Comparing Performance Measures for the Trade Off of Flow Time and Throughput in Complex Manufacturing Systems

M. Macchi

Politecnico di Milano, Dipartimento di Ingegneria Gestionale
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
WWW home page: http://www.dig.polimi.it

Abstract. Management of the trade off of logistic performances has always been a key issue in industry. The trade off results from the contrasting business objectives defined at strategic management level and is part of the subsequent decisions taken at shop floor level. Therein, performance measures influence decisions. These are, in general, operational measures and might be more or less aware of the trade off fixed at strategic level. The present paper aims to demonstrate how awareness may change in a production context where both flow time and throughput assumes a strategic importance. The demonstration is achieved by means of 2 case studies: the first one is simulated from scratch in laboratory, the second one is a simulation from a real industrial setting.

1   Introduction

Performance measurement of the trade off of logistic performances is clearly not a novelty. If one looks over literature, in the industrial engineering and operations management arena, 2 approaches for performance measurement can be identified. A typical approach is to define a cockpit of measures (throughput, flow time, …) [3]. A relatively new approach is to define some indicators of logistic efficiency, proposed to characterise the trade off of logistic performances in a synthetic measure, instead of a cockpit. Also these indicators are derived from basic measures (throughput, flow time, …) [2,7]. In this paper, the 2 approaches are compared to show the different decisions that they may lead to. The decisions are also evaluated with respect to defined strategic objectives, to deduce if the performance measures permit to be aware of them. The performance measures are, firstly, selected after a state of the art analysis (par. 2). Their theoretical and experimental comparison is shown (par. 3, 4). Concluding remarks are eventually pointed out for future works (par. 5).

2   State of the art

A cockpit of performance measures is the typical approach to support decisions concerned with the trade off. This approach can be interpreted as a natural follow up of theoretical backgrounds of the Little's law. According to this approach, a decision at operational level can be reached after a concurrent analysis of throughput and flow time. The trade off of these measures can be shown in typical diagrams, such as flow time – throughput diagram. This diagram is referred to a production system or some of its subsystems and clearly leads a decision maker to a concurrent analysis of the 2 measures. Hence, decisions are somewhat aware of the trade off. Other alternative measures can be introduced in the cockpit. OEE [5,6] is well known. It is adopted, in TPM, for reducing efficiency losses at equipments of the production system. Its metric is correlated to the system throughput. Besides, even if the improvement in OEE is expected to bring benefit in flow time, OEE does not really count for it. OEE is then usually adopted aside other measures, such as flow time itself, and analysed concurrently with them to take into account the trade off. E metric [1] is a slight modification of OEE: similar notes can be pointed out to its concern. OTE [4], instead, complements OEE, by directly measuring an efficiency ratio of the achieved system throughput with respect to the maximum achievable throughput (achievable by its bottleneck resources). OTE still lacks of consideration of the flow time efficiency. Again, it requires other measures so that the decision is aware of the trade off. The second approach is to measure system performances by means of synthetic indicators of logistic efficiency. 2 indicators may be cited. The first one is the manufacturing performance P [2], the second one is OFE (Overall Fab Efficiency [7]). These are intended to measure, in a unique model, the efficiency of flow time and throughput. P metric is detailed in next par. 3. OFE is built as a product of 3 factors. The production efficiency (pe) is the factor that measures the trade off, normalised according to a reference situation, the so called "Practical Worst Case", defined as a production line, whose stations are single machines and balanced, with exponential distribution of processing times and CONWIP control.

3   Theoretical comparison

This section compares how decision is reached using the flow time – throughput and P – throughput diagrams. Attention is paid to one of the main decision making problem faced when solving the trade off: to decide the best operating point of the production system, if there is need to improve its efficiency (by increased utilisation or, equivalently, throughput) or to improve its responsiveness (by reduced utilisation, leading to a reduced flow time). Decision is supposed constrained by 2 objectives fixed at strategic level, and constraining the operational level. The first constraint is "max flow time". This cannot be overcome being it not acceptable for the operations strategy. It can be, e.g., the case of a MTO (make to order) operations strategy when order delivery time overcoming "max flow time" is not competitive in the market. The second constraint ("min throughput") may regard, e.g., the Break Even Point of

a production system. These 2 constraints are included in the flow time – throughput and P – throughput diagrams, to create awareness of the strategic objectives.

