REAL-TIME DATA-DRIVEN AVERAGE ACTIVE PERIOD METHOD FOR BOTTLENECK DETECTION

M.SUBRAMANIYAN[1], A. SKOOGH[1], M. GOPALAKRISHNAN[1] & A. HANNA[2]
[1]Department of Product and Production Development, Chalmers University of Technology, Gothenburg, Sweden
[2]Manufacturing Research and Advanced Engineering, Volvo Group Trucks Operations, Gothenburg, Sweden

ABSTRACT
Prioritising improvement and maintenance activities is an important part of the production management and development process. Companies need to direct their efforts to the production constraints (bottlenecks) to achieve higher productivity. The first step is to identify the bottlenecks in the production system. A majority of the current bottleneck detection techniques can be classified into two categories, based on the methods used to develop the techniques: analytical and simulation based. Analytical methods are difficult to use in more complex multi-stepped production systems, and simulation-based approaches are time-consuming and less flexible with regard to changes in the production system. This research paper introduces a real-time, data-driven algorithm, which examines the average active period of the machines (the time when the machine is not waiting) to identify the bottlenecks based on real-time shop floor data captured by Manufacturing Execution Systems (MES). The method utilises machine state information and the corresponding time stamps of those states as recorded by MES. The algorithm has been tested on a real-time MES data set from a manufacturing company. The advantage of this algorithm is that it works for all kinds of production systems, including flow-oriented layouts and parallel-systems, and does not require a simulation model of the production system.
Keywords: average active duration, bottleneck detection, data-driven algorithm, maintenance, production system.

1 INTRODUCTION
Digitalization has recently gained substantial attention in the manufacturing industry and is seen as a corner stone of future production. Manufacturing companies currently capture almost every instant of time in machine activities with the help of Manufacturing Execution Systems (MES) [1]. This means that manufacturing collects huge amounts of data on the machines, often referred to as 'Big Data'. A preliminary study at an automotive company in Sweden shows that, on average, 100 data rows per hour are recorded per machine. This means 500 000 recorded data rows per year for just one machine [2]. Many companies have started looking for better ways to support the real-time processing of big data with the help of advanced analytics [3].

In order to stay competitive, manufacturing industries must focus on improving the productivity of the production system. Prioritisation of improvements and maintenance activities is an important part of the production management process to improve productivity [4]. The improvement activities should be carried out on the constraint resources in the production system to get the maximum value out of the investments made. The constraints of the production system can be traced to one or more machines, called bottlenecks [5]. Thus, identifying these bottlenecks from a systems perspective is crucial.

Today, many companies identify the bottlenecks in the production system based on the gut feelings of the production personnel and this is partly because of the lack of data-driven methodologies. Most

of the available methodologies for bottleneck detection are developed in a simulation environment, which takes more time to build, execute and analyse a model to support the decision-making process [6]. Also, this needs expert skills in modelling, which are lacking in many companies. Simulation methods are also difficult to use as a day to day real-time decision support tool. Furthermore, simulation-based tools do not provide real-time bottleneck detection results, and this leads to the loss of maintenance opportunities and consequent loss of production [7]. On the other hand, there is an increasing need for real-time data analytics in manufacturing industries, which can be used to detect the bottlenecks of complex production systems in relatively shorter time.

The overall purpose of this paper is to improve the productivity in manufacturing industries by facilitating the production and maintenance decision-making process using the real-time data-driven approach. The aim of the paper is to propose a real-time data-driven algorithm for an existing simulation-based bottleneck detection method, called the average active period method, as proposed by Roser *et al.* [8].

2 LITERATURE REVIEW

2.1 Bottleneck detection methods

There are various types of bottleneck detection methods in the manufacturing context available in the literature. Table 1 summarises the different bottleneck detection methods, providing a clear classification of the existing literature, based on the types of the tools used to develop the method.

From Table 1, it can be seen that most bottleneck detection methods were built using simulation as the support tool, that is, these methods are built using the model of the production line and applied the simulation techniques to develop the methods. The turning point method is the only available method in the literature to identify the bottlenecks from the real-time data without resorting to the simulation model of the production system.

