A dynamic system model for proactive control of dynamic events in full-load states of manufacturing chains

H. -Y. Huang [a]; Y. -C. Chou [a]; S. Chang [a]
[a] Graduate Institute of Industrial Engineering, National Taiwan University, Taiwan

First Published on: 22 October 2007

PLEASE SCROLL DOWN FOR ARTICLE

A dynamic system model for proactive control of dynamic events in full-load states of manufacturing chains

H.-Y. Huang, Y.-C. Chou* and S. Chang

Graduate Institute of Industrial Engineering, National Taiwan University, Taiwan

Risk management is a major concern in supply chains that have high levels of uncertainty in product demand, manufacturing process or part supply. The uncertainties frequently manifest as dynamic events that pose a threat to interrupting supply chain operation. Depending on the nature and severity of uncertainty, the impact of dynamic events can be distinguished into three categories: deviation, disruption, and disaster. Many studies in literature addressed modelling of deviation events. In this paper, a dynamic system model of supply chains is described which can be applied to managing disruptive events in full-load states of manufacturing chains. An example of disruptive events is given which arises from demand shocks in distribution channel. The procedure to construct full-load production functions of complex manufacturing nodes with internal queuing delay is described. Analytic optimal solution is derived for the dynamic model. Given an unordinary event of demand shock, this model can be used to determine if demand shock can be absorbed by a manufacturing chain and the level of contingent resources that must be synchronously activated in multiple nodes of the chain. This model can be used to reduce what could have been a disruptive event into a deviation event, thus enhancing risk management.

Keywords: dynamic systems; full-load production function; risk management; supply chain management; disruption events; optimal control

1. Introduction

In the literature of supply chain management, inconsistent use of uncertainty and risk is very common. There are many definitions and classifications. In this section, we will discuss a conceptual framework for these concepts before describing the problem that this paper addresses. We make no claim that our conceptual framework is comprehensive. As was found in an industry survey by Jüttner *et al.* (2003), these concepts are industry-, supply chain-, or even firm-dependent. They advocated future studies on 'different supply chains and/or industries and apply a contingency perspective'. In the following paragraphs, uncertainty sources, risk and exposure are first defined.

Uncertainty and risk management are major concerns in recent literature of supply chain management. There are at least two reasons. The first reason is theoretical. Uncertainty and its modelling are intellectually intriguing in and by themselves. Almost a century after the seminal work of Frank Knight (1921), discussions on uncertainty and risk

*Corresponding author. Email: ychou@ntu.edu.tw

still appear frequently in the literature, such as in (LeRoy and Singell 1987, Runde 1998, Holton 2004). The second reason is practical. That supply chain systems are vulnerable to the influence of many factors of uncertainty is well known. Recent events, such as the terrorist attack in New York City, the 1999 earthquake in Taiwan, and the electricity blackout that struck North America in 2003, repeatedly disrupted global supply chain operation and their awakening consequences have triggered a spate of theoretical and applied research (Sheffi 2001, Kleindorfer and Saad 2005, Blackhurst *et al.* 2005).

Frank Knight's definitions of risk and uncertainty are based on a threefold classification of unknown outcomes: *a priori* probabilities, statistical probabilities, and estimates. *A priori* probabilities are those that can be derived deductively, such as in rolling dice; statistical probabilities are those that are generated by empirical evaluation of relative frequencies; estimates are subjective judgments that are necessitated in the situations in which the probabilistic calculus is inapplicable. Thus, the term uncertainty can be, and has been, used in two ways: broad or narrow senses. In the broad sense, uncertainty is used to refer to any situations with unknown outcomes. In the narrow sense, uncertainty refers to the last class, with risk referring to the first two classes. Uncertainty is differentiated from risk by our incapability (or capability) to construct probabilities on possible outcomes. If well-defined probabilities can be constructed, the underlying phenomenon is considered to be in the realm of risk. Otherwise, it is in the domain of uncertainty. This distinction can also be associated with the concepts of profit and windfalls in economics. Whereas profit comes from taking risk, windfalls are associated with uncertainty. In our survey, we found that the majority of models in the literature of supply chain management are used to address risk factors. The focus of this paper, however, is on non-risk uncertainty. In this paper, to avoid ambiguity, the term uncertainty will be used in the broad sense except for a few noted occasions. We will use risk factors and events for the narrow sense whenever ambiguity might otherwise arise.