3.1   Flow time – throughput diagram

Next figure 1 shows an example of flow time – throughput diagram with strategic constraints. According to such a diagram, a decision maker should decide to operate the system at an operating point over 21 orders / shop calendar day, so the Break Even Point is guaranteed. He/she should not exceed 26, otherwise the achieved flow time would not be competitive in the market. On the other hand, the diagram does not provide any criterion to decide if there is a best operating point in between.
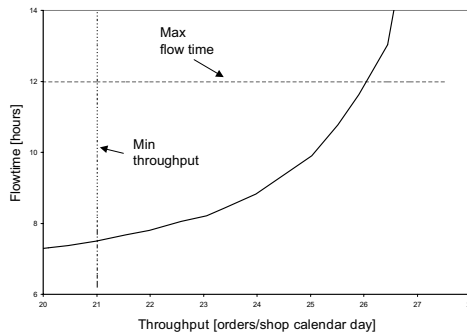

Fig. 1. Flow time – throughput diagram with strategic constraints (example)

3.2   P – throughput diagram

P metric is product of 2 efficiencies, flow time and throughput efficiency, calculated with respect to reference measures. Hence, $P = \eta_\varphi * \varepsilon_\delta$, where: $\eta_\varphi = \varphi_{ref} / \varphi$; $\varepsilon_\delta = \delta / \delta_{ref}$. φ_{ref} is normally fixed at the minimum achievable flow time (sum of all processing times required in process plans). δ_{ref} is normally fixed as the maximum achievable throughput (achievable by the bottleneck resources of the system). Next figure 2 shows an example of P – throughput diagram with strategic constraints. The representation of the strategic constraints is obtained by fixing acceptable thresholds such as a minimum throughput ($\delta = \delta_{min}$) or a maximum flow time ($\varphi = \varphi_{max}$). According to such a diagram, the operating point should be again between 21 and 26 orders / shop calendar day. The decision maker is then aware of the strategic objectives, similarly as with the flow time – throughput diagram. On the other hand, this diagram provides a new criterion to decide the best operating point in between: optimum point of P is reached at a throughput of about 22 orders / shop calendar day. Moreover, P is close to optimum in the range between 21 and 22. This was not fixed in the flow time – throughput diagram.
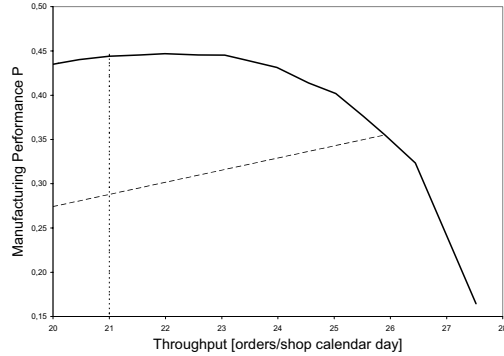
Fig. 2. P – throughput diagram with strategic constraints (example)

3.3   Remarks on the strategic constraints

Other constraints may be identified at strategic level. Diverse levels of flow time may be, e.g., competitive in a market, depending on how much the client is sensible to the system responsiveness: e.g., "min flow time" might be the level over which a client is eager to accept a worse logistic service (reduced delivery time), but only in the case that the supplier offers discounts or pays some penalties with respect to the full cost; whilst the "max flow time" might still stand as a not acceptable level. A similar situation is considered in the second case study analysis (par. 4.2).

4   Experimental comparison

2 case studies were simulated in ARENA tool for experimental comparison. The simulation results were then used to build flow time – throughput and P – throughput diagrams. The simulation results are reported in next subsections: the first case study is used to compare decisions resulting from flow time – throughput and P – throughput diagrams; the second case study integrates strategic / economic issues, a comparison on how decision changes is then done by adopting the P – throughput diagram (with strategic constraint) and a Gross margin – throughput diagram.

