SEMI Draft Document 3454
PROVISIONAL GUIDELINE FOR DEFINITION AND CALCULATION OF OVERALL FACTORY EFFICIENCY (OFE) AND OTHER ASSOCIATED FACTORY-LEVEL PRODUCTIVITY METRICS

1 Purpose

1.1 This document describes metrics that show how well a factory is operating compared to how well it could be operating (for the given product mix). These metrics can be used for tracking factory performance (in value-added production) in a way that rewards good operational decisions and that is not easy to adversely manipulate. They can be used in a process of ongoing improvement that can be visible to all levels of a semiconductor manufacturing organization.

1.2 The metrics in this document are intended for evaluating factory productivity after a factory is in production, not for capacity analysis while the factory is being designed or redesigned. However, some of these metrics can be used in factory simulations for choosing equipment sets and scheduling policies.

2 Scope

2.1 To evaluate the overall effectiveness of a factory, there are at least three things in need of measurement: production, utilization of assets, and costs. This document focuses on evaluating production; utilization of assets and costs (as well as other economic factors) are outside its scope. For measuring utilization of assets, an average (over all of the equipment in the factory) of overall equipment efficiency (OEE) weighted by cost of ownership (COO) as defined in SEMI E79 and E35 (respectively) can be defined in one of those documents. However, utilization of consumables, utilities, and human resources would still need to be comprehended. For measuring costs, some other new standard might define a new metric (like COO) for the cost of factory ownership in such terms as \$/(good wafers), \$/(good chips), \$/(metal levels), \$/(good transistors), \$/circuit, or \$/bit.

2.2 This standard is provisional because *overall factory efficiency* (*OFE*) is a new concept. Once the metrics have been validated by collecting data in production factories and computing the metrics, this standard should be modified and upgraded from provisional status (as specified in the SEMI regulations).

2.3 This standard does not purport to address safety issues, if any, associated with its use. It is the responsibility of the users of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.

3 Limitations

3.1 These metrics are not intended for diagnosing problems (or opportunities for improvement) in the factory, although some component metrics can be used that way.

NOTE 1: These metrics should indicate whether the factory is running poorly (like taking a person's temperature tells whether they are sick) while some of its components and other diagnostic metrics might indicate what the cause of the problem is (like doing a blood analysis in the lab identifies the disease). The following are examples of metrics not in this standard that can be used for diagnosis:

- ratio of turns to work in process (WIP) at key operations or for blocks of operations.
- overall WIP distribution.
- daily starts and output.
- defect density.
- throughput, utilization, and available up-time of bottleneck equipment.

3.2 This standard provides metrics and calculations for measuring the overall productivity only of manufacturing environments (such as wafer fabs, flat panel factories, and some disk-drive production facilities) in which product substrates move through the factory with no assembly or disassembly processes. However, in the future this document may be extended to comprehend other, more complex manufacturing environments.

3.3 In the context of this document, it is important to note that factory-level productivity is impacted greatly by factors far beyond the factory itself, including material availability, device designs, and customer demand.

4 Referenced Standards

4.1 SEMI Standards

SEMI E10 — Specification for Definition and Measurement of Equipment Reliability, Availability, and Maintainability (RAM)

SEMI E35 — Cost of Ownership for Semiconductor Manufacturing Equipment Metrics

SEMI E79 — Standard for Definition and Measurement of Equipment Productivity

NOTE 2: As listed or revised, all documents cited shall be the latest publications of adopted standards.

5 Terminology

NOTE 3: Definitions for calculated metrics can be found in Section 6.

5.1 *availability efficiency* (time divided by time) — the fraction of *total time* that the equipment is in a condition to perform its intended function (SEMI E79).

5.2 *average cycle time* — the average (over all of the units of production in *finished units out*) of *cycle time*.

5.3 *cycle time* — the amount of time a unit of production spends as WIP in the factory.

5.4 *finished units out* — the number of units of production that finish processing and testing during the period being measured.

5.5 *operational efficiency* (time divided by time) — the fraction of equipment *uptime* that the equipment is processing actual units (SEMI E79).

5.6 *overall equipment efficiency* (*OEE*) (time divided by time) — a metric of equipment performance, expressing the theoretical production time for the effective unit output divided by the *total time* (SEMI E79).

5.7 *quality efficiency* (time divided by time) — the theoretical production time for effective units divided by the theoretical production time for actual units (SEMI E79).

5.8 *set of bottleneck equipment* (F_{e*}) — the collection of production equipment of the type that has the highest average *operational efficiency* in the factory during the period being measured. Elements of this set are indicated by "f", and the equipment type is indicated by "$e*$".

NOTE 4: This set of bottleneck equipment might not be the equipment set (often the expensive lithography exposure equipment) that was planned to be the bottleneck in the factory, but rather the equipment set with the highest average available utilization during the period being measured. If another equipment set experiences significantly lower availability than expected, it might become the bottleneck.

5.9 *set of equipment types* (E) — the collection of the different types of production equipment in the factory, including metrology equipment and material handling vehicles and conveyors. Elements of this set are indicated by "e".

NOTE 5: If units are transported manually between process process steps, then the human transporters (and any carts or mechanized vehicles that they operate to perform the movement) should be considered a type of equipment for the purpose of computing the metrics in this document.

5.10 *set of process steps of product type p on equipment type e* (S_{pe}) — the collection of the different process steps (including metrology inspection and material handling transport) planned for a unit of production of product type p on equipment of type e in the factory. Elements of this set are indicated by "s".

5.11 *set of product types* (P) — the collection of the different types of products manufactured in the factory. Elements of this set are indicated by "p".

5.12 *set of equipment of type e* (F_e) — the collection of production equipment of type $e \in E$ in the factory. Elements of this set are indicated by "f".

5.13 *theoretical production time per unit* (*THT*) (time per unit) — for a given production recipe performed by a given processing module, the minimum time to complete processing on one unit of production assuming no efficiency losses are present. The determination of *theoretical production time per unit* is based on continuous operation of the processing module, where the module is assumed to operate in an ideal condition. For equipment cycles that simultaneously process more than one unit, *theoretical production time per unit* is the minimum time to perform the cycle on an equipment load whose size is optimized for throughput divided by the number of units in that optimized load (SEMI E79).

5.14 *theoretical unit throughput by recipe* (units per time) — for a given production recipe, the number of units per period of time that theoretically could be processed by the equipment. For each recipe, theoretical unit throughput is equal to the reciprocal of theoretical production time per unit (SEMI E79).

5.15 *throughput rate* — the number of units of production that pass through a process per period of time.

