughput and flow time as estimated by the model to those observed in the actual system, since one is interested in how well the lumped parameter model describes the behaviour of the actual system. If the estimated throughput and flow time are accurate enough, the ELP model and the EPT-distributions are accepted. If they are rejected, distribution fitting and model building are reconsidered. Possible changes include enhancing the level of detail of the model or using more parameters to fit more accurate distributions.

If the EPT-distributions and the ELP model are accepted, they can be used for performance analysis and optimisation. A bottleneck analysis can be carried out based on the distribution parameters t_e and c_e^2 of the various workstations. The effect of suggested improvements can be evaluated using the ELP model by accordingly adjusting the EPT distribution parameters in the model.

Implementation of the EPT-approach provides several significant advantages. First of all, shop-floor realities are included in the EPT-distributions and thus do not have to be included explicitly in the ELP model. Now, an ELP model can be obtained that is accurate, yet simple when compared to the detailed models that are typically used. Second of all, since the processing disturbances are included in the EPT-distributions, directly obtained from industrial data, the EPT parameters t_e and c_e^2 readily give insight in the behaviour of the flowline, allowing for straightforward bottleneck analysis.

3 Measuring EPT

The EPT was introduced by Hopp and Spearman (2001) to be used in queueing models. Similar concepts, such as completion time (Chen and Chen 1990, Dallery and Gershwin 1992, Buzacott and Shanthikumar 1993), are used in sample path equations. In all references, the respective distributions are assumed to be known *a priori*, and then the sample path equations are used to derive properties concerning flow time, throughput, etcetera. None of the authors, however, specifies how these distributions should be estimated from industrial data.

Jacobs *et al.* (2003) presented a method to compute EPT-realizations for infinitely buffered, isolated workstations from industrial data. Their method does not assume the effective process time distributions *a priori*, but uses a sample path equation to determine these distributions. For a single machine

workstation, the sample path equation is:

$$\text{EPT}_{i,j} = \text{AD}_{i,j} - \max(\text{AA}_{i,j}, \text{AD}_{i-1,j}), \quad (1)$$

where $\text{EPT}_{i,j}$ denotes the EPT realisation of lot i on workstation j , $\text{AD}_{i,j}$ is the departure of lot i from workstation j and $\text{AA}_{i,j}$ is the arrival of lot i on workstation j . From equation (1), one sees that an EPT realisation encompasses all time during which the server could have been processing the lot. For the events, the following equations are known: $\text{AD}_{i,j-1} = \text{AA}_{i,j}$ and $\text{AA}_{i,j} \leq \text{AD}_{i,j}$.

Algorithmic extensions have been presented for workstations with multiple parallel servers (Jacobs *et al.* 2003) and with batching (Jacobs 2004). However, the algorithms are only applicable to workstations with an infinitely large buffer. This paper studies finite buffers, which gives rise to blocking. Due to blocking, $\text{EPT}_{i,j}$ depends on events occurring on workstation $j - 1$, rendering the previous algorithms inapplicable.

For finitely buffered workstations, two additional events are introduced. Let $\text{PD}_{i,j}$ denote the possible departure of lot i from workstation j (i.e. the moment in time where workstation j finishes processing lot i) and $\text{PA}_{i,j}$ the possible arrival of lot i on workstation j (i.e. the moment in time where workstation $j - 1$ finishes processing the lot). For these new events we know that $\text{PD}_{i,j-1} = \text{PA}_{i,j}$ and $\text{PA}_{i,j} \leq \text{AA}_{i,j} \leq \text{PD}_{i,j} \leq \text{AD}_{i,j}$. Then, the sample path equation is:

$$\text{EPT}_{i,j} = \text{PD}_{i,j} - \max(\text{PA}_{i,j}, \text{AD}_{i-1,j}), \quad (2)$$

As one can see, possible occurrences of blocking (which account for the differences between a possible and actual event) are not included in the EPT realisation.

4 Lumped parameter modelling

Distribution fitting is the second phase of the EPT–approach. The relevant distribution parameters are estimated based on the measured EPT–realisations and appropriate distribution functions are proposed.

Process time distributions based on the first two moments of the distribution are often used in models of manufacturing systems consisting of workstations with infinitely large buffers. The two–moment fits are supported by queueing theory, see Buzacott and Shanthikumar (1993) and Curry *et al.* (2003).

For workstations in a flowline with finite buffer sizes, distribution fitting could be more complicated. Due to blocking, workstations are expected to affect one another more prominently. Therefore, extra information may be needed. Regardless, in queueing theoretical approaches, two moment distribution fits are used for computational reasons. However, once simulation techniques are used, the necessity of more information can be reconsidered. A typical example thereof is presented by Kim and Alden (1997). They study constant natural process times with exponentially distributed times to failure and times to repair. In the EPT–approach, all sources of disturbances are included. In addition, no assumptions regarding the distribution of the process times or disturbances are required. The necessity of additional distributional information in ELP models will be studied here.

Using simulations, the influence of the offset parameter is investigated. The offset parameter is chosen since, in practice, many operations require at least a minimum amount of time. The offset refers to the smallest possible value of a distribution. The simulation model is a flowline consisting of three unbuffered single server workstations in which lots do not overtake. The three workstations have process times distributed according to a shifted Gamma distribution. The distributional parameters are $t_e = 1.0$, $c_e^2 = 1.0$ and offset Δ_e . The offset parameter is varied from $\Delta_e = 0.0$ to $\Delta_e = 0.9$.

The corresponding simulation results are presented in figure 2. The results show that for large offsets, significant differences in throughput (δ) and flow time (φ) are observed. Increasing Δ_e from 0.0 to 0.9 results in a throughput increase of 50% and a flow time decrease of 21% (see figure 2).

The observed phenomenon can readily be explained by considering the nature of the offset. An offsetted distribution consists of a constant part, Δ_e , that is increased by a random variable with mean t_l and squared coefficient of variation c_l^2 , where $t_l = t_e - \Delta_e$. Since the variance of the process time distribution does not change, one knows that $t_e^2 c_e^2 = t_l^2 c_l^2$. Now, if $t_l = 0.1 t_e$ (i.e. $\Delta_e = 0.9$), $c_l^2 = 100 c_e^2$.

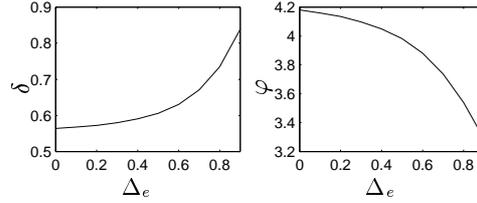


Figure 2: Influence of the offset parameter on δ and φ

Due to the large c_l^2 , most process times will be small ($\gtrsim \Delta_e$), and sporadically a value greatly exceeding the average ($\gg t_e$) will occur. The sporadic large process time realisation therefore causes massive amounts of blocking on preceding and starvation on successive workstations. If $\Delta_e = 0.0$ however, all process times will be centred around t_e . Process times will thus often be larger than t_e , frequently causing some blocking and starvation on preceding or successive workstations.

Additional simulation results, presented in Kock (2003), show that the offset parameter grows increasingly important as the amount of blocking and starvation in the flowline increases.

5 Industrial case

A case from an automotive manufacturing plant will be used to illustrate the practical applicability of the EPT–approach as described in Section 2.

System description

Experimental data has been obtained from one of the clients of Steelweld B.V. This particular client produces two types of cars, called pt_0 and pt_1 in the remainder of this section. Focus is on a small semi–synchronous flowline within the manufacturing plant. On this flowline, referred to as FL in the remainder of this section, lots are produced according to a constant product mix, i.e. $pt_0/(pt_0 + pt_1) = 0.57$. The actual sequence of lots is determined by an overhead scheduler. Since the scheduler is not considered in this case, the stream of lots entering the system will have a random lot type sequence.

FL consists of a transport system and eleven workstations in tandem (i.e. sequential). The workstations are labelled WS_0 to WS_{10} . Here, WS_1 and WS_2 are manual workstations, served by one operator. Workstations WS_7 and WS_8 are single buffer spaces. Workstation WS_{10} is used for (occasional) manual quality checks. All other workstations in the line are used for hotmelting.

First stage of the EPT–approach

The event data needed for the EPT analysis is obtained from the programmable logic controllers (PLCs) within FL . In their present configuration, only possible departures and actual arrivals can be measured using the PLCs; the actual departures and possible arrivals thus have to be reconstructed. Reconstruction is done according to equation (3). Note that $\mathbf{AD}_{i,j}$ ($\mathbf{PA}_{i,j}$) is reconstructed using the value of $\mathbf{PD}_{i,j}$ ($\mathbf{AD}_{i,j-1}$) since the necessary value of $\mathbf{AA}_{i,j+1}$ ($\mathbf{PD}_{i,j-1}$) might be unknown or wrong.

$$\begin{aligned} \mathbf{AD}_{i,j} &= \max(\mathbf{PD}_{i,j}, \mathbf{AA}_{i,j+1} - \Delta_{\min}), \\ \mathbf{PA}_{i,j} &= \min(\mathbf{PD}_{i,j-1}, \mathbf{AD}_{i,j-1}), \end{aligned} \quad (3)$$

where Δ_{\min} is the minimum amount of time required for transportation and docking. By setting $\mathbf{AD}_{i,j}$ equal to $\mathbf{AA}_{i,j+1} - \Delta_{\min}$, the transportation and docking time is included in the EPT realisation on workstation WS_j . Based on the experiences of Steelweld, Δ_{\min} is taken as a constant (the value of which is not reported here in order to respect the confidentiality of the case data).

Collection of data was not equally reliable on all workstations. For WS_1 and WS_2 , bad measurements were frequently obtained. Furthermore, no data was available for WS_7 and WS_8 . Since actual arrivals on WS_7 are unknown, actual departures on WS_6 cannot be reconstructed. Thus only the first six workstations are analysed for the case.

Since not all gathered events are correct, the data must be filtered. First of all, a number of the events result in EPT–realisations that are unrealistically low or even negative if either possible or actual arrivals are registered too late. Furthermore, since the machines are reliable, large EPT–realisations due to failures and repairs only occur sporadically. Since only a few of these realisations occur, no reliable statistics concerning these high realisations can be obtained. The EPT realisation for lot i on workstation j is thus only used during the analysis if it satisfies equation (4).

$$\Delta_{\min^*,j} \leq \text{EPT}_{i,j} \leq \Delta_{\max^*,j} \quad (4)$$

Second stage of the EPT–approach

Distribution fitting, the second stage of the EPT–approach, is done by computing the values for Δ_e , t_e and c_e^2 per workstation from the obtained filtered EPT realisations, as presented in table 1. The data in table 1 have been slightly altered, using a scaling factor, in order to respect the confidentiality of the data. Based on this data, shifted Gamma distributions were fitted for all workstations.

WS_i	t_e	c_e^2	Δ_e	$\frac{t_e}{\Delta_e}$
WS_0	82.73	0.106	57.19	1.45
WS_1	77.13	1.316	19.72	3.91
WS_2	91.89	0.894	7.89	11.65
WS_3	120.00	0.135	80.85	1.48
WS_4	112.46	0.069	70.99	1.58
WS_5	111.90	0.033	94.65	1.18

Table 1: Fitted distributions for the industrial case

Third stage of the EPT–approach

In the third stage, the shifted Gamma distributions with parameters as presented in table 1 are used as input for an ELP model, a discrete event simulation model in this case. The structure of the model is identical to the structure of FL , i.e., six unbuffered single server workstations in a flowline.

A distribution capturing the starvation observed on the first workstation has been obtained from the data to model the starvation of the first workstation in the flowline. In order to obtain this starvation distribution, a filter similar to equation (4) has been applied. The starvation distribution has properties $t_s = 63.63$, $c_s^2 = 2.564$ and $\Delta_s = 29.58$. If it is starving, the first workstation requests a lot from the generator. The generator sends a lot on to the first workstation after an appropriate period of starvation. Similarly, for the final workstation in the flowline, a distribution capturing the observed blocking is obtained. The parameters of this blocking distribution are $t_{b,5} = 15.10$, $c_{b,5}^2 = 8.04$ and $\Delta_{b,5} = 1.97$.

The true mean flow time φ of FL is determined by computing the individual flow times from the obtained data and deleting the unrealistic flow times. Flow time realisations are thus again filtered using a filter similar to equation (4). Due to filtering, some EPT–realisations are discarded during data analysis. Consequently, the mean throughput cannot be computed as the amount of bodies produced during the measured time period. Instead, mean throughput δ will be computed by determining the mean interdeparture time of bodies on workstation WS_0 .

The ELP model underestimates the throughput $\tilde{\delta}$ by 1.6%, whereas the flow time $\tilde{\varphi}$ is overestimated by 3.0% (simulation results presented in this section have a confidence level of 99% and a relative width of less than 0.1% of the mean). As can be seen, only a small error remains in the approximation. A part of this error can be explained as follows. Firstly, the ELP model assumes identically and independently distributed (iid) process times on all workstations. In the case considered here, each lot is typically of a different type than the preceding one. Since $t_{i,0}$ differs from $t_{i,1} \forall i$, a correlation is expected for successive process times on a workstation. Due to the assumption of iid process times in the ELP model,

these correlations between successive process times on a workstation are neglected. Secondly, in the ELP model, the process times of one lot on the successive workstations are assumed to be independent. In the original model however, process times for one lot on successive workstations are correlated due to the type-specific natural process times. The lumped parameter model again does not incorporate this correlation.

To improve on this, type specific EPT-distributions can be fitted, as presented in table 2. The new distributions are used in the ELP model. The model now overestimates both $\tilde{\delta}$ and $\tilde{\varphi}$ by 0.3%. By adding more detail, the approximations have become more accurate.

WS_i	pt_0				pt_1			
	t_e	c_e^2	Δ_e	$\frac{t_e}{\Delta_e}$	t_e	c_e^2	Δ_e	$\frac{t_e}{\Delta_e}$
WS_0	86.01	0.139	59.16	1.45	78.42	0.041	57.19	1.37
WS_1	40.99	1.670	19.72	2.08	128.01	0.373	67.05	1.91
WS_2	136.50	0.224	69.02	1.98	36.11	0.538	7.89	4.58
WS_3	116.40	0.216	80.85	1.44	124.80	0.037	94.65	1.32
WS_4	107.19	0.027	70.99	1.51	119.42	0.102	86.77	1.38
WS_5	115.50	0.023	100.60	1.15	107.30	0.042	94.65	1.13

Table 2: Type specific fitted distributions for the industrial case

Fourth stage of the EPT-approach

A bottleneck analysis is performed, after which the suggested improvements are simulated by accordingly adapting the EPT-distributions. It is used to determine which workstations are the major restrictions on throughput and flow time. Workstations with high t_e or c_e^2 are potential bottlenecks since they may cause starvation or blocking.

Using the information of table 2, one can see that the values of t_e range from 36.11 to 136.50. Out of this range, acceptable values of t_e seem to lie between 100 and 120 seconds (although lower values are obviously desirable). Therefore, parameters $t_{e_{1,1}}$, $t_{e_{2,0}}$ and $t_{e_{3,1}}$ are reduced to 120.00 seconds

Furthermore, table 2 illustrates that for most situations, $c_e^2 < 0.25$. Reduction of $c_{e_{1,1}}^2$ and $c_{e_{2,1}}^2$ to 0.25 is assumed to be feasible, whereas it is assumed that $c_{e_{1,0}}^2$ can be reduced to 0.75.

The suggested changes have been implemented in the ELP model. Implementation of these changes would, according to the ELP model result in an increase of 4.6% in δ and a decrease of 3.0% in φ . The simulation study with the unscaled data predicted improvements of the same order of magnitude; which was further confirmed (for the throughput) during implementation; the flow time was not studied during implementation.

6 Conclusions

A new method for performance analysis and lumped parameter modelling of single server flowlines subject to blocking has been proposed. The method is based on the effective process time (EPT). Previously, EPT has only been considered for infinitely buffered, isolated workstations. Here, a calculation method for EPT-realizations for single server flowlines subject to blocking has been presented and validated. The method translates event data (actual and possible arrivals and departures of lots) into EPT-realizations using sample-path like equations.

The EPT of a lot is the time experienced by the lot on a workstation from a logistical perspective. It is implemented by means of an approach consisting of four stages, the so-called EPT-approach. In the first stage, EPT-realizations are gathered from industrial data. Next, the realisations are translated into distributions. Typically, distributions are fitted using the first two moments (t_e , c_e^2). Simulation results

however show that the offset Δ_e should be used as an additional distribution parameter. In the third stage, an ELP model can be built and validated. Finally, in the fourth stage, the flowline can be optimized.

The EPT–approach has been applied to a case study taken from automotive industry. The ELP model accurately estimated both throughput and flow time. Adding more detail to the ELP model (i.e., including product type specific shifted Gamma distributions) further reduced errors to less than 0.3%. Based on the EPT–approach, changes in t_e and c_e^2 were proposed to increase throughput and to decrease flow time.

Acknowledgment

The authors would like to thank Onno Boxma, Johan Jacobs, Ton de Kok, Erjen Lefebber, Frank Nijssse (of Steelweld B.V.), Marcel van Vuuren, Sven Weber and Steelweld B.V. for their assistance.

References

- BUZACOTT, J.A. and SHANTHIKUMAR, J.G. (1993). *Stochastic models of manufacturing systems*. Prentice Hall, Englewood Cliffs, New Jersey, 1st edn.
- CHEN, L. and CHEN, C.L. (1990). A fast simulation approach for tandem queueing series. In O. Balci, R.P. Sadowski & R.E. Nance, eds., *Proceedings of the 1990 Winter Simulation Conference*, 539–546.
- CURRY, G.L., PETERS, B.A. and LEE, M. (2003). Queueing network model for a class of material-handling systems. *International Journal of Production Research*, **41**, 3901–3920.
- DALLERY, Y. and GERSHWIN, S.B. (1992). Manufacturing flow line systems: a review of models and analytical results. *Queueing Systems: Theory and Applications*, **12**, 3–94
- HOPP, W.J. and SPEARMAN, M.L. (2001). *Factory physics: foundations of manufacturing management*. London: Irwin McGraw-Hill, 2nd edn.
- HOPP, W.J., SPEARMAN, M.L., CHAYET, S., DONOHUE, K. and GEL, E.S. (2002). Using an optimized queueing network model to support wafer fab design. *IIE Transactions*, **34**, 119–130.
- JACOBS, J.H., ETMAN, L.F.P., CAMPEN, E.J.J. VAN and ROODA, J.E. (2001). Quantifying operational time variability: the missing parameter for cycle time reduction. In *2001 IEEE/SEMI Advanced semiconductor manufacturing conference*, 1–10.
- JACOBS, J.H., ETMAN, L.F.P., CAMPEN, E.J.J. VAN and ROODA, J.E. (2003). Characterization of operational time variability using effective process time. *IEEE Transactions on Semiconductor Manufacturing*, **16**, 511–520.
- JACOBS, J.H. (2004). *Performance quantification and simulation optimization of manufacturing flowlines (Ph.D. thesis)*. Eindhoven University of Technology, Eindhoven.
- KIM, D.S. and ALDEN, J.M. (1997). Estimating the distribution and variance of time to produce a fixed lot size given deterministic processing times and random downtimes. *International Journal of Production Research*, **35**, 3405–3414.
- KOCK, A.A.A. (2003). *Performance Evaluation and Simulation Meta Modelling of Single Server Flow Lines subject to Blocking: an Effective Process Time Approach*. Master's thesis, Eindhoven University of Technology, Department of Mechanical Engineering, Systems Engineering Group, SE 420367-A.
- ROSSETTI, M.D. and CLARK, G.M. (2003). Estimating operation times from machine center arrival and departure events. *Computers and Industrial Engineering*, **33**, 493–514.
- SABUNCUOGLU, I., EREL, E. and KOK, A.G. (2002). Analysis of assembly systems for interdeparture time variability and throughput. *IIE Transactions*, **34**, 23–40.