4.1   Case study 1 – job shop system

The first case study (fig. 3) is a job shop with 3 shops (WS1, WS2, WS3). Each WS is a set of parallel machines (with equivalent processing capabilities). The order mix is fixed for type (P1, P2, P3, P4) and percentage (30%, 35 %, 25 % and 10%). Unit processing and set up times are fixed as well. Inter-arrival times at the job shop are exponentially distributed and a FIFO dispatching rule is used at each machine. Other issues (transport, failures, buffer spaces, …) are neglected.
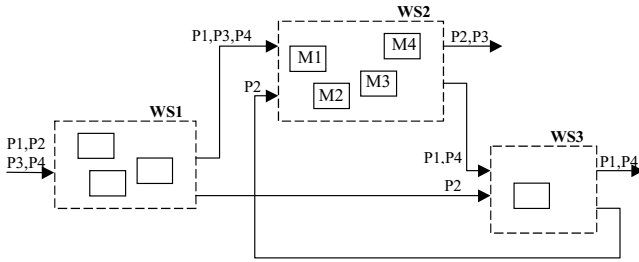
Fig. 3. Production flows and machines in the job shop

Being all fixed, some alternatives for product routing through the WS2 machines (WS2 is the bottleneck resource) are then compared: rule 1 (routing "all products to all machines" using a random rule), rule 2 (routing "all products to all machines" using a rule for selection of the machine with "minimum queuing times") and rule 3 (with rigid product-machine allocations, i.e. M1 and M2 allocated to products with high work loads – M1 allocated to P1, M2 to P2 –, M3 to products with low work loads – P3, P4 –, M4 to all products, to provide a degree of routing flexibility –). The simulation results are shown in the following P – throughput and flow time – throughput diagrams (fig. 4).
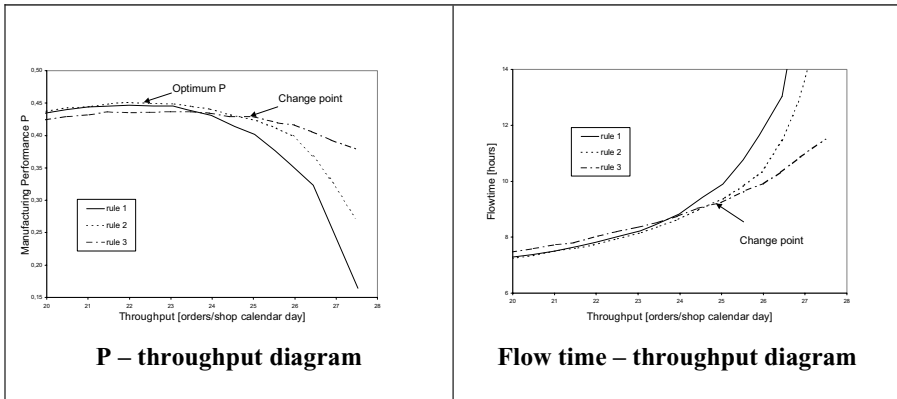

Fig. 4. Selecting the best operating point in case of alternative routing rules

According to the P - throughput diagram, rule 2 is preferred at lower throughputs, rule 3 outperforms other rules at higher throughputs. In fact, at lower throughputs, rule 2 outperforms rule 1 thanks to its routing criterion ("minimum queuing time") instead of random routing. Rule 3 is, any how, the worst one, since it suffers from waiting times subsequent to the constraints in product routings created by the rigid product-machine allocations: an order may risk to wait much of the times being its dedicated resource busy, the worst flow time then results which, finally, reduces the P performance. A change point comes out at higher levels (close

to 25 orders / shop calendar day). The losses of the flow time efficiency are now due to the set up times (plus the induced waiting times) much more registered with the flexible solutions (rule 1, 2). These losses are smoothed with dedicated product-machine allocations (rule 3). The same ranking (rule 2 preferred at throughputs lower than 25, rule 3 preferred at throughputs higher than 25) is reached with the flow time – throughput diagram. Conversely, the flow time – throughput diagram does not help to fix the position of the optimum of the trade off as done accordingly with the P criterion. Indeed, the optimum P is achieved by the flexible routing rule 2, at throughput equal to 22 orders / shop calendar day.