In supply chains, the sources of uncertainty exist in supply nodes, production processes, and demand and distribution nodes (Figure 1). Most papers include some modelling elements of risk factors. In modelling a two-stage inventory system for determining optimal base stock level, Bollapragada *et al.* (2004) used a Gamma distribution to represent uncertain supplier capacity in each period. In modelling a multi-stage serial production system for work release control of light-emitting diode fabrication, Lee and Yano (1988) used a yield rate distribution (histogram) to represent the uncertainty of yield loss in production. Demands or job arrivals are commonly modelled by random variables, stochastic processes (Chou *et al.* 2007) or smoothed forecast process (Graves *et al.* 1998, Ni *et al.* 2003). These models suggest that the underlying risk factors are regarded as fundamental and exogenous in nature. (Although production equipment resides inside production systems, equipment breakdowns and process yield problems are exogenous in so far as supply chain management is concerned.)

The effects of uncertainty factors propagate in supply chains, and it is a key objective of supply chain management to design new processes or systems to mitigate or pre-empt those effects that are potentially damaging. Kleindorfer and Saad (2005) described an analysis and mitigation framework for managing disruption risks, formulated 10 guiding principles of implementation, and presented an application in the chemical industry. Vorst and Beulens (2002) analysed uncertainty factors in the food industry and discussed effective supply chain re-design strategies for various sources of uncertainty. In their typology of uncertainty sources, the planning and control (P&C) function of supply chains is a fourth area of uncertainty. However, as pointed out by Jüttner (2005), the control
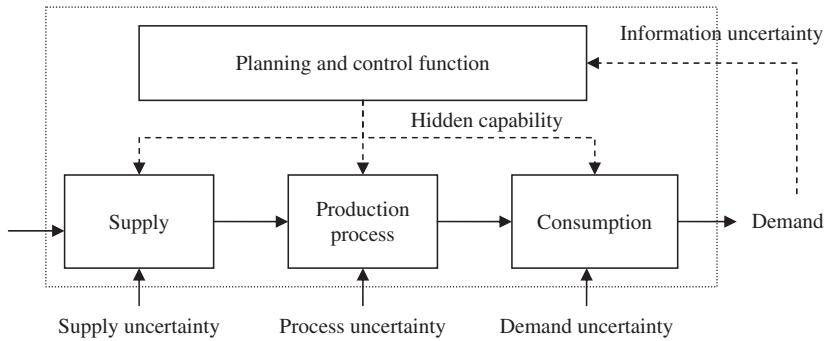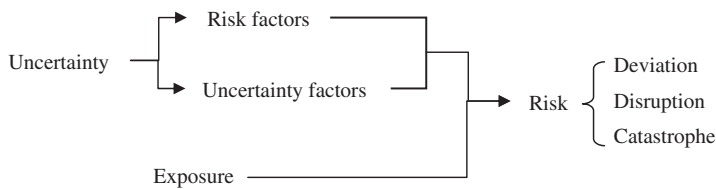
Figure 1. Sources of uncertainty factors.

Figure 2. Uncertainty factors, exposure and risk.

function can act as an absorber or amplifier of risk effects. The performance of P&C is hinged on its planning logic, control policy and requisite information. A supply chain will be robust if its control function is bestowed with sound logic and policy and provided with requisite information. If it is the case, the control function will be able to absorb the impact caused by uncertain events. In the opposite case, the control function might, on the contrary, be an amplifier and the supply might not be stable. That P&C is a fourth source of uncertainty can be traced to information availability, accuracy and acquisition time lag. Finally, there is a fifth category of uncertainty which covers natural accidents, socio-political actions, and natural disasters. Based on the above analysis, the sources of uncertainty factors can be located in two levels as depicted in Figure 1. At the physical flow level, the three sources of uncertainty factors are exogenous to the production firm. At the planning and control level, the risk factors in planning logic, control policy and information availability are not exogenous. Instead, they can be associated with economic tradeoffs and organisational constraints. The fifth class of uncertainty is in the environment.

As defined previously, risk factors are a term we use to describe a subclass of uncertainty. The term risk is frequently used in the literature to refer to the consequence of uncertainty. We will use Figure 2 to clarify our usage of these terms to avoid ambiguity. The existence of uncertainty sources does not necessarily constitute a risk to a firm or a supply chain. Risk has two constituent elements: uncertainty and exposure (Holton 2004). For example, a firm can protect itself from demand uncertainty by elevating the safety stock level or by reserving capacity buffer. Thus, two firms might have different exposure to the same exogenous uncertainty source, and, consequently, two different risk levels. Exposure can be reduced by supply chain re-configuration, information sharing, or other strategies. If exposure is contained, risk will be reduced.