Overall, it is clear that there is only one bottleneck detection technique which is completely data-driven, and it can be inferred from Table 1 that there is limited research reported in determining the applicability of simulation-based bottleneck detection methods over real-time data without building a simulation model. Hence, in this paper, the real-time data-driven algorithm is developed for the simulation-based average active period method proposed by Roser *et al.* [8] without building a simulation model of the production line. This method was chosen out of the other simulation-based approaches because it not only detects the long-term bottlenecks with high accuracy, but is also capable of finding a group of bottleneck machines in the production line.

Table 1: Summary of different bottleneck detection methods and tools used for their development.

Analytical model	Simulation based	Real-time data-driven
Sensitivity based bottleneck detection [9–12]	Average waiting time [13]	Turning point method [20]
	Longest Queue [14]	
	Longest waiting time [15]	
	Active period percentage [8]	
	Average active period [8]	
	Shifting bottlenecks [16]	
	Utilisation [17]	
	Sensitivity based bottleneck detection [18]	
	Interdeparture time variance [19]	

2.2 Average active period method

The average active period method is used for identification of the long-term static bottleneck [8]. A long-term bottleneck is defined as that machine in the production line which has the highest bottleneck time, when compared to all other machines, throughout a given interval of time. The first step in this method is to classify each machine status into active and inactive states. An active state is that state of the machine when the machine produces a part or is being serviced or setup, whereas the inactive state of the machine is when the machine is waiting for the part or waiting to be serviced or waiting for the removal of the parts from the machine [8]. Consecutive active periods are taken as a single active period as shown in Fig. 1 (adapted from Roser *et al* [8]). According to this method, the machine with the highest average active time duration is the bottleneck machine in the production line.

The following equations are proposed by Roser *et al.* [8] to calculate the average active period of the machine.

Let A = {a_1, a_2…, a_{n-3}… a_n} be a set consisting of each active period duration of the machine *i*. The average active period of A_i is denoted by, the standard deviation of the active period set is denoted by $\overline{A}$ and the confidence interval at 95% is denoted by *CI*.

$$\overline{A} = \left(a_1 + a_1 + a_3 \ldots + a_{n-3} \ldots a_n\right)/n \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{i=n}\left(a_i - A\right)^2}{n-1}} \tag{2}$$

$$CI = t_{a/2}\,\frac{\sigma}{\sqrt{n}} \tag{3}$$

Bottleneck machines are determined based on the average of all machines. The confidence interval is calculated for each machine so as to estimate the accuracy of the bottleneck measurement and to determine the ties between the machines if the confidence intervals of two machines overlap.

3 OVERVIEW OF THE APPROACH

The primary focus of this paper is the investigation and development of the average active period algorithm for real-time applications and to test the algorithm with the real-world empirical data. Figure 2 represents the over view of the work flow.

The development phase of the algorithm involves a thorough literature study of the simulation-based average active period method. In addition to the literature study, a representative single case study approach was also adopted [21]. This involved detailed study of the structure of empirical
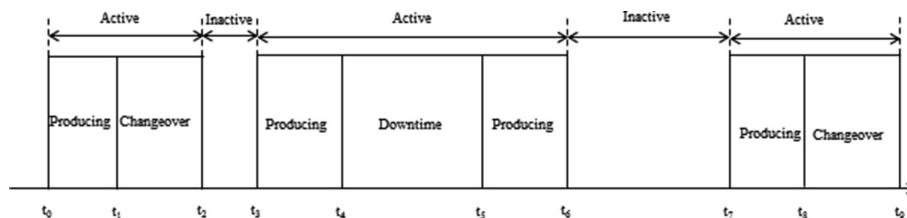

