TSS: a Daily Production Target Setting System for Fabs[1]

Guan-Liang Wu, Kang Wei, Chih-Yang Tsai, Shi-Chung Chang
Control and Decision Laboratory
Department of Electrical Engineering
National Taiwan University, Taipei, Taiwan, R.O.C.
Email: scchang@ac.ee.ntu.edu.tw

Nan-Jye Wang, Rong-Long Tsai, Huei-Ping Liu
MFG-2B
Taiwan Semiconductor Manufacturing Corporation
Hsin-Chu, Taiwan, R.O.C.

Abstract

In this paper, a Target Setting System (TSS) is designed as a computer aided decision support, which includes three modules: (1) Master Production Schedule Conversion (MPSC), (2) Capacity Estimation, and (3) Target Setting Main Function (TSMF). MPSC converts wafer release and output schedule into daily demanded moves of each production stage while CE estimates tool capacity by taking auto-regression of the actual move data. TSMF then takes the outputs of MPSC and CE and WIP distribution as its inputs, and adopts PULL, PUSH and iterative proportional capacity allocation schemes to calculate the daily target of each stage. TSS has been implemented for daily application. The scheduled total moves by TSS are within 5% of what have been actually achieved in the shop floor and about 50% of scheduled stage targets are within 15% difference with actual stage moves.

I. Introduction

Functions of a decision hierarchy for target-based production flow control in a fab is depicted in Figure 1. Major issues include (1) wafer release and output scheduling, (2) daily target setting and (3) lot dispatching [2, 5]. Wafer release/output aims at controlling the wafer-in-process (WIP) level and cycle time while meeting delivery requirements. It schedules the quantities of wafer release and output using a day as a time unit over two to four months. Lot dispatching determines which lot to process next when a tool is available. Daily target setting bridges between these two functions. It combines daily release/output schedule, present fab production states, and tool capacity forecast to properly set daily production targets at each production stage, which then serves as a guideline for dispatching.

Operational complexity of a foundry fab is rooted in its high-variety, low-volume production and very complicated fabrication processes. There are typically multiple operation objectives for setting daily targets:

(o1)   to meet target monthly output volume,

(o2)   to balance the production line,

(o3)   to reduce WIP and cycle time,

(o4)   to maximize tool utilization, and

(o5)   to facilitate on-time delivery of products.

In addition, there are often time varying operation constraints and manually maintained data, which may not be completely captured by the manufacturing execution system (MES) due to very dynamic and complex operations. Daily target setting is therefore quite challenging and a computer-aided decision support tool is highly desirable for daily operations.

In the past few years, Chang et al have developed a series of algorithms based on deterministic problem formulation for daily target generation and tool allocation (TGMA) of a foundry fab. The core method [1] of these algorithms adopts an iterative, proportional resource allocation approach that combines flow balance relationship, deterministic queueing analysis and some empirical rules to capture the important decision factors such as maximizing production moves, balancing the line, reducing WIP levels and prediction of detailed production flows. After the first successful field application at TSMC, the core method was extended to weekly short-term scheduling (WTGMA) [6] and to incorporating priority (hot and regular) lot scheduling (TGMA-PR) [2,3]. All these algorithms

were implemented into tools and were applied in the field. Some theoretical analysis of the core method is also available [4].

Although TGMA-based tools have achieved some successes, there are four major deficiencies that hindered the long lasting applications of these tools:

1) Exact numbers of tools at the beginning of a day are used in capacity calculation, which actually vary during the day and differ a lot from those taken for calculation;

2) Standard WIP of each stage requires empirical input and adjustment;

3) Standard processing times are used for production flow and cycle time calculation, which are difficult to be accurately maintained in a foundry fab of high product variety, low volume per product and frequent tool change;

4) Insufficient accuracy in capacity modeling of some critical tools.

Deficiencies 1) and 3) are due to unnecessary data accuracy requirements, which need not be so accurate for the purpose of daily target setting. Deficiencies 2) and 4) require a new algorithm and a detailed model respectively.