5.16 *total time* — all time (at the rate of 24 hours per day, seven days per week) during the period being measured. In order to have a valid representation of *total time*, all six basic equipment states must be accounted for and tracked accurately (SEMI E10). For factory-level productivity metrics, *total time* must be larger than the *average cycle time* (and is recommended to be significantly larger).

5.17 *total units out* — the number of units of production (including scrap, external rework, etc.) that exit the factory during the period being measured.

5.18 *unit* (*of production*) — the basic entity in the factory (a wafer in a fab, a glass pane in a flat panel factory, and a die in a post-wafer back-end chip production facility) which acts as a product substrate (and moves through the factory with no assembly or disassembly processes). Only product units are included (as opposed to test wafers or other non-product devices).

NOTE 6: This definition is more restrictive than that given in SEMI E10 in order to be sufficiently specific.

NOTE 7: Production lot sizes can (and typically do) include multiple units of production, and units can (and typically do) contain multiple product devices.

5.19 *uptime* (*equipment uptime*) — the hours when the equipment is in a condition to perform its intended function. It includes productive, standby, and engineering time, and does not include any portion of non-scheduled time (SEMI E10).

5.20 *WIP capacity* ($W_{\max}$)— the maximum number of units of production the factory can contain (including on shelves, in stockers, on material handling transport vehicles, on equipment load ports, in internal carrier buffers, and in process chambers).

NOTE 8: This is not a practical WIP level, because it represents total gridlock of the factory.

5.21 *work in process* (*WIP*) — the number of units of production that have been released into the factory but have not yet been scrapped, sent out for external rework, or finished processing through all of their production steps.

6 Calculated Metrics

6.1 *Main Productivity Metric* — Figure 1 shows how terminology definitions from Section 5 (shown in the top row) feed into the calculated metrics in this section. Arrows go from subordinate terms to the term in which they are cited as a part of the primary definition. Colors indicate the sub-section in which the definition is given.

6.1.1 *overall factory efficiency* (*OFE*) — This metric shows how well a factory is operating compared to how well it could be operating (for the given product mix).

$$OFE = \begin{pmatrix} \mathit{production} \\ \mathit{efficiency} \end{pmatrix} \times \begin{pmatrix} \mathit{yield} \\ \mathit{efficiency} \end{pmatrix}$$

NOTE 9: Unlike OEE, OFE and its factors are not dimensioned in time divided by time, because not all equipment in the factories is present or operating for the same amount of time.

NOTE 10: Similar to OEE, this metric:

- is dependent on product mix, process flow, operations, and time period (so users should be careful when comparing different factories or even comparing different time periods in the same factory when the product mix or process has changed).
- does not comprehend down-stream demand or the varying importance of different products (which might be addressed by a separate metric).
- varies between zero (total chaos or gridlock) and one (unobtainable perfection).
- is a product of independent dimensionless efficiencies.

6.2 *Quality Metrics* — These metrics show the efficiency of the process with respect to use of materials.

6.2.1 *yield efficiency* — This metric shows the overall material efficiency and is again a dimensionless product of independent efficiencies.

$$\begin{pmatrix} \mathit{yield} \\ \mathit{efficiency} \end{pmatrix} = \begin{pmatrix} \mathit{line} \\ \mathit{yield} \end{pmatrix} \times \begin{pmatrix} \mathit{test} \\ \mathit{yield} \end{pmatrix} = \frac{\mathit{equivalent\ good\ units\ out}}{\mathit{total\ units\ out}}$$

NOTE 11: This metric is similar to but not the same as the *quality efficiency* metric from SEMI E79 (the definition of which is given in Section 5).

6.2.2 *line yield* — This metric measures relative material losses before processing is finished.

$$\begin{pmatrix} \mathit{line} \\ \mathit{yield} \end{pmatrix} = \frac{\mathit{finished\ units\ out}}{\mathit{total\ units\ out}}$$

6.2.3 *test yield* — This metric measures relative losses due to failure at final test.

$$\begin{pmatrix} \mathit{test} \\ \mathit{yield} \end{pmatrix} = \frac{\mathit{equivalent\ good\ units\ out}}{\mathit{finished\ units\ out}}$$

6.3 *equivalent good units out* — This metric is the (possibly non-integer) number of equivalent units of production required to contain all of the good product that exits the factory during the period being measured.

$$\begin{pmatrix} \mathit{equivalent} \\ \mathit{good} \\ \mathit{units\ out} \end{pmatrix} = \sum_{p \in P} \frac{\begin{pmatrix} \text{total number of good} \\ \text{product devices of type } p \\ \text{in } \mathit{finished\ units\ out} \end{pmatrix}}{\begin{pmatrix} \text{number of product devices} \\ \text{on each unit of type } p \end{pmatrix}}$$

6.4 *Production Metrics* — These metrics show the efficiency of the process with respect to factory dynamics (without the effects of yield and scrap losses). For an exposition of the underlying science behind these metrics, see Appendix 1.

6.4.1 *production efficiency* — This metric is a normalized measure of the efficiency of the process with respect to factory dynamics.

$$\begin{pmatrix} \mathit{production} \\ \mathit{efficiency} \end{pmatrix} = \begin{pmatrix} \mathit{throughput\text{-}rate\ and} \\ \mathit{cycle\text{-}time\ efficiency} \end{pmatrix}^{\begin{pmatrix} \mathit{normalizing} \\ \mathit{exponent} \end{pmatrix}}$$