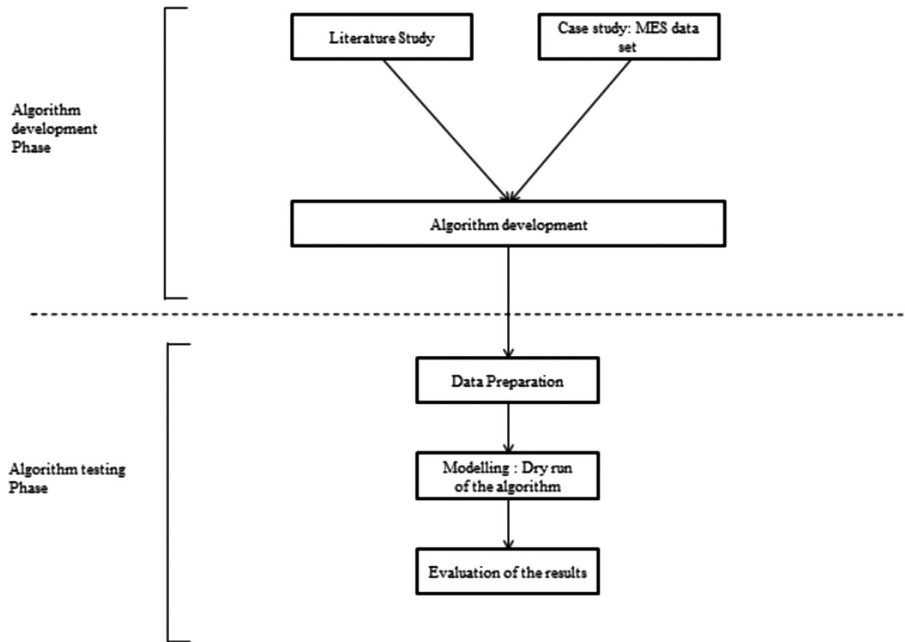Figure 1: Active periods and inactive states of the machine.

Figure 2: Overview of the work flow.

MES data of one production line at a case company as shown in Table 2. The output from this phase is the detailed algorithm and the definition of the inputs and the outputs of the algorithm.

The testing phase of the algorithm is based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) model. This is a hierarchical process model for data analysis and provides a five-step standard methodology for a data mining project [22]. Specifically, steps 3, 4, and 5 described in the model are used in the testing phase as shown in Fig. 2 and the algorithm is tested on the MES data of the case company of the production line. The testing phase is also a part of the validation of the algorithm. The strategy is to test the algorithm in a real-world scenario and thereby evaluates the applicability of the algorithm over the real-time data. The data preparation step in the testing phase includes the selection of the data formats from the MES system and clean-ing of data if necessary. The modelling phase includes running the algorithm over the cleaned

Table 2: Sample MES record of the machine status with time stamps.

Production area	Work area	Date with time	State of the machine
Line 1	M1	01-09-2014 06:28:02	Not Active
Line 1	M1	01-09-2014 06:28:25	Comlink Up
Line 1	M1	01-09-2014 06:29:20	Not Active
Line 1	M1	01-09-2014 06:29:34	Waiting
Line 1	M1	01-09-2014 06:29:34	Waiting
Line 1	M1	01-09-2014 06:42:46	Producing

data. The evaluation phase includes the interpretation of the results obtained from the application of the algorithm.

4 STUDY OF THE PROPOSED REAL-TIME DATA-DRIVEN AVERAGE ACTIVE PERIOD ALGORITHM

After a thorough literature review and a study of the MES data structure of the case company, the average active period algorithm was developed. The proposed algorithm uses the MES data from individual machines and analyses this to identify the bottleneck machines. The algorithm requirement in terms of inputs, consists of the following three types of information:

- *Period interval*: Period for which the analysis has to be made to identify the bottleneck machine
- *Time interval*: Start and the stop of the shift each day
- *States of the machines*: Different machine states classified into two categories; active and inactive for each time instant in the time interval (See Section 2.2)

4.1 Notations

d = Start day of the period
D = End day of the period
t = Start time instant of the production shift
T = End time instant of the production shift
j = Increment variable for time instant
n = Increment variable for the number elements in the set
$\{a\}$ = Active period set; $(a_1, a_2, a_3, a_4, \ldots)$