In this paper, we develop a new Target Setting System (TSS) as a computer decision-aid for setting the daily target of each production stage. TSS is designed by simplifying the data requirements and unnecessary algorithmic complications, enriching functionality and enhancing user interface of the existing TGMA-based tools. It includes three major function modules: (1) Master Production Schedule Conversion (MPSC), (2) Capacity Estimation (CE), and (3) Target Setting Main Function (TSMF). Section 2 first gives a system overview. MPSC and CE functions are then described in Section 3 and TSMF described in Section 4. Section 5 presents field test results. Finally, concluding remarks are summarized in Section 6.

II. System Overview

In addition to the five objectives (o1-o5) listed in the Introduction, TSS is also designed to

(o6) track higher level production targets;

(o7) be able to estimate wafer flow-in and forecast WIP distribution;

(o8) support dynamic calculation of standard WIP at each production stage;

(o9) minimize requirements for manually maintained input data by modification of algorithm; and

(o10) provide a friendly user interface.

Accomplishment of these objectives is essential to long lasting applicability of the tool as a fab

evolves. However, the problem is very challenging because of the dynamic nature and complexity of a foundry fab.

To reduce the problem complexity and focus on the purpose of daily target setting, we make a few assumptions:

A1. A MPS is available, which is based on reasonable estimation of mid-term tool capacity and production demands.

A2. Detailed characteristics and constraints of tool operations are neglected; empirical statistics about tool capacity are used as aggregate characterization.

A3. There exists no priority difference among prodcuts.

A4. The production unit is a wafer.

A5. The fabrication process of each part type is fixed.

A6. Stage is adopted as the basic unit for describing process flows, where a stage of a product type is obtained by aggregating a few consecutive fabrication steps of the product type.

A7. All the stages of various fabrication processes can be arranged into a global sequence of totally J stages in a way that if stage k precedes a stage k' in one process, the stage k also procedes stage k' in the global stage sequence.

A8. Each stage has a corresponding key tool group. As steps of a stage may require tools from different tool groups and different stage may also share the same amchine groups, the tool group that has the highest ratio of processing time over number of tools among the steps in a stage is selected to be the key tool of the stage.

A9. Buffer space is large and can be considered as infinite.

Figure 2 depicts the TSS system functional flow chart. The MPSC module converts wafer release and output schedule of MPS into daily demanded moves of each production stage according to the average cycle time of each stage and the process flow of each product. It also recommends the reference WIP levels for next day production. As long as all daily demanded moves are achieved every day, on-time delivery of MPS demands can be accomplished and cycle time variation controlled. The MPSC module is designed to run daily. The CE module estimates the capacity (number of wafer moves) of each tool type per day. It takes a moving average of the actual move data of individual tool over a certain period of time as the estimated capacity. And finally, TSMF takes daily stage demands from MPSC, tool capacity estimates from Capacity Estimation, and WIP distribution from the MES to calculate the daily target moves of each stage.

III. Demanded Move Calculation and Capacity Estimation

MPSC and CE modules correspond to higher levels in the decision hierarchy of production flow control. They set the environment of TSMF and significantly affect the accuracy of target setting. The two function modules are now described respectively.

III. 1 MPSC Module

Inputs of the MPSC module are
- WIP levels by stage by part type,
- actual wafer output by type up to date,
- process flow of individual part types,
- cycle times by stage by type, and
- MPS schedule (wafer start and output) for the coming three months,

which can all be obtained form the manufacturing execution system (MES) or the management information system(MIS). The outputs include
- demanded numbers of wafer moves by stage by type during the day under consideration,
- delayed numbers of wafer moves by stage by type at the beginning of the day under consideration, each of which is the differences between the demanded and the actual cumulative moves to-date for a type at a stage, and
- the reference WIP levels by stage by type for next day production.

Let us now define some notations. As MPSC is designed for individual part types, the part type index is omitted in the following notational definitions.