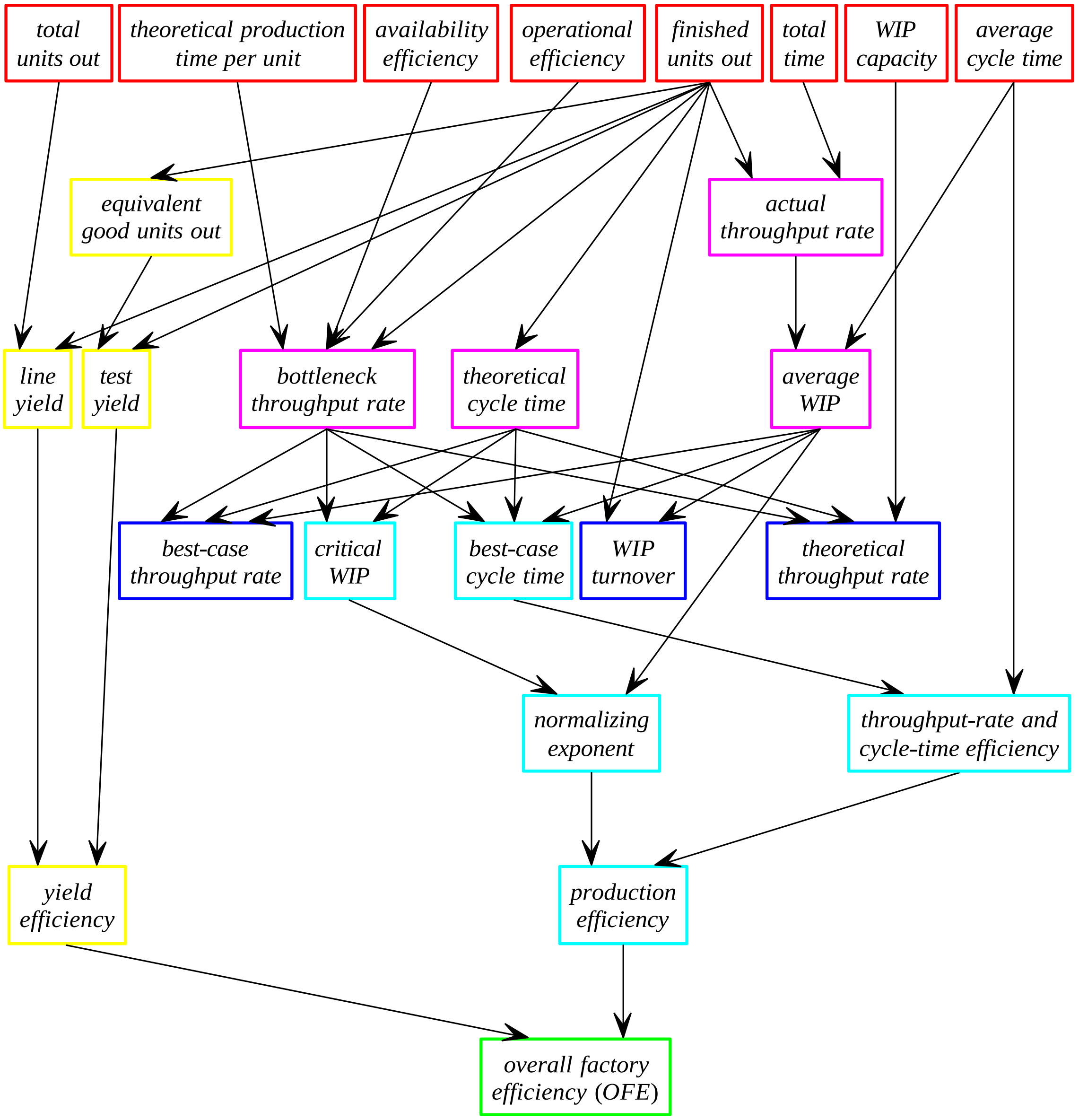

Figure 1
Definition Tree for Calculated Metrics

6.4.2 *normalizing exponent* — This power normalizes the *throughput-rate and cycle-time efficiency* so that a value of ½ for *production efficiency* indicates that the factory is performing at the threshold level (which divides a well run factory from one badly operated). This threshold level is also known as the practical worst case, because it represents what the best operating procedures can do in a maximally random factory. See the *Factory Physics* book[1] for more on this case.

$$\binom{\textit{normalizing}}{\textit{exponent}} = \frac{1}{\log_2\left[\dfrac{\binom{\textit{average}}{\textit{WIP}} + \binom{\textit{critical}}{\textit{WIP}} - 1}{\max\left\{\binom{\textit{average}}{\textit{WIP}}, \binom{\textit{critical}}{\textit{WIP}}\right\}}\right]}$$

NOTE 12: For any $x>0$, $\log_2(x) = \log_{10}(x)/\log_{10}(2) = \ln(x)/\ln(2)$.

---

[1] *Factory Physics: Foundations of Manufacturing Management*; 2nd edition (April 4, 2000), by Wallace J. Hopp and Mark L. Spearman, Irwin/McGraw Hill, New York, ISBN: 0256247951.

6.4.3 *throughput-rate and cycle-time efficiency* — This metric shows the relative performance of the factory with respect to throughput rate and cycle time.

$$\left(\begin{array}{c}\textit{throughput-rate and}\\ \textit{cycle-time efficiency}\end{array}\right)=\frac{\textit{best-case cycle time}}{\textit{average cycle time}}$$

The efficiency of both throughput rate and cycle time can be measured by the same metric, as shown in the following derivation. This derivation also gives alternative definitions for *throughput-rate and cycle-time efficiency* for use when cycle time information is not available (such as in resource-based simulations).

$$\begin{aligned}
&\left(\begin{array}{c}\textit{throughput-rate and}\\ \textit{cycle-time efficiency}\end{array}\right)\\
&=\frac{\textit{best-case cycle time}}{\textit{average cycle time}}\\
&=\frac{\max\left\{\left(\begin{array}{c}\textit{theoretical}\\ \textit{cycle time}\end{array}\right),\dfrac{\textit{average WIP}}{\textit{bottleneck throughput rate}}\right\}}{\textit{average cycle time}}\\
&=\frac{\max\left\{\left(\begin{array}{c}\textit{theoretical}\\ \textit{cycle time}\end{array}\right),\dfrac{\left(\begin{array}{c}\textit{average}\\ \textit{WIP}\end{array}\right)}{\left(\begin{array}{c}\textit{bottleneck}\\ \textit{throughput rate}\end{array}\right)}\right\}\div\left(\begin{array}{c}\textit{average}\\ \textit{WIP}\end{array}\right)}{\left(\begin{array}{c}\textit{average}\\ \textit{cycle time}\end{array}\right)\div\left(\begin{array}{c}\textit{average}\\ \textit{WIP}\end{array}\right)}\\
&=\frac{\max\left\{\dfrac{\textit{theoretical cycle time}}{\textit{average WIP}},\dfrac{1}{\left(\begin{array}{c}\textit{bottleneck}\\ \textit{throughput rate}\end{array}\right)}\right\}}{1\div(\textit{actual throughput rate})}\\
&=\frac{\textit{actual throughput rate}}{\min\left\{\dfrac{\textit{average WIP}}{\textit{theoretical cycle time}},\left(\begin{array}{c}\textit{bottleneck}\\ \textit{throughput rate}\end{array}\right)\right\}}\\
&=\frac{\textit{actual throughput rate}}{\textit{best-case throughput rate}}\\
&=\frac{(\textit{finished units out})\div(\textit{total time})}{(\textit{average WIP})\div(\textit{theoretical cycle time})}\\
&=\frac{\textit{theoretical cycle time}}{\textit{total time}}\times\frac{\textit{finished units out}}{\textit{average WIP}}\\
&=\left(\begin{array}{c}\textit{theoretical cycle time}\\ \text{as a fraction of }\textit{total time}\end{array}\right)\times\left(\begin{array}{c}\textit{WIP}\\ \textit{turnover}\end{array}\right)
\end{aligned}$$