4.2   Case study 2 – flow shop with re-entrant flows

The case study concerns repairing of radar components of air planes' fleets. The turn time (time to receive a radar component, test, complete repairs and return it to the fleet operator) is a strategic objective, to avoid the risk of lengthy AOG (aircraft on the ground). This requires high responsiveness from the repair shop. The quick response is rewarded: a constant premium price is rewarded by the fleet operator, if the flow time is lower than a threshold ("min flow time"). Otherwise, the premium price is reduced by a penalty cost, with an almost linear penalty cost function: if the flow time increases ft % over the "min flow time", the premium price is reduced at a p %. Different ratio p % ÷ ft % have been experimented in simulation for sensitivity analysis. When LOW, the ratio p % ÷ ft % is 5 % ÷ 10 %; when MID, the ratio p % ÷ ft % is 7,7 % ÷ 10 %; when HIGH, the ratio p % ÷ ft % is 10 % ÷ 10 %.

The repair shop is a flow shop with re-entrants: the repair cycle is, at best case, a sequence of "inspect – test – repair – test – inspect" operations, before delivering the repaired item to a fleet operator. The P – throughput diagram (fig. 5) compares then 2 rules to find improvements of the bottleneck operation ("test"). The 2 rules concern human tasks allocation. In fact, the test benches are devices with specific testing capabilities that cannot be changed due to the high investment costs. The competencies of the human operators may instead be changed. It is then possible to change the existent human tasks allocation – "operators dedicated to specific test devices" (rule 1) –, to improve the system flexibility – with "operators able to work on all test benches" (rule 2) –. An inefficiency may be reasonably expected for rule 2: the unit testing time was in fact estimated 15 % shorter in the most efficient case of rule 1 ("dedicated operators"). The P – throughput diagram is a follow up: at throughputs higher than 30 repair orders / months, the "dedicated" solution (rule 1) is preferred thanks to its efficient unit testing time; at lower throughputs, the "flexible" solution for the product routings (rule 2) wins over the advantages of the testing efficiency and should be chosen. The best operating point, according to optimisation of P, is again achieved at low throughputs (between 25 and 26 repair orders / month) by the "flexible" solution (rule 2). This may not correspond to the economic optimum if penalty costs are taken into account beside P.

Indeed, the gross margin diagram (see again fig. 5) leads to the same optimum as with P only in the case of MID p % ÷ ft %. The penalty cost is too high in the HIGH case: if the throughput is incremented over the "min flow time", the increase in repair volumes is not sufficient to counterbalance the penalty cost resulting from

worse flow time efficiency, hence the gross margin is reduced. In this case, then, the economic optimum corresponds exactly to the "min flow time" point, where major gains result from achieving premium price whilst avoiding penalty cost. The LOW case is the opposite situation: if the throughput is incremented over the "min flow time", the increase in repair volumes brings more benefit than reduction in unit gross margin due to the penalty cost resulting from worse flow time; so the gross margin is augmented. The economic optimum, in such a case, is reached by pushing production as maximum as possible according to capacity constraints.
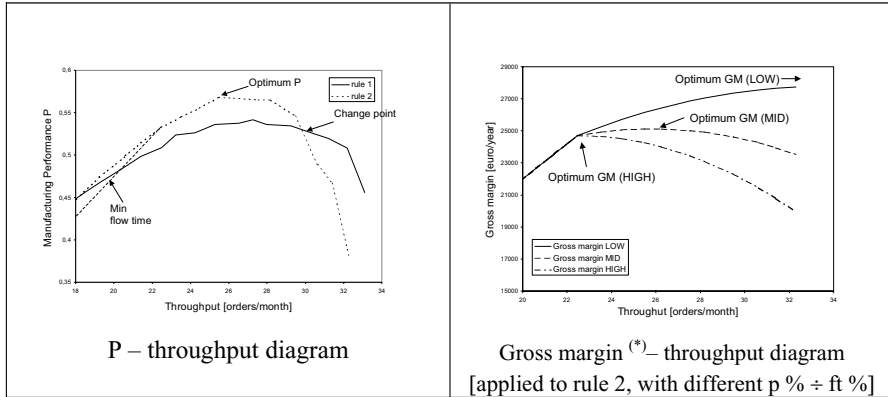

| P – throughput diagram | Gross margin (*)– throughput diagram [applied to rule 2, with different p % ÷ ft %] |