Notations

J: total number of stages;

j: stage index, $j=1,\ldots,J$;

T: time horizon;

t: time index, $t=1,\ldots,T$;

d: current time;

CT_j: empirical cycle time of stage j, $j=1,\ldots,J$;

$WIP_j(t)$: empirical cycle time of stage j, $j=1,\ldots,J$;

$MPS\_OUT(t)$: the delivery quantity of day t in the MPS;

$QC\_OUT(t)$: the delivery quantity of day t in the MPS;

$DD_j(t)$: demanded moves of stage j during day t ;

$Cum\_DD_j(t)$: cumulative demanded moves from day 1 to the end of day t-1 at stage j;

$MPS\_day\_demand_j(d)$: total demanded moves of current day including $DD_j(t)$ and the delayed number of moves up to the end of t-1.

The MPSC algorithm converts, for each part type, wafer release and output schedule of MPS into demanded moves of each production stage. It is described as follows.

MPSC Algorithm

Step 1: Compute, for all the MPS outputs, the available time of $MPS\_OUT(t) = t - d$, t=1,...T.

Step 2: Compute the remaining cycle time from a stage j to output backward in stage sequence,

$$RCT_J = CT_J,$$

$$RCT_j = CT_j + RCT_{j+1}, \quad \text{for } j=J-1,...,1.$$

Step 3: Compute the slack times of individual stages,

$$SL_j = t - d - RCT_{j+1}, \quad \text{for } j=1,...,J.$$

Step 4: Calculate the demands by MPS for the current day.

The demand for stage j due to $MPS\_OUT(t)$, t=1,...,T, is

$$MPS\_OUT(t) * \Pi(SL_j), \text{ where } \Pi(x) = \begin{cases} x, & 0 < x \le 1 \\ 2 - x, \text{ if } 1 < x \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Total demand by MPS for stage j is therefore

$$DD_j(d) = \sum_{t=1}^{T} MPS\_OUT(t) * \Pi(SL_j), j = 1,...,J,$$

and the cumulative demand up to the beginning of the current day is

$$Cum\_DD_j(d) = \sum_{t=1}^{d-1} DD_j(t), j = 1,...,J.$$

Step 5: Calculate the demanded (to be fulfilled) moves by MPS for the current day.

The cumulative moves of stage j that have been accomplished up to the beginning of the current day can be calculated by

$$Actual\_move_j(d) = \sum_{j'=j+1}^{J} WIP_{j'}(d) + QC\_OUT(d), j = 1,...,J,$$

where $QC\_OUT(d)$ is the amount of output wafers up to the beginning of the current day.

The demanded (to be fulfilled) moves by MPS for the current day is

$$MPS\_day\_demand_j(d) = Max\{DD_j(d) + Cum\_DD_j(d) - Actual\_move_j(d), 0\}, j = 1,...,J.$$

Output is generated for each day and Output format of MPSC is given in Table 1.

III.2 CE Module

Inputs of the CE module include

- tool ID, type and stage mapping tables,

- multi-chamber tool ID table,

- tool time data of the previous day such as up, lost, backup and hold times,

- actual moves of the previous day,

- capacity estimate for the previous day, and

- process flow of individual part types.

Let $c_m(t)$ be the capacity estimate of day t for tool group m and $C_m(t)$ be the auto-regression (AR) estimate for tool group m. Previous day utilization of each tool is first calculated by utilizing the tool time data. The value of $c_m(t)$ is obtained by

$$c_m(t) = \sum_{k \in M_m} \frac{\text{actual moves of tool k}}{\text{utilization of tool k}},$$

where the summation is over all the tools belonging to tool group m. The AR estimate for tool group m during day t is then