6.4.4 *critical WIP* (W_0) — This metric gives the WIP level that (theoretically) allows the factory to have the highest throughput rate with the shortest cycle time.

$$\left(\begin{array}{c}\textit{critical}\\ \textit{WIP}\end{array}\right)=\left(\begin{array}{c}\textit{theoretical}\\ \textit{cycle time}\end{array}\right)\times\left(\begin{array}{c}\textit{bottleneck}\\ \textit{throughput rate}\end{array}\right)$$

6.4.5 *best-case cycle time* — This metric shows the best cycle time that the factory can do given the WIP loading.

$$\left(\begin{array}{c}\textit{best-case}\\ \textit{cycle time}\end{array}\right)=\max\left\{\left(\begin{array}{c}\textit{theoretical}\\ \textit{cycle time}\end{array}\right),\frac{\left(\begin{array}{c}\textit{average}\\ \textit{WIP}\end{array}\right)}{\left(\begin{array}{c}\textit{bottleneck}\\ \textit{throughput rate}\end{array}\right)}\right\}$$

6.5 *Auxiliary Metrics* — These metrics give some additional measures that might be useful in managing a factory.

6.5.1 *best-case throughput rate* — This metric shows the best throughput rate that the factory can do given the WIP loading.

$$\left(\begin{array}{c}\textit{best-case}\\ \textit{throughput}\\ \textit{rate}\end{array}\right)=\min\left\{\frac{\left(\begin{array}{c}\textit{average}\\ \textit{WIP}\end{array}\right)}{\left(\begin{array}{c}\textit{theoretical}\\ \textit{cycle time}\end{array}\right)},\left(\begin{array}{c}\textit{bottleneck}\\ \textit{throughput}\\ \textit{rate}\end{array}\right)\right\}$$

6.5.2 *theoretical throughput rate* — This metric gives an (unobtainable) upper bound on the factory throughput rate.

$$\left(\begin{array}{c}\textit{theoretical}\\ \textit{throughput}\\ \textit{rate}\end{array}\right)=\min\left\{\frac{\left(\begin{array}{c}\textit{WIP}\\ \textit{capacity}\end{array}\right)}{\left(\begin{array}{c}\textit{theoretical}\\ \textit{cycle time}\end{array}\right)},\left(\begin{array}{c}\textit{bottleneck}\\ \textit{throughput}\\ \textit{rate}\end{array}\right)\right\}$$

6.5.3 *WIP turnover* — This metric shows how often the inventory of work in process was replaced during the period being measured.

$$\left(\begin{array}{c}\textit{WIP}\\ \textit{turnover}\end{array}\right)=\frac{\textit{finished units out}}{\textit{average WIP}}$$

6.6 *Fundamental Quantities* — These metrics (along with some definitions in Section 5) give the basic building blocks for constructing previous metrics.

6.6.1 *average WIP* — This metric shows how many units of production fill the "pipeline" on average.

$$\left(\begin{array}{c}\textit{average}\\ \textit{WIP}\end{array}\right)=\left(\begin{array}{c}\textit{average}\\ \textit{cycle time}\end{array}\right)\times\left(\begin{array}{c}\textit{actual}\\ \textit{throughput rate}\end{array}\right)$$

6.6.2 *actual throughput rate* — This metric shows how fast finished wafers flow out of the factory.

$$\left(\begin{array}{c}\textit{actual}\\ \textit{throughput rate}\end{array}\right)=\frac{\textit{finished units out}}{\textit{total time}}$$

6.6.3 *theoretical cycle time* (T_{min}) — This metric is the minimum time required to process a unit of production through the factory (including material handling transport time) if the unit never has to wait for equipment or a vehicle to become available. This is also known as the raw process time. If a process change for a product causes this metric to change, it should be considered a different product for the purposes of performing these computations. If more than one product (or process flow) is represented in the output, an average is taken over each of the products' *theoretical cycle time* weighted by the fraction of that product found in *finished units out* (as shown in the following computation):

$$\begin{pmatrix}\textit{theoretical}\\ \textit{cycle time}\end{pmatrix} = \frac{\displaystyle\sum_{p\in P}\left\{\begin{pmatrix}\text{number}\\ \text{of units}\\ \text{of product}\\ \text{type } p \text{ in}\\ \textit{finished}\\ \textit{units out}\end{pmatrix}\times\sum_{e\in E}\left[\sum_{s\in S_{pe}}\begin{pmatrix}\text{minimum}\\ \text{cycle}\\ \text{time of}\\ \text{a single}\\ \text{unit of}\\ \text{product}\\ \text{type } p \text{ in}\\ \text{step } s \text{ on}\\ \text{equipment}\\ \text{type } e\end{pmatrix}\right]\right\}}{\textit{finished units out}}$$

NOTE 13: This metric is similar to but not the same as the *theoretical production time per unit* (*THT*) metric from SEMI E79 (the definition of which is given in Section 5).

6.6.4 *bottleneck throughput rate* (R_{max}) — This metric gives the upper bound on the factory throughput rate imposed by the current bottleneck equipment set. If a process change for a product causes this metric to change, it should be considered a different product for the purposes of performing these computations.

$$\begin{pmatrix}\textit{bottleneck}\\ \textit{throughput}\\ \textit{rate}\end{pmatrix} = \frac{\begin{pmatrix}\textit{finished}\\ \textit{units}\\ \textit{out}\end{pmatrix}\times\begin{pmatrix}\text{average}\\ \text{number of}\\ \text{tools in}\\ \text{bottleneck}\\ \text{equipment}\\ \text{type } e^*\end{pmatrix}\times\begin{pmatrix}\text{average}\\ \textit{availability}\\ \textit{efficiency}\\ \text{of bottleneck}\\ \text{equipment}\\ \text{type } e^*\end{pmatrix}}{\displaystyle\sum_{p\in P}\left[\begin{pmatrix}\text{number of}\\ \text{units of}\\ \text{product}\\ \text{type } p \text{ in}\\ \textit{finished}\\ \textit{units out}\end{pmatrix}\times\sum_{s\in S_{pe^*}}\begin{pmatrix}\textit{theoretical}\\ \textit{production}\\ \textit{time per unit}\\ \text{for product type}\\ p \text{ in step } s \text{ on}\\ \text{bottleneck}\\ \text{equipment}\\ \text{type } e^*\end{pmatrix}\right]}$$

NOTE 14: This metric is similar to but not the same as the *theoretical unit throughput by recipe* metric from SEMI E79 (the definition of which is given in Section 5).