Risk (and uncertainty sources) can be classified into at least three categories of increasing severity: deviation, disruption, and catastrophe (Gaonkar and Viswanadham 2004). If supply chains are regarded as systems, deviation refers to variations in system parameters or performance outcomes. In general, we would be able to cope with deviations by containing or protecting exposure. However, it is not economical, or possible, probabilistically speaking, to have complete protection. In an uncertain environment, there will always be remnant risk. At other times, the impact of uncertainty events is so severe and far reaching that normal operation is disrupted for some duration of time and system performance becomes unpredictable. It will take contingent remedial actions to bring the system back to a more stable state. Such events, which we will call dynamic events, do not occur on a regular basis. Thus, probabilistic distributions cannot be reasonably constructed. We will describe an example of such events in Appendix A. The third category is catastrophes. Natural disasters, political instability, and industry accidents have the potential of altering the configuration of supply chains. Although several papers have emphasised the need to mitigate and to contain catastrophic events (Kleindorfer and Saad 2005), we have found no mathematical models of supply chain control for this class of problems. In Appendix B, we will describe an example of unplanned changes in supply chain organisation which result from a single industry accident at a factory.

Several recent studies on research needs of supply chain risk management have yielded very insightful and practical suggestions. Through a multi-form, multi-stage study of the automotive supply chain, Blackhurst *et al.* (2005) compiled an agenda of critical research issues for managing disruptions. The agenda encompasses three sections: disruption discovery, disruption recovery, and supply-chain redesign. Their major findings of research needs are:

(1) Visibility enhancement and the economics of visibility.
(2) Prediction of capacity bottleneck in transportation networks and manufacturing facilities.
(3) Predictive analysis tools for the possibility of occurrence and impact of disruption events (such as for predicting lead-time).
(4) Supply chain reconfiguration.
(5) Reachability analysis for damage control.
(6) Understanding the tradeoffs between cost and flexibility.
(7) Dynamic supply-chain optimisation tools.

Jüttner *et al.* (2003) reported on an exploratory survey of research agenda for supply chain risk management by using semi-structured interviews with managers in manufacturing, retail and logistic service industries. Their study suggests future research in:

(1) Assessing the risk sources.
(2) Defining the risk concept and adverse consequences.
(3) Identifying the risk drivers of the supply chain strategy.
(4) Mitigating risk.

They found out that concepts about risk exposure and consequence are firm-, network-, and industry-dependent. They also noted that risk can come from a drive to be lean and efficient, to globalise production, and to outsourcing. Thus, there is a need to investigate tradeoffs between risk and efficiency. Of the 10 guiding principles formulated in Kleindorfer and Saad (2005) for managing disruption risk, the sixth principle is particularly relevant to this paper and it states that: 'Establishing backup systems,

contingency plans, and maintaining reasonable slack, can increase the level of readiness in managing risk.'

In practice, when supply operation is disrupted, firms will seek to find alternative production methods or to modulate and temper demand. Tang (2006) described several approaches of using strategic stocks, flexible supply base, economic supply incentives, flexible transportation, dynamic pricing to influence demand, and dynamic product assortment management. These approaches amount to utilising contingent or backup resources.

We have found no mathematical models in the literature for assessing the impact of disruptive events. In this paper, we discuss a type of supply chain problems in which dynamic events that arise from uncertainty sources can not be satisfactorily modelled by probability distribution. Those events cannot be known *a priori*, but they have the potential of disrupting supply chain operations. Therefore, when they do occur, firms must take immediate actions to assess their impacts and, if necessary, to activate contingent plans for their mitigation. Two fundamental questions arise from this capability of risk management. One, how to determine if a dynamic event is disruptive, rather than deviational? Two, how to mitigate the impact of a disruptive event? In this paper, a dynamic system model is described as solutions to these two questions. Of the research needs identified in Blackhurst *et al.* (2005), this model can be used to predict capacity bottleneck, and analyse the impact and the reach of disruption events. This model can also be applied to the sixth principle formulated in Kleindorfer and Saad (2005) for establishing backup systems, contingency plans, and maintaining reasonable slack to increase the level of readiness in managing risk.