$$C_m(t) = \alpha\, c_m(t) + (1-\alpha)C_m(t-1),\ 0 \le \alpha \le 1.$$

IV. Target Setting Main Function

Reference WIP levels are critical to target setting. In TSMF, the reference WIP level of a stage is set as the average WIP level under the MPS demanded moves of the next day, which can be calculated by applying Little's formula. Such a definition is intended to prepare for the next day operation through efforts of the current day. A PULL procedure is first used to calculate the up-to-date cumulative production demand of each part at each stage. Tool capacity is allocated to such demands proportionally among individual stages[1]. For the residual tool capacity and WIPs after PULL, a PUSH procedure then tries to push as many as possible residual WIPs of individual stages to their respective downstream stages so that the tool utilization can be maximized and cycle times can be reduced. Target moves of individual part types at a stage are obtained by adding up, across all part types, the PULL targets and PUSH targets of the stage. Note that in the one hand, the available WIP of individual stages affects the setting of reasonable targets; on the other hand, target setting/tool allocation affects the flow-in from up-stream stages and hence available WIPs of individual stages. Cycle time statistics are also exploited to pose a heuristic constraint on updating the amount of flow-in

wafers to each stage. The maximum number of flow-in wafers for a stage is the sum of initial WIPs of upstream stages that are within one-day cycle time to the stage under consideration. The PULL, PUSH and Flow-in updating procedures constitute an iteration of TSMF and are repeated till the calculated set of stage targets converge. Figure 1 gives the functional flow chart of TSMF.

To formalize the description of the algorithm, let us add the part type index to the previously defined notations and further define some new notations.

Notations

I: total number of part types;

i: part type index, $i = 1, \ldots \ldots I$;

Tom_Dem$_{ij}$: daily MPS demand for type i at stage in the next day;

WIP$_{ij}$: the WIP level of type-i parts at stage j at the beginning of the day;

Ref_WIP$_{ij}$: the desirable WIP level preparing for next day production at stage j.

Flow_in$_{ij}$: number of type-i wafers flowing to stage-j from its up-stream stages during the day;

R$_i$: amount of type i wafer start for the day;

M: the total number of tool groups;

m: the tool group index, $m=1, \ldots, M$;

C$_m$: estimated capacity of a tool in group m in term of wafers per day;

S$_m$: the set of all part type and stage pair (i, j) that require tool group m for processing;

Decision Variable

Target$_{ij}$: number of type-i wafers leaving stage j to its down-stream stages.

TSMF Algorithm

Step 0: Initialization

Input all the necessary data.

Set Flow_in$_{i1}$ = R$_i$ for all i, and Flow_in$_{ij}$ = 0 for all i and j=2,…J.

For $i = 1, \ldots, I$

Step1: Compute the reference WIP levels.

Ref_WIP$_{ij}$ = Tom_Dem$_{ij}$ * CT$_{ij}$, j=1,…,J.

Step 2: Compute the maximum number of flow-in wafers to each stage,

For j=J, J-1, …, 2,

Max_flow$_{ij}$, = 0,

Time = CT$_{ij}$,

$$j' = j,$$

While Time < 24 hours and j > 1,

$$j' = j' - 1,$$

$$\text{Max\_flow}_{ij'} = \text{Max\_flow}_{ij'} + \text{WIP}_{ij'},$$

$$\text{Time} = \text{Time} + \text{CT}_{ij'}.$$

Pull Procedure

    Step 3: Calculate the PULL demand for each stage by considering the demand of the stage itself, the downstream stage demand and the number of available WIP.

        For j=1,...,J

$$\text{Pullf}_{ij'} = \text{Max} (0,\ \text{MPS\_day\_demand}_{I(k+1)} + \text{Ref\_WIP}_{I(k+1)} - \text{WIP}_{I(k+1)}),$$

$$\text{Pullf}_{ij'} = \text{Max} (\text{MPS\_day\_demand}_{ij},\ \text{Pullf}_{ij}),$$

$$\text{Pullf}_{ij'} = \text{Min} (\text{Flow\_in}_{ij} + \text{WIP}_{ij'},\ \text{Pullf}_{ij}).$$

Step 4: Allocate tool capacities by part type and by stage proportionally to their WIPs.