NOTE 15: This set of bottleneck equipment might not be the equipment set (often the expensive lithography exposure equipment) that was planned to be the bottleneck in the factory, but rather the equipment set with the highest average *operational efficiency* (the available utilization) during the period being measured. If another equipment set experiences significantly lower availability than expected, it might become the bottleneck.

NOTE 16: One of the factors in this metric is the average number of available tools in the current bottleneck equipment set, not the total number of tools in the set.

NOTE 17: This metric is not an average over the bottleneck throughput rates of each product.

7 Related Documents

7.1 SEMI Standards

SEMI E58 — Automated Reliability, Availability, and Maintainability Standard (ARAMS): Concepts, Behavior, and Services

NOTICE: SEMI makes no warranties or representations as to the suitability of the standard set forth herein for any particular application. The determination of the suitability of the standard is solely the responsibility of the user. Users are cautioned to refer to manufacturer's instructions, product labels, product data sheets, and other relevant literature respecting any materials mentioned herein. These standards are subject to change without notice.

The user's attention is called to the possibility that compliance with this standard may require use of copyrighted material or of an invention covered by patent rights. By publication of this standard, SEMI takes no position respecting the validity of any patent rights or copyrights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of any such patent rights or copyrights, and the risk of infringement of such rights, are entirely their own responsibility.

APPENDIX 1
MANUFACTURING SCIENCE BACKGROUND

NOTE: The material in this appendix is an official part of this document.

A1-1 To understand the Production Metrics given in Section 6.3, we need to understand the underlying science behind factory dynamics. The most important concept is Little's Law given in the following equation:

$$\begin{pmatrix} \mathit{average} \\ \mathit{WIP} \\ (\mathit{units}) \end{pmatrix} = \begin{pmatrix} \mathit{actual} \\ \mathit{throughput\ rate} \\ (\mathit{units}/\mathit{time}) \end{pmatrix} \times \begin{pmatrix} \mathit{average} \\ \mathit{cycle\ time} \\ (\mathit{time}) \end{pmatrix}$$

A1-2 *Factory Physics* calls this identity the "*F = ma*" of manufacturing science. Little's Law relates the three most significant fundamental quantities of production systems. Unfortunately, it says that all three metrics cannot be optimized simultaneously. Little's Law is shown graphically in Figure A1-1. In all of the figures in this Appendix, the color bands denote different values of the *production efficiency* metric with red representing values close to zero, yellow representing values close to ½ (the threshold case), and green representing values close to one. Note that the boundaries of the operating region are determined by the *theoretical cycle time* ($T_{\min}$), the *bottleneck throughput rate* ($R_{\max}$), and the *WIP capacity* ($W_{\max}$).

A1-3 If we look at two of these fundamental quantities at a time, the factory dynamics become more clear. For example, Figure A1-2 on the next page shows *actual throughput rate* as a function of *average WIP* levels. Here the diagonal solid black lines represent different constant cycle times, but no known strategy will keep the factory operating exactly on one of these lines.