4.2 Algorithm and flow chart

The following is a general description of the step-by-step procedure in the algorithm:
    Step 1: Specify the time period for which the calculation to be made (specification of the d and D)
    Step 2: Specify the shift timings of the day (specification of t and T)
    Step 3: Classify the states of the machine in terms of active and inactive states
    Step 4: If $d< = D$, then check whether $t< = T$
    Step 5: Check whether the machine is active at the time instant of t
    Step 6: Store the active periods in the matrix
    Step 7: Continue until $t>T$ and $d>D$
    Step 8: Calculate the mean active period according to eqn (1)
    Step 9: Calculate the standard deviation according to eqn (2)
    Step 10: Calculate the confidence interval at 95% according to eqn (3)
    Step 11: Perform Step 3 to Step 10 for all machines and compare the results

In addition n to the algorithm, a flow chart as shown in Fig. 3 is drawn to show the iteration process of different steps in the algorithm to calculate the mean active period with the standard deviation for a machine. When carried out separately for each machine in the production line and by comparing the average active period durations of each machine, the bottleneck machine is determined, because the bottleneck is that machine with the highest average active period among other machines. Also, if the confidence interval of the bottleneck machine overlaps with other machines, then a group of bottleneck machines can be revealed.
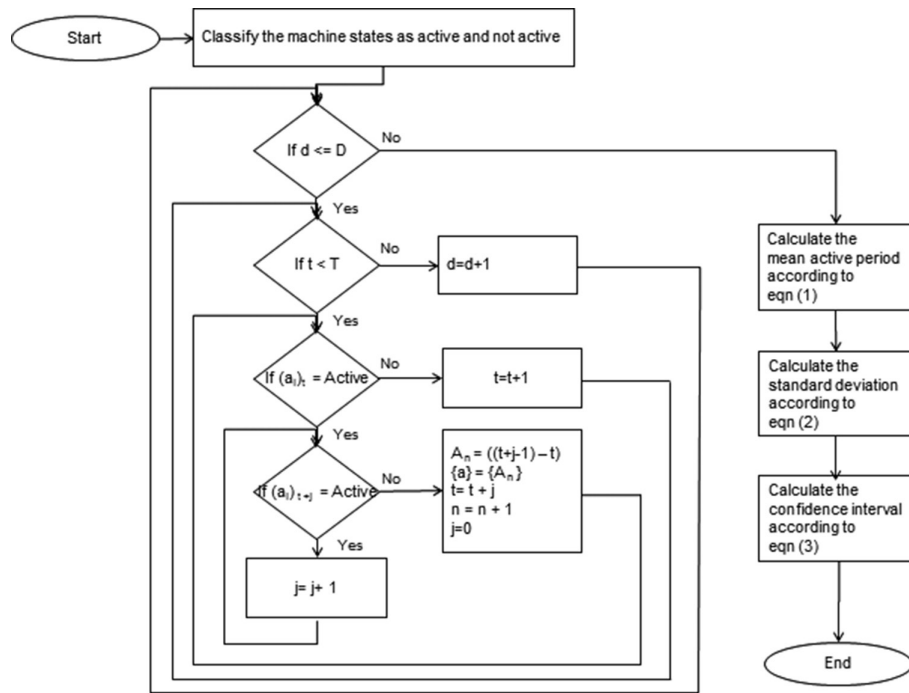
Figure 3: Flow chart for data-driven average active period algorithm.

5 TEST PHASE OF THE ALGORITHM

To test the algorithm, MES data from a production line was collected from the case company. The production line has 17 machines arranged in a production layout as shown in Fig. 4. The production line starts with machine M1 and M2 and ends at machine M16 and M17. Machines M1 and M2, M4 and M5, M7 and M8, M10 and M11, M13 and M14, M16 and M17 are parallel machines. Robots (M3, M6, M9, M12, M15, and M18) are used for handling the product between the parallel set of machines. There are buffers between each parallel set of machines. All machines are automatic, so no operator is required to run the machines.

Data preparation: The MES data for each machine and robots were collected for 44 scheduled production days corresponding to 748 hours of scheduled production time. The MES data consist of the time stamps of the machine along with the states of the machines. The time stamps were checked for overlaps and the data were cleaned by removing the data rows of information outside the scheduled production days. The sample MES information describing the machine status record on the shop floor is shown in Table 2.