        For m=1,...,M,

            For each part type i that uses tool group m at stage j,

$$\text{Pull}_{ij} = \begin{cases} \text{Pullf}_{ij}, \text{if } \sum_{\substack{i=1 \\ }}^{I} \sum_{\substack{j'=1 \\ (i,j') \in S_m}}^{J} \text{Pullf}_{ij'} \le C_m, \\[2em] \text{Pullf}_{ij} \times \dfrac{C_m}{\sum_{\substack{i=1 \\ }}^{I} \sum_{\substack{j'=1 \\ (i,j') \in S_m}}^{J} \text{Pullf}_{ij'}}, \text{otherwise.} \end{cases}$$

Push Procedure

For i = 1,...,I

    Step 5: Calculate the remaining WIP of each stage that can be pushed to the down stream stages.

        For stages j=1,...,J,

$$\text{Push}_{ij} = \text{Max} [0,\ \text{Flow\_in}_{ij} + \text{WIP}_{Ij} - \text{Pullf}_{ij}].$$

Step 6: Allocate residual capacity of tools by part type and by stage proportionally to their remaining WIPs.

        For m=1,...,M,

$$\overline{C}_m = C_m - \sum_{\substack{i=1 \\ }}^{I} \sum_{\substack{j'=1 \\ (i,j') \in S_m}}^{J} \text{Pull}_{ij'}.$$

For each part type i that uses tool group m at stage j,

$$
Push_{ij} = \begin{cases} Pushf_{ij}, & \text{if } \displaystyle\sum_{\substack{i=1 \\ (i,j')\in S_m}}^{I}\sum_{j'=1}^{J} Pushf_{ij'} \leq \overline{C}_m, \\[4ex] Pushf_{ij} \times \dfrac{\overline{C}_m}{\displaystyle\sum_{\substack{i=1 \\ (i,j')\in S_m}}^{I}\sum_{j'=1}^{J} Pushf_{ij'}}, & \text{otherwise.} \end{cases}
$$

Step 7: Generate new targets and check convergence.

$Target_{ij} = Pull_{ij} + Push_{ij}$, I=1,...,I, j=1,...,J.

If $\displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\left| Target_{ij} - Flow\_in_{i(j+1)} \right| \leq \Theta$, then

Output $Target_{ij}$ and Stop.

Step 8: Update Flow-in and iterate.

$Flow\_in_{ij} = Min\ (\ Target_{i(j-1)}\ ,\ Maxflow_{ij}\ )$

Go to Step 3 for the next iteration.

V. Implementation Results

In designing and implementing TSS, the desirable features include (1) meeting the production demands of MPS, (2) line balancing, (3) mean cycle time and variance reduction, (4) manual data maintenance reduction, (5) maximization of daily moves, (6) wafer flow-in forecast, and (7) target adjustment for banking and scrap.   Table 5.1 summarizes the coverage of these features by individual TSS functions.   Note that although the data files of key tool, tool subtype and its mapping to production stages require manual maintenance, they are maintained at a monthly or quarterly frequency.

TSS has been implemented for daily application.   Production supervisors briefly review and adjust the targets generated by TSS every morning. Targets are then delivered to shop floor after the production meeting.   There have been less then 20 out of a total of 140 production stage targets each day that require some adjustments.   Figure 2(a) depicts that the scheduled total moves by TSS (before manual adjustments) are mostly within 5% of what were actually achieved in the shop floor.   Since

the daily capacity variation of tool is about 10 - 15% of the forcasted capacity, we consider the scheduled daily target of a stage accurate if it is within 15% of the actually achieved number of moves. The performance of TSS on stage target accuracy is given by Figure 2(b), which shows that about 50% of scheduled stage targets are accurate.

Detailed analyses indicate two potential cause of the by-stage inaccuracy of TSS. The first is the discrepancy between the demanded moves calculated from MPS and what are desired in a day by the shop floor managers. As TSS is driven by MPS, such an observation suggests further integration between MPS and shop floor scheduling in operational philosophy. The second is the use of empirical cycle times, which are very well known to be in accurate; they vary a lot from day to day with respect to actual WIP distribution and available tool capacity. TSMF does not adopt the stage of penetration estimation algorithm (SOPEA) as the previous TGMA algorithms do because SOPEA requires accurate stage processing times which are difficult to maintain in a foundry fab. This second cause is a tradeoff between accuracy and maintainability. In addition, coarse modeling of key tools of individual stages may also contribute to the inaccuracy.