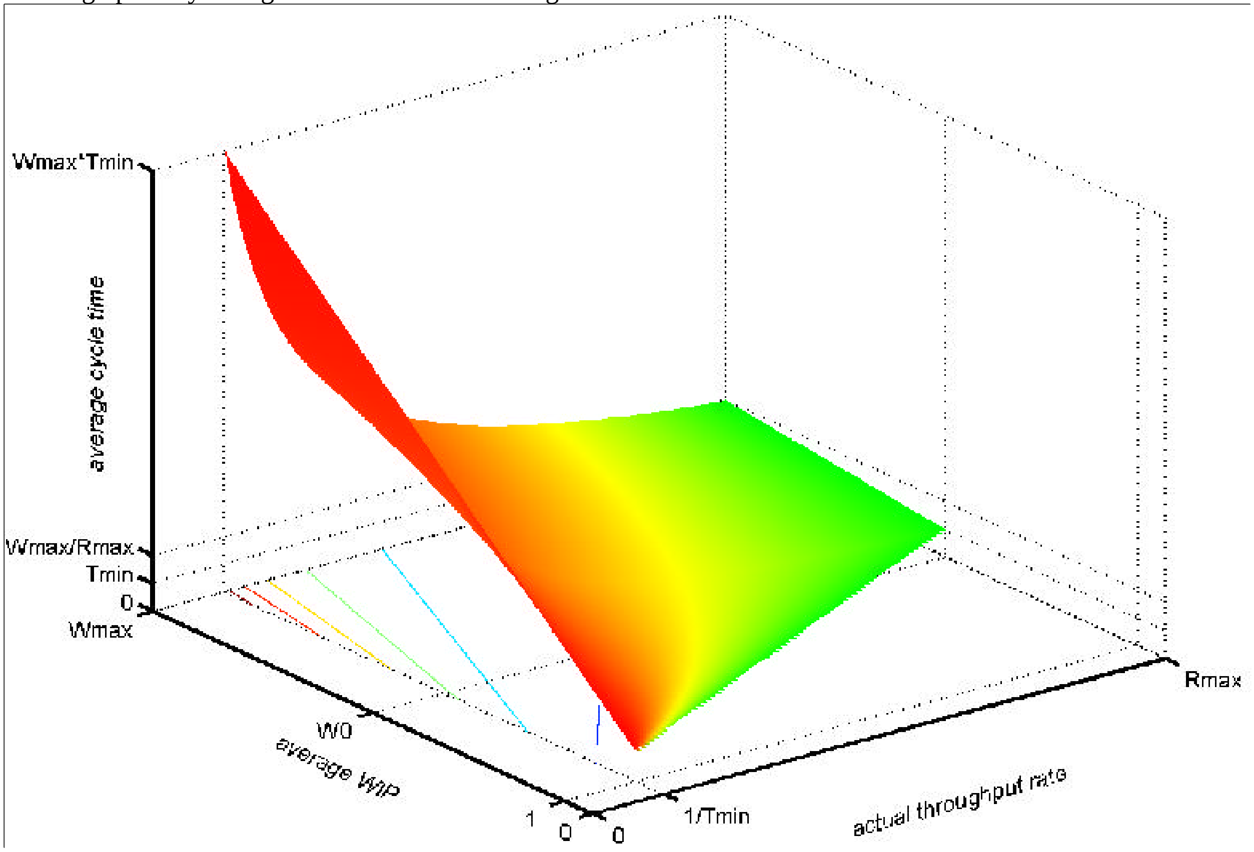

Figure A1-1
Surface Plot of Little's Law

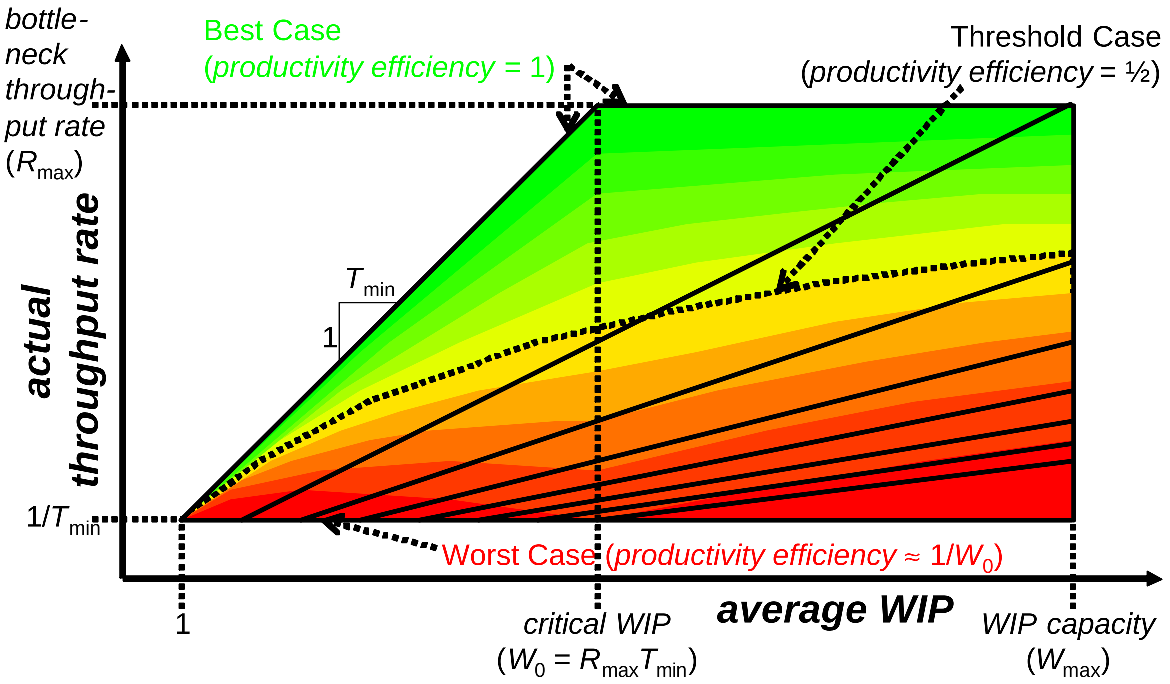

Figure A1-2
Plot of *actual throughput rate* vs. *average WIP*

A1-4 Now suppose the factory is managed with a push strategy where a constant throughput rate is enforced so that WIP levels are allowed to reach their equilibrium state. As shown below in Figure A1-3, this amounts to choosing to operate the factory on one of the diagonal solid black lines (each of which represent different constant throughput rates). We then try to drive the factory along the line toward the bottom left (for lower *average cycle time* and *average WIP* levels) by using better operating principles, but we are resisted by the inherent variability of the factory.

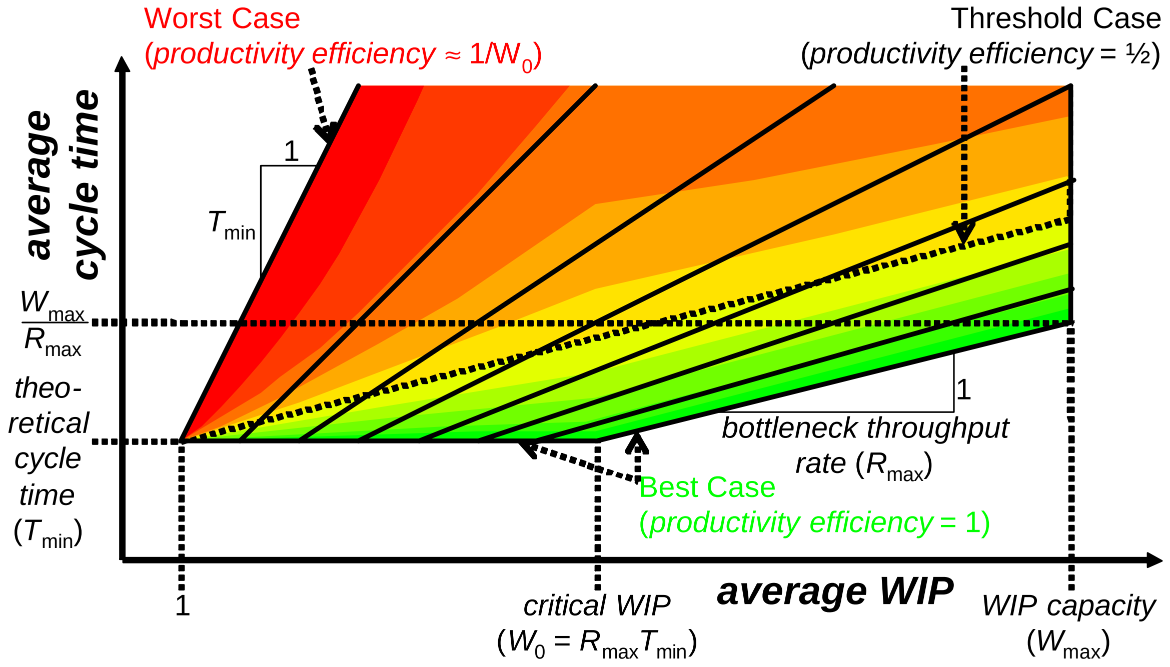

Figure A1-3
Plot of *average cycle time* vs. *average WIP*

A1-5 Now suppose the factory is managed with a pull strategy where a constant WIP level is enforced so that throughput rates are allowed to reach their equilibrium state. In general, this is a better strategy, because studies have shown that a constant WIP level will result in a higher average throughput rate than the constant throughput rate that results in the same average WIP level. As shown below in Figure A1-4, this constant WIP strategy (known as CONWIP) amounts to choosing to operate the factory on one of the solid black curves (each of which represents a different constant WIP level). We then try to drive the factory along the curve toward the bottom right (for lower *average cycle time* and a higher *actual throughput rate*) by using better operating principles, but we are resisted by the inherent variability of the factory.