Modelling: The first step in the algorithm is to classify machine states as active and inactive states. There are eight states defined for each machine and robot in the production line. Those are *Producing, Part Changing, Error, Comlink Down, Comlink Up, Waiting, Not Active, and Empty Run.* The definitions of these states are explained in Table 3. All the machines and robots are at any given time in one of the eight given states. The machine states are divided into active or inactive states based on the definition of active and inactive periods. The active states of the machine include *Producing, Part Changing, Error, Comlink Down, Comlink Up, and Empty Run.* The states, *Comlink Up and Comlink Down* are classified as active states because the machine
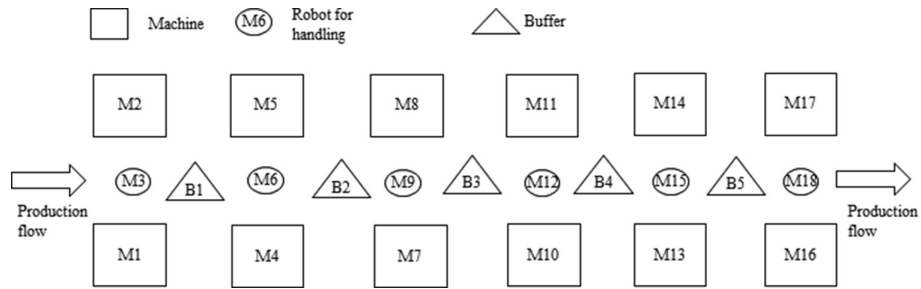
Figure 4: Layout of the production line.

Table 3: Definitions of the states of the machine.

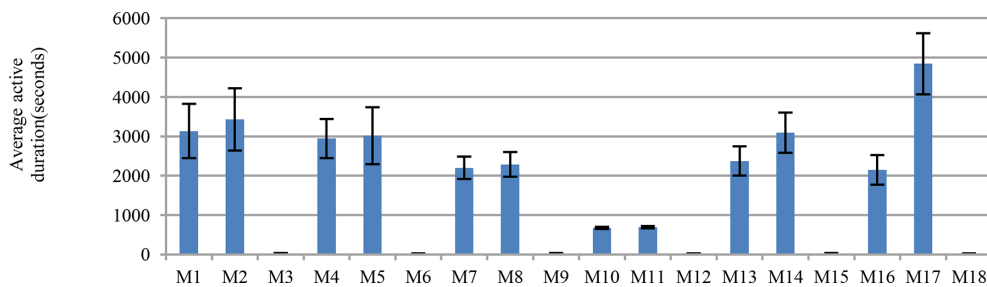| Machine states	Explanation
Producing	Machine is engaged in producing a product
Part changing	Machine in undergoing a setup
Error	Machine is taken down due to tool or machine failure
Comlink down	Connection between MES system and the main server is interrupted
Comlink up	Connection between MES system and the main server is resumed
Waiting	When the machine is blocked or starved
Not active	When the machine stops apart from above mentioned reasons
Empty run	When the machine produces new products or trial products

Figure 5: Average active duration of the machines in the production line.

does not stop when the MES of the machine is down. The inactive states include *Not active* and *Waiting*.

The proposed data-driven algorithm is applied over the MES data of the production line, and the average active duration of each machine and of robots with a 95% confidence interval was revealed as shown in Fig. 5.

Evaluation: From Fig. 5, it can be seen that the duration of the active period M17 is the largest duration in the production line but the confidence interval of machine M17 overlaps with the confidence interval of machine M2 (~3.6% overlap), which indicates a tie in identifying the primary bottleneck. Hence, it could be concluded that there is a group of potential primary bottlenecks in the production system and that the sole primary bottleneck is difficult to identify.

6 DISCUSSION

From this study, it is clear that the real-time data-driven average active period algorithm can be developed from the empirical MES data to identify a group of potential bottlenecks in the production system. The results from the case study demonstrate that the average active period can also be used in the production system with parallel machines to detect the bottlenecks. This is because the data-driven approach can correlate between the parameters and interactions between the machines in the production system. Also, the average active period method points to a group of bottlenecks which could be very useful when prioritising maintenance activities, as pointed out by Gopalakrishnan *et al.* [4].