VI. Conclusions

In this paper, a computer aided decision support system, TSS, was designed for daily target setting. It includes three modules: (1) Master Production Schedule Conversion (MPSC), (2) Capacity Estimation, and (3) Target Setting Main Function (TSMF). MPSC and the PULL procedure of TSMF drives to meet the higher level, mid-term (MPS) targets, while the PUSH procedure of TSMF tries to maximize capacity utilization. The auto-regression based CE automatically yields good estimates for target setting purpose. Capacity allocation proportional to PULL and PUSH demands among stages leads to line balancing effect and PULL demands are given a higher priority in allocation. Modifications on flow-in estimation heuristics and fine-tuning of tool models are now being conducted to improve the by-stage taget setting accuracy.

References

[1] S.-C. Chang, L.-H. Lee, L.-S. Pang, T. Chen, Y.-C. Weng, H.-D. Chiang, D. Dai, "Iterative Capacity Allocation and Production Flow Estimation for Scheduling Semiconductor Fabrication," *17th IEEE International Electronics Manufacturing Technology Symposiu*m, Austin TX, Oct., 1995, pp. 508 – 512.
[2] S.-C. Chang et al, "Smart Monitoring and Control Aide – Final Report," *Technical Report,* Dept.

of Electrical Engrg., National Taiwan University, Taipei, Taiwan, R.O.C., March, 1996.

[3] W.-L. Jan, S.-C. Chang, I.-C. Hsu, S.-R. Huang, K.-C. Lin , "TG&MA-Priority," *Technical Report*, Dept. of Electrical Engrg., National Taiwan University, Taipei, Taiwan, R.O.C., 1995.

[4] W.-L. Jan, "Analysis of Proportional Tool Capacity Allocation Scheme in a Deterministic Re-entrant Line," M.S. Thesis, Dept. of Electrical Engrg., National Taiwan University, Taipei, Taiwan, R.O.C., June, 1994.

[5] R. Uzsoy, C. Lee, L. A. Martin-Vega, "A Review of Production Planning and Scheduling Models in the Semiconductor Industry," *IIE Trans.*, Vol.26, Sept. 1994, pp. 44-55.