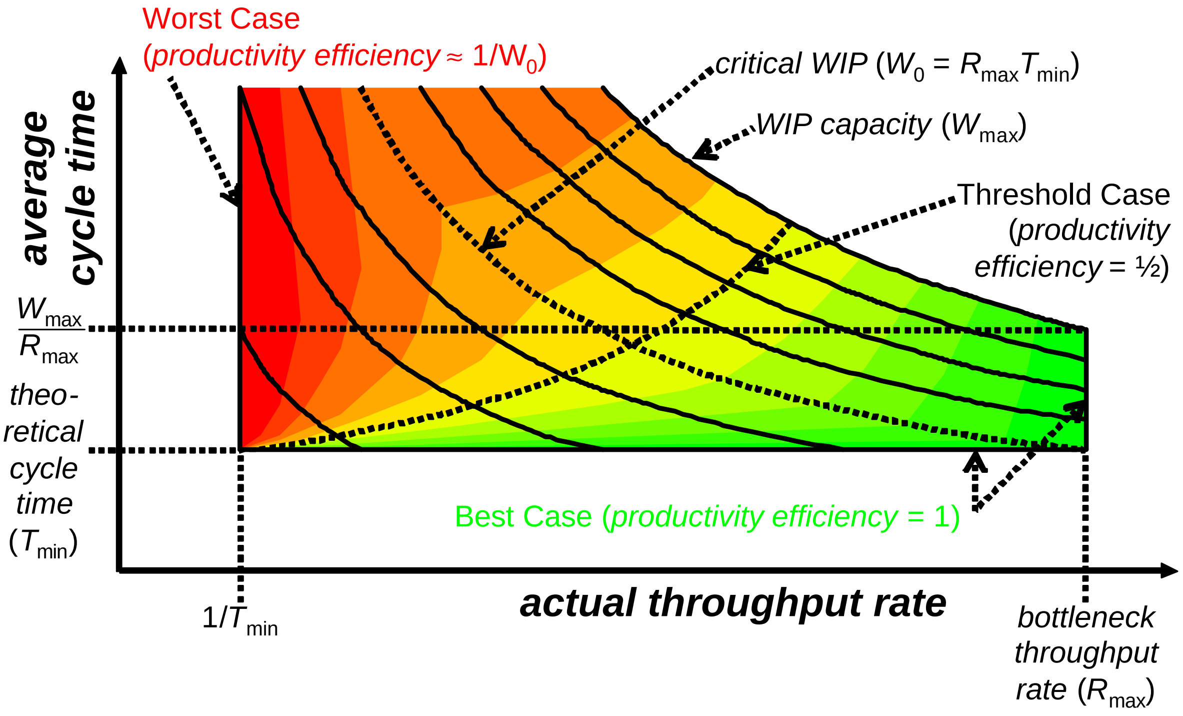

Figure A1-4
Plot of *average cycle time* vs. *actual throughput rate*

APPENDIX 2
EXAMPLE APPLICATION

NOTE: The material in this appendix is an official part of this document.

A2-1 As an example to show how the metrics in this document are applied, the diagram in Figure 6 shows the process flow for a grossly simplified model of a wafer fab (with only five machines in three tool sets and one process flow with six steps) that was developed by Karl Kempf at Intel.

A2-2 The data for this model are shown in the first few columns of Table A2-1 (where all times are in minutes) and Table A2-2. A simulation gave the following additional run data:

total time = 9 years = 4,733,640 minutes

average cycle time = 1.8 days = 2592 minutes

total units out = 39420 lots = 985500 wafers

finished units out = 938571 wafers

equivalent good units out = 891642.1 wafers

A2-3 In the next-to-last column of Table A2-1, we added together (for each step) all of the times (to load, process batch, unload, and travel to next step) and summed the results to get a *theoretical cycle time* of 812.4 minutes. In the last column of Table A2-1, we computed the *theoretical production time per unit* for each step. These values were then used in the last column of Table A2-2 to compute the *throughput rate* for each equipment set. The *throughput rate* for the Lithography equipment set (0.218182 wafers/minute) is the *bottleneck throughput rate*, not because it is the smallest *throughput rate* (it is), but because the Lithography equipment set has the highest average *operational efficiency*.

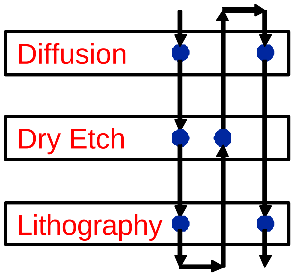

Figure A2-1
MiniFab Process Flow

A2-4 The remaining Terminology is given by:

$$\begin{aligned}\begin{pmatrix} WIP \\ capacity \end{pmatrix} &= \sum_{e \in E} \left[\begin{pmatrix} Buffer \\ Size \end{pmatrix} + \begin{pmatrix} Process \\ Batch\ Size \end{pmatrix} \times \begin{pmatrix} Size \\ of\ \ Set \end{pmatrix} \right] \\ &= [450 + 75 \times 2] + [300 + 25 \times 2] + [300 + 1 \times 1] \\ &= [450 + 150] + [300 + 50] + [300 + 1] \\ &= 600 + 350 + 301 \\ &= 1251 \text{ wafers} \end{aligned}$$

A2-5 The remaining Calculated Metrics are computed on the following page.

Table A2-1 Process Data

Process Step Number	*Equipment Set Name*	*Time to Load*	*Time to Process Batch*	*Time to Unload*	*Time to Travel to Next Step*	*cycle time = (sum of 4 Time columns at left)*	*theoretical production time/unit = (Time to Process Batch) ÷ (Process Batch Size)*
1	Diffusion	20	225	40	8	293 minutes	3.0 minutes/wafer
2	Dry Etch	15	30	15	4	64 minutes	1.2 minutes/wafer
3	Lithography	10	2.2	10	4	26.2 minutes	2.2 minutes/wafer
4	Dry Etch	15	50	15	8	88 minutes	2.0 minutes/wafer
5	Diffusion	20	255	40	4	319 minutes	3.4 minutes/wafer
6	Lithography	10	2.2	10	-	22.2 minutes	2.2 minutes/wafer
Sum	-	90	564.4	130	28	812.4 minutes	-