The proposed data-driven algorithm for average active periods has many advantages over the simulation-based average active period method. Firstly, the bottleneck machines can be revealed using the real production data, without building a simulation model of the production system. Hence, bottleneck detection can be made quickly and the application of algorithm and interpretation of results does not demand detailed expert skills such as in simulation as indicated by Fowler and Rose [6]. Secondly, the proposed data-driven algorithm is scalable and can be used to find the bottlenecks on a daily basis or on an hourly or weekly basis. Thus, the algorithm can also detect short-term bottlenecks (the machines which impede the production system performance for short durations of time) which can be used for maintenance task prioritisation according to Li and Ni [7]. Thirdly, the algorithm is independent of the nature of the production system. In other words, it is dependent only on the machine. This allows a complete flexibility to remove certain machines from the analysis if needed. The algorithm uses only simple machine information as input to reveal the bottlenecks and can be easily linked to the MES system. Lastly, the algorithm could be applied to complex production systems with parallel machines and with buffers. Although there are a number of advantages, there are also some limitations with the data-driven approach. The data-driven approach can only be used when the historical machine data provides a clear description of machine activity, and when the time stamps of those activities are available. Moreover, input data need to be cleaned prior to the algorithm run. For example, the overlaps of the time stamps of the different states of the machine need to be corrected, and the unscheduled production times need to be removed from the data set prior to the application of the algorithm.

With the use of this data-driven algorithm, group of bottlenecks from a production system perspective can be revealed. Maintenance planning activities (scheduling preventive maintenance etc.) and improvement activities (cycle time reduction etc.) can then be prioritised in the bottleneck machines to obtain the maximum value from the investments made and to continuously improve the productivity.

7 FUTURE WORK

Our future work would consist of applying the proposed algorithm over other case companies within the StreaMod project to investigate the generalisability of the algorithm and obtain a deeper understanding of the online production measurements as recorded by MES. And furthermore, to develop data-driven algorithms for other types of bottleneck detection approaches.

8 CONCLUSION

A real-time data-driven average active period bottleneck detection algorithm was proposed and developed in this paper. Further, it was tested over the real-time production empirical data set of a case company in the manufacturing industry. This work addresses the problem of detecting the bottlenecks without building a simulation model and suggests a practical and simple approach to accomplish that end. It is also proven that the algorithm can easily detect the bottlenecks in the

production line for any given time period and hence shows the scalability of the algorithm. Also, this algorithm can be integrated with the existing MES system within the manufacturing companies, so that decisions on mitigating the bottlenecks can be taken based on real-time data.

ACKNOWLEDGMENTS

REFERENCES

[1] Kletti, J., (ed), *Manufacturing Execution Systems – MES*, Springer: Berlin, 2007.
http://dx.doi.org/10.1007/978-3-540-49744-8

[2] Subramaniyan, M., Production data analytics – to identify productivity potentials. *Master Thesis, Chalmers University of Technology,* 2015.

[3] Donovan, P.O., Lahy, K., Bruton, K. & Sullivan., D.T.J.O., Big data in manufacturing: a systematic mapping study. *Journal of Big Data*, **2**(1), pp. 2–20, 2015.

[4] Gopalakrishnan, M., Skoogh, A. & Laroque, C., Simulation-based planning of maintenance activities by a shifting priority method. *Proceedings of 2014 Winter Simulation Conference,* eds A. Tolk, S.Y. Diallo, I.O. Ryzhov, L. Yilmaz, S. Buckley & J.A. Miller, pp. 2168–2179, 2014.
http://dx.doi.org/10.1109/wsc.2014.7020061

[5] Goldratt, E.M. & Cox, J., *The Goal: A Process on Ongoing Improvement*. North River Press, 1992.