The remainder of this paper is organised as follows. In section 2, the problem is defined and a dynamic system model for supply chains is described. Because the impact of disruptive events is the most severe when supply chains are approximately fully loaded, a procedure for constructing full-load production functions is presented in section 3. section 4 is devoted to solution derivation for the dynamic system model and a numerical example. Finally, discussions and conclusions are given in section 5.

2. Problem description

Figure 3 shows a representation of 2-node manufacturing chains. A node can be considered as a manufacturing shop or factory and the numerical subscript is used to index the nodes. The input to the manufacturing chain is a planned demand process $d(t)$. Release rate decision $r_1(t)$ is made based on $d(t)$. Normally, $r_1(t)$ either is the same as $d(t)$ or closely follows $d(t)$. Each node ($k = 1$ or 2) has a capacity C_k, a state variable of work-in-process W_k, and an output rate $\Omega_k(t)$. The release rate decision at node 2 also closely follows $\Omega_1(t)$. The output of the chain is sent to the distribution channel. The channel inventory is represented as $I(t)$. Market demand is usually different from the forecast $d(t)$. Their difference is represented by a random process $\phi(t)$. In many industries, distribution channels have multi-tiers. Information delay and opacity prevent full assessment of $\phi(t)$ in real time. Thus, the difference will accumulate over time and reveals itself in the form of demand information shock. In the following, we will first describe these dynamics of supply and demand using semiconductor IC chip manufacturing as an example.

Semiconductor manufacturing typically has four stages of manufacturing plants. The distribution channels of IC chips comprise foundry manufacturers, microchip product
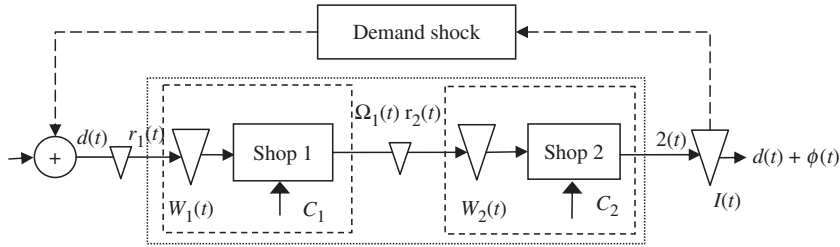
Figure 3.   Supply and demand dynamics of manufacturing chains.

companies, electronic manufacturing service companies, distributors and retailers. Being at the starting point of the supply chains for the computer, consumer electronics, communication equipment and many other industries, the demand facing semiconductor manufacturing is very dynamic. The inventory in the channel fluctuates in an unpredictable way. (Please refer to Appendix A for illustrating data.) It is a repeated phenomenon that channel inventory swings from excessive to below norm and back to excessive. Since replenishing channel inventory is a source of demand origination, the inventory level will affect the demand $d(t)$ as indicated by the outer feedback loop in Figure 3. This feedback, however, is not continuous, but is in the form of *ex post* events. In the figure, manufacturing facilities 1 and 2 together can be regarded as a system, with input $r_1(t)$ and output $\Omega_2(t)$.

Channel inventory information can arrive as a shock input to the manufacturing chain. When channel inventory is discovered to be lower than what has been anticipated or planned for, it could be inferred that true market demand of the end products is stronger than expected. Urgent orders thus would follow to replenish the inventory. If the ϕ is a white noise or if the manufacturing chain has ample built-in reserves, the system will be stable. If the ϕ is neither a white noise nor a well-behaved random process and if the chain is lean, the system will be unstable. The manufacturing chain is in fact a system that provides manufacturing services to the customers by fulfilling their needs. The predictability of the services will be dependent on the characteristics of $d(t)$ and when $d(t)$ has a large fluctuation, the controllability of fulfilment services will be reduced.

The impact of disrupting events is most severe when the chain is fully loaded. We assume that the chain is highly loaded and demonstrate how to assess and mitigate the impact of shock input by coordinating the production decisions at the node levels. Figure 4 illustrates a general framework for proactively managing dynamic events. Suppose that channel inventory is surveyed at discrete time t_0. The discovered excessive inventory is represented as $S(t_0)$ or simply S, which can be regarded as a scalar (for one product type) or a vector (for multiple product types). It is converted to some urgent demand forecast $E(t)$ over a short run horizon $[t_0, t_0 + T]$ after its nature is analysed. This analysis process is likely to involve management judgments. The urgent demand is a second source of demand stream. It is distinct from the nominal demand input $\bar{d}(t)$, but is merged with $\bar{d}(t)$. (Note that in Figure 3, the $d(t)$ includes the demand shock, whereas in Figure 4 the demand shock is separated from the nominal stream of demand.) The thesis of this paper is that the stability of the system can be enhanced and performance be improved by putting in place a shorter feedback loop and by coordinating $E(t)$ with production decisions at the production nodes.
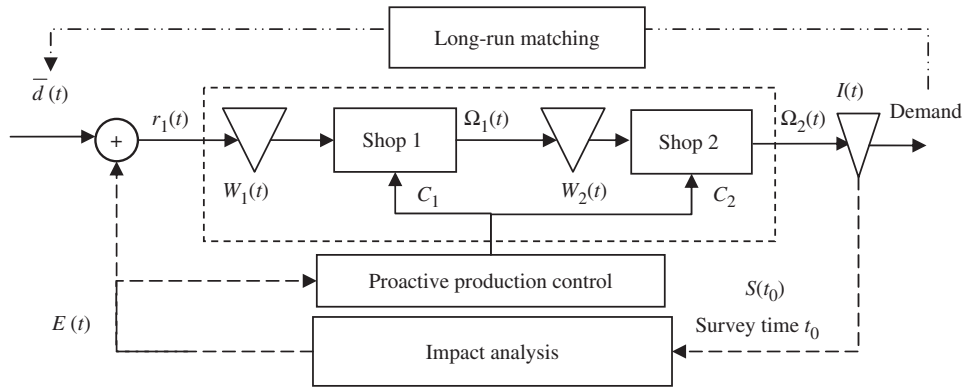
Figure 4. Proactive production control for managing dynamic events.

The problem setting can be summarised as follows at the abstract level.

(1) The occurrence of dynamic events cannot be predicted. But when a dynamic event occurs, it becomes useful leading information to the manufacturing chain in planning and tuning the supply capability. Let the time of occurrence be t_0.

(2) In the long run, the demand input $\bar{d}(t)$ is matched with market demand. But in the medium run, the difference might be significant. This implies that in the short run ϕ is a random variable, in the medium run $S = \int_0^{t_0} \phi(t)dt \neq 0$, but in the long run $\int_0^{t_0=\infty} \phi(t)dt = 0$.

(3) The demand input to the system has two parts. One is nominal demand $\bar{d}(t)$ and the other is marginal demand $E(t)$, $t \in [t_0 t_0 + T]$, with the following relationship $\int_{t_0}^{t_0+T} E(t)dt = \int_0^{t_0} \phi(t)dt$, where T is the duration used to digest S.

The production control function takes $E(t)$ as input information and exerts control on the shop operation. To design the proactive production control function, the dynamics of the chain system must be analysed and modelled. A production operation model of supply chains has two constituent elements: production functions of constituent production units, and a linking mechanism for integrating multiple production functions. A production function is a model for the behaviour of the production unit. In economics, a production function relates throughput to capital and labour. Similarly, a production function in supply chain management describes the relationship between performance measures and production decisions in a factory. We will describe production function modelling in the next section. The linking mechanism will be discussed in section 4.

3. Full-load production functions for queuing manufacturing systems

A production function is a mathematical relationship between system output (such as throughput rate and performance measures) and system input and operation state (such as input rate, capacity, variety, etc.). Time-delay models and clearing models are two approaches to modelling production functions. The focus of time-delay models is on the flow time that a job order will spend in passing through a production unit. Many complex production units, such as wafer fabrication plants (Conners and Yao 1996), exhibit
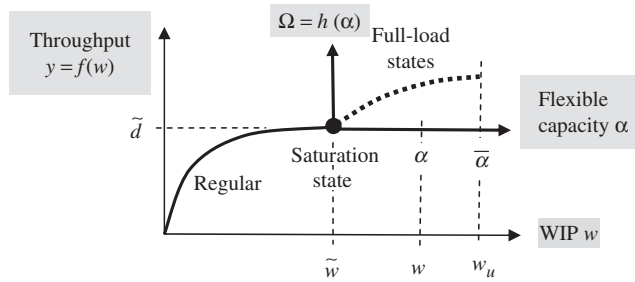
Figure 5. Full-load production function of queuing manufacturing systems.

significant queuing delay phenomenon. The flow time includes raw processing