Table A2-2 Equipment Data

Equipment Set Name	*Size of Set*	*Process Batch Size*	*Buffer Size (wafers)*	*Average operational efficiency*	*availability efficiency (each tool)*	*Average availability efficiency*	*throughput rate = (Size of Set) × (Ave. availability efficiency) ÷ S(theoretical prod. time/unit)*
Diffusion	2	75	450	88%	93% & 97%	95%	0.296875 wafers/minute
Dry Etch	2	25	300	78%	81% & 85%	83%	0.518750 wafers/minute
Lithography	1	1	300	91%	96%	96%	0.218182 wafers/minute

$$\begin{aligned}\begin{pmatrix}\textit{actual}\\ \textit{throughput rate}\end{pmatrix} &= \frac{\textit{finished units out}}{\textit{total time}}\\ &= \frac{938{,}571 \text{ wafers}}{4{,}733{,}640 \text{ minutes}}\\ &= 0.19827\frac{\text{wafers}}{\text{minute}}\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{average}\\ \textit{WIP}\end{pmatrix} &= \begin{pmatrix}\textit{average}\\ \textit{cycle time}\end{pmatrix}\times\begin{pmatrix}\textit{actual}\\ \textit{throughput rate}\end{pmatrix}\\ &= (2592 \text{ minutes})\times\left(0.19827\ \frac{\text{wafers}}{\text{minute}}\right)\\ &= 513.9 \text{ wafers}\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{WIP}\\ \textit{turnover}\end{pmatrix} &= \frac{\textit{finished units out}}{\textit{average WIP}}\\ &= \frac{938{,}571 \text{ wafers}}{513.9 \text{ wafers}} = 1826.25 \text{ turns}\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{theoretical}\\ \textit{throughput}\\ \textit{rate}\end{pmatrix} &= \min\left\{\frac{\begin{pmatrix}\textit{WIP}\\ \textit{capacity}\end{pmatrix}}{\begin{pmatrix}\textit{theoretical}\\ \textit{cycle time}\end{pmatrix}}, \begin{pmatrix}\textit{bottleneck}\\ \textit{throughput}\\ \textit{rate}\end{pmatrix}\right\}\\ &= \min\left\{\frac{1251 \text{ wafers}}{812.4 \text{ minutes}}, 0.21818\frac{\text{wafers}}{\text{minute}}\right\}\\ &= \min\{1.53988, 0.21818\}\frac{\text{wafers}}{\text{minute}}\\ &= 0.21818\ \frac{\text{wafers}}{\text{minute}}\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{best-case}\\ \textit{throughput}\\ \textit{rate}\end{pmatrix} &= \min\left\{\frac{\begin{pmatrix}\textit{average}\\ \textit{WIP}\end{pmatrix}}{\begin{pmatrix}\textit{theoretical}\\ \textit{cycle time}\end{pmatrix}}, \begin{pmatrix}\textit{bottleneck}\\ \textit{throughput}\\ \textit{rate}\end{pmatrix}\right\}\\ &= \min\left\{\frac{513.9 \text{ wafers}}{812.4 \text{ minutes}}, 0.21818\frac{\text{wafers}}{\text{minute}}\right\}\\ &= \min\{0.63261, 0.21818\}\frac{\text{wafers}}{\text{minute}}\\ &= 0.21818\ \frac{\text{wafers}}{\text{minute}}\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{best-case}\\ \textit{cycle time}\end{pmatrix} &= \max\left\{\begin{pmatrix}\textit{theoretical}\\ \textit{cycle time}\end{pmatrix}, \frac{\begin{pmatrix}\textit{average}\\ \textit{WIP}\end{pmatrix}}{\begin{pmatrix}\textit{bottleneck}\\ \textit{throughput rate}\end{pmatrix}}\right\}\\ &= \max\left\{812.4 \text{ minutes}, \frac{513.9 \text{ wafers}}{0.21818\frac{\text{wafers}}{\text{minute}}}\right\}\\ &= \max\{812.4 \text{ minutes}, 2355.5 \text{ minutes}\}\\ &= 2355.5 \text{ minutes}\end{aligned}$$

NOTE A2-1: For this *average WIP* level, the *best-case cycle time* was not determined by the *theoretical cycle time*, but by the *bottleneck throughput rate*. Thus, a simple metric like *average cycle time* ÷ *theoretical cycle time* underestimates how well the factory is doing.

$$\begin{aligned}\begin{pmatrix}\textit{throughput-rate and}\\ \textit{cycle-time efficiency}\end{pmatrix} &= \frac{\textit{best-case cycle time}}{\textit{average cycle time}}\\ &= \frac{2355.5 \text{ minutes}}{2592.0 \text{ minutes}} = 90.87\%\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{critical}\\ \textit{WIP}\end{pmatrix} &= \begin{pmatrix}\textit{theoretical}\\ \textit{cycle time}\end{pmatrix}\times\begin{pmatrix}\textit{bottleneck}\\ \textit{throughput rate}\end{pmatrix}\\ &= (812.4 \text{ minutes})\times\left(0.218182\ \frac{\text{wafers}}{\text{minute}}\right)\\ &= 177.25 \text{ wafers}\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{normalizing}\\ \textit{exponent}\end{pmatrix} &= \frac{1}{\log_2\left[\dfrac{\begin{pmatrix}\textit{average}\\ \textit{WIP}\end{pmatrix}+\begin{pmatrix}\textit{critical}\\ \textit{WIP}\end{pmatrix}-1}{\max\left\{\begin{pmatrix}\textit{average}\\ \textit{WIP}\end{pmatrix},\begin{pmatrix}\textit{critical}\\ \textit{WIP}\end{pmatrix}\right\}}\right]}\\ &= \frac{1}{\log_2\left[\dfrac{513.93+177.25-1 \text{ wafers}}{\max\{513.93,\ 177.25\}\text{wafers}}\right]}\\ &= \frac{1}{\log_2\left[\dfrac{690.18 \text{ wafers}}{513.93 \text{ wafers}}\right]}\\ &= \frac{1}{\log_2[1.34294]} = \frac{1}{0.4254} = 2.35072\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{production}\\ \textit{efficiency}\end{pmatrix} &= \begin{pmatrix}\textit{throughput-rate and}\\ \textit{cycle-time efficiency}\end{pmatrix}^{\begin{pmatrix}\textit{normalizing}\\ \textit{exponent}\end{pmatrix}}\\ &= (0.9087)^{2.35072} = 79.86\%\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{test}\\ \textit{yield}\end{pmatrix} &= \frac{\textit{equivalent good units out}}{\textit{finished units out}}\\ &= \frac{891{,}642.1 \text{ wafers}}{938{,}571.0 \text{ wafers}} = 94.00\%\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{line}\\ \textit{yield}\end{pmatrix} &= \frac{\textit{finished units out}}{\textit{total units out}}\\ &= \frac{938{,}571 \text{ wafers}}{985{,}500 \text{ wafers}} = 95.24\%\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}\textit{yield}\\ \textit{efficiency}\end{pmatrix} &= \begin{pmatrix}\textit{line}\\ \textit{yield}\end{pmatrix}\times\begin{pmatrix}\textit{test}\\ \textit{yield}\end{pmatrix}\\ &= (0.9524)\times(0.9400) = 90.48\%\end{aligned}$$

$$\begin{aligned}OFE &= \begin{pmatrix}\textit{production}\\ \textit{efficiency}\end{pmatrix}\times\begin{pmatrix}\textit{yield}\\ \textit{efficiency}\end{pmatrix}\\ &= (0.7986)\times(0.9048) = 72.26\%\end{aligned}$$