[6] T.-H. Wang, "Design and Analysis of Daily Production Control Methods for Semiconductor Wafer Fabrication," M.S. Thesis, Dept. of Mech. Engrg., National Taiwan University, Taipei, Taiwan, R.O.C., June, 1994.
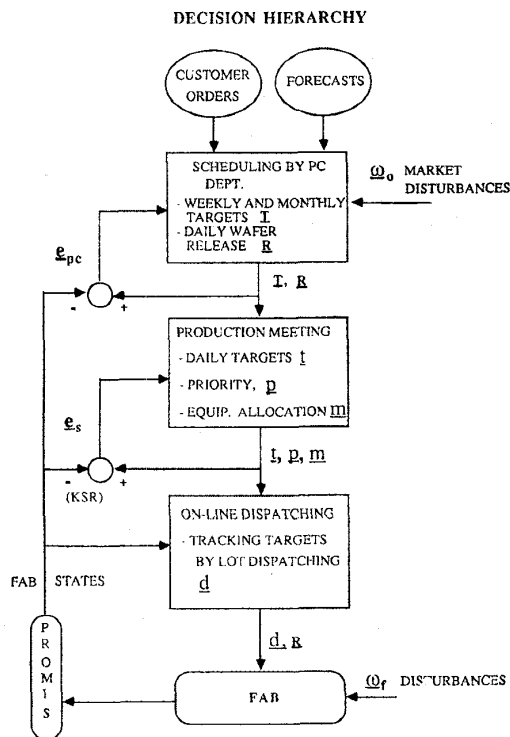
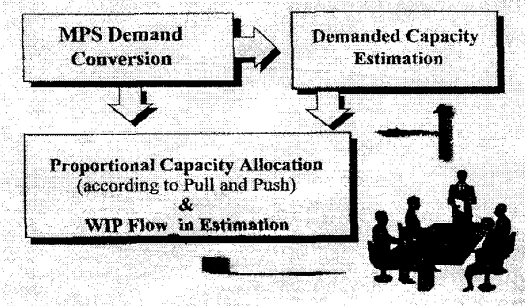Figure 1: decision Hierarchy of Production Scheduling
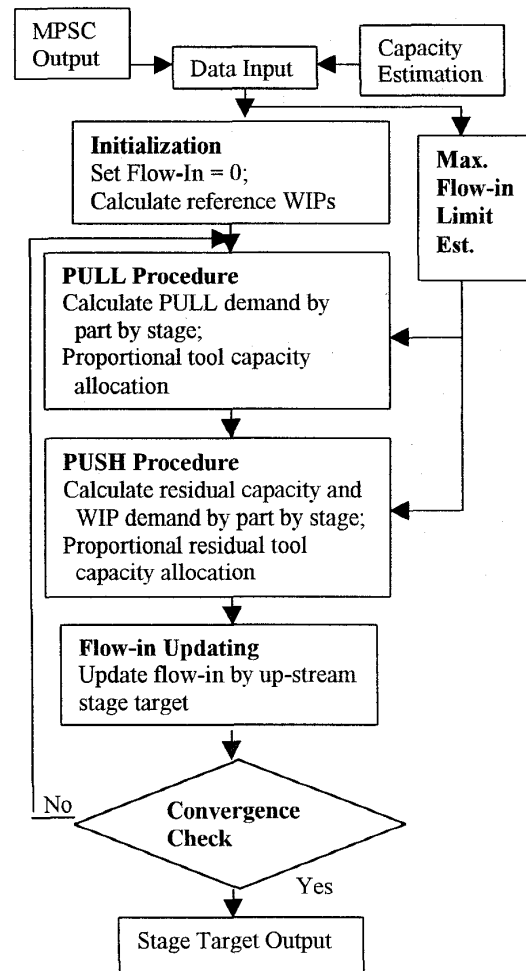

Figure 2: TSS System Functional Flow Chart
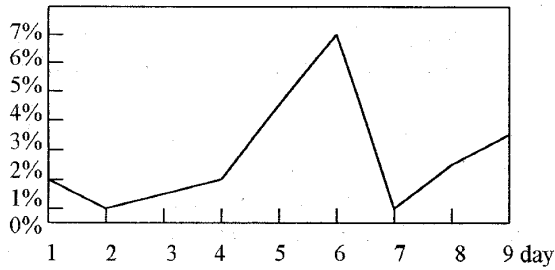

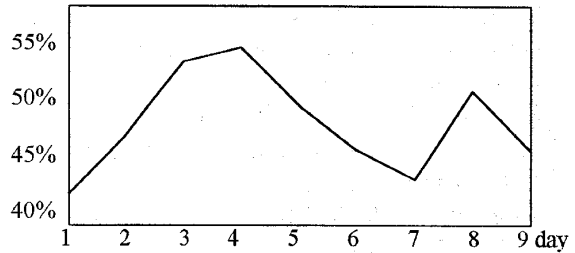Figure 1: TSMF Flow Chart

Figure 4(a): TSS Performance on Total Moves

Figure 4(b): TSS Performance on Stage Targets

Table 1: MPSC Output Format

PART	STAGE	Cum_Dem	Cum_Mov	To_Dem	Tom_Dem	WIP	CT	Tool
NTU001	WAF-START	96	96	0	0	0	3.84	Marker
xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx

To_Dem: MPS_day_demand for today　　　Tom_Dem: MPS_day_demand for tomorrow

Table 2 Production Support Features by TSS Functions

	Meeting MPS Demands	Line Balancing	C/T mean & Variance Reduction	Mannual Data Maintenance Reduction	Maximization of Total Moves	Wafer Flow-in Forecast	Banking and Scrap Re-setting
MPSC	direct	direct	indirect				direct
CE				direct			
PULL	direct	indirect					
PUSH			direct		direct		
Flow_in Calc.						direct	