Fig. 5. Selecting the best operating point in case of alternative rules for human tasks allocation; (*) Gross Margin = [(premium price - penalty cost - other variable costs) * throughput * working days per year] (Note 1. premium price fixed at a fictitious level of 100, Note 2: other variable costs constant at different throughputs).

Thereafter, comparing with P, it can be concluded that P is a sufficient criterion to identify the best operating point when the strategic setting is in a specific situation. In the case study, being "min flow time" fixed, P can be used to optimise the trade off only when the stipulated contracts maintain a ratio p % ÷ ft % around 7,7 % ÷ 10 %, after a more detailed sensitivity analysis, it was established that P would also be sufficient in a range between 7,5 % ÷ 10 % and 8 % ÷ 10 %. In other cases: (i) better to push production to the maximum capacity, when ratio is lower than 7,5 % ÷ 10 %, using rule 1 ("dedicated operators"); (ii) better to work at minimum throughput, correspondent to the "min flow time", when ratio is higher than 8 % ÷ 10 %.

4.3   Remarks on the strategic constraints

The second case study helps to point out another constraint that may exist beside those already issued: the repair demand clearly depends on how many contracts are stipulated, so from a decision strictly related with strategic level. This repair demand might, however, lead to throughput requirements which are not consistent with either the P (logistic) or Gross Margin (economic) optimum. E.g., an high repair demand

would be good in case of ratio lower than 7,5 % ÷ 10 %: it favours the economic optimisation. In other cases, it is not good. A reconciliation with the strategic level should then be planned. In the case study, 2 situations are now under consideration to reconcile: (i) to outsource the excess of repair demand, so to make only the optimum throughput; (ii) to re-negotiate contracts, to reduce the penalty cost (lower than 7,5 % ÷ 10 %).

5   Conclusions

The paper showed the possibility to adopt P as synthetic measure to optimise the trade off of flow time and throughput. Its optimisation, however, considers only a logistic perspective, economic optimisation may be integrated to consider the cost effects of the trade off. Thereafter, a practice of integrated use of P and cost criteria for decision making should be envisioned. Next research steps will aim to this end: (i) by extending the experimentation in case studies with different strategic and cost settings and (ii) by building, from their empirical evidence, a much more general framework of the practice of integrated use of P and cost criteria.

References

1. De Ron, A. J. and Rooda, J. E. Equipment Effectiveness: OEE revisited. IEEE Transaction on Semiconductor Manufacturing, vol. 18, no. 1, 190-196, 2005.

2. De Ron, A. J. and Rooda, J. E. Fab Performance. IEEE Transaction on Semiconductor Manufacturing, vol. 18, no. 3, 399-405, 2005.

3. Hopp, W. and Spearman, M.L. Factory physics: foundation of manufacturing management. McGraw-Hill, Boston, 2000.

4. Huang, S. H., Dismukes, J. P., Shi, J., Su, Q., Razzak, M. A., Bodhale R. and Robinson, D.E. Manufacturing productivity improvement using effectiveness metrics and simulation analysis. International Journal of Production Research, vol.41, no.3, 513–527, 2003.

5. Nakajima, S. Introduction to TPM Total Productive Maintenance. Productivity Press, Cambridge, Massachusetts, Norwalk, Connecticut, 1988.

6. SEMI. Standard for definition and measurement of equipment productivity, SEMI E79-0200. Semiconductor Equipment and Material International, Mt. View, California, 2000.

7. SEMI. Provisional guideline for definition and calculation of overall factory efficiency (OFE) and other associated factory-level productivity metrics. Technical Report SEMI E124, International SEMATECH, Inc, 2003.