[6] Fowler, J.W. & Rose, O., Grand challenges in modeling and simulation of complex manufacturing systems. *Simulation*, **2**(9), pp. 469–476, 2004.
http://dx.doi.org/10.1177/0037549704044324

[7] Li, L. & Ni, J., Short-term decision support for maintenance task prioritisation. *International Journal of Production Economics,* **121**(1), pp. 195–202, 2009.
http://dx.doi.org/10.1016/j.ijpe.2009.05.006

[8] Roser, C., Nakano, M. & Tanaka, M., A practical bottleneck detection method. *Proceedings of 2001 Winter Simulation Conference,* eds B.A. Peters, J.S. Smith, D.J. Medeiros & M.W. Rohrer, pp. 949–953, 2001.
http://dx.doi.org/10.1109/WSC.2001.977398

[9] Kuo, C.-T., Lim, J.-T. & Meerkov, S.M. Bottlenecks in serial production lines: a system-theoretic approach. *Mathematical Problems in Engineering,* **2**, pp. 233–276, 1996.
http://dx.doi.org/10.1155/S1024123X96000348

[10] Chiang, S.-Y., Kuo, C.-T. & Meerkov, S., Bottlenecks in markovian production lines: a systems approach. *IEEE Transactions on Robotics and Automation*, **14**(2), pp. 352–359, 1998.
http://dx.doi.org/10.1109/70.681256

[11] Chiang, S.Y., Kuo, C.-T. & Meerkov, S., DT- bottlenecks in serial production lines: theory and application. *IEEE Transactions on Robotics and Automation*, **16**(5), pp. 567–580, 2000.
http://dx.doi.org/10.1109/70.880806

[12] Chiang, S.Y., Kuo, C.-T. & Meerkov, S., c-Bottlenecks in serial production lines: identification and application. *Mathematical Problems in Engineering*, **7**, pp. 543–578, 2001.
http://dx.doi.org/10.1155/S1024123X01001776

[13] Law, A.M. & Kelton, D.W., *Simulation Modeling and Analysis*, 3rd edn.*,* McGraw Hill: New York, 1991.

[14] Lawrence, S.R. & Buss, A.H., Shifting bottlenecks: causes, cures and conundrums. *Production and Operations Management*, **3**(1), pp. 21–37, 1994.
http://dx.doi.org/10.1111/j.1937-5956.1994.tb00107.x

[15] Law, A.M. & Kelton, D.W., *Simulation Modeling and Analysis*, 3rd edn.*,* McGraw Hill: New York, 2000.

[16] Roser, C., Nakano, M. & Tanaka, M., Shifting bottleneck detection. *Proceedings of 2002 Winter Simulation Conference,* eds E. Yucesan, C.H. Chen, J.L. Snowdon & J.M. Charnes, pp. 1079–1086, 2002.
http://dx.doi.org/10.1109/wsc.2002.1166360

[17] Roser, C., Nakano, M. & Tanaka, M., Comparison of bottleneck detection methods for AGV systems. *Proceedings of 2003 Winter Simulation Conference,* eds S. Chick, P.J. Sanchez, D. Ferrin & D.J. Morrice, pp. 1192–1198, 2003.
http://dx.doi.org/10.1109/WSC.2003.1261549

[18] Chang, Q., Ni, J., Bandyopadhyay, P., Biller, S. & Xiao, G., Supervisory factory control based real-time production feedback. *Journal of Manufacturing Science and Engineering*, **129**(3), pp. 653–630, 2007.
http://dx.doi.org/10.1115/1.2673666

[19] Betterton, C.E. & Silver, S.J., Detecting the bottlenecks in serial production lines – a focus on interdeparture time variance. *International Journal of Production Research*, **50**(15), pp. 4158–4174, 2012.
http://dx.doi.org/10.1080/00207543.2011.596847

[20] Li, L., Chang, Q., Ni, J. & Biller, S., Data driven bottleneck detection of manufacturing systems. *International Journal of Production Research,* **47**(18), pp. 5019–5036, 2009.
http://dx.doi.org/10.1080/00207540701881860

[21] Yin, R.K., *Case Study Research: Design and Methods*, Sage Publications Inc: Thousand Oaks, CA, 2003.

[22] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. & Wirth, R., CRISP-DM 1.0 Step-by-Step Data Mining Guide, SPSS, available at https://www.the-modeling-agency.com/crisp-dm.pdf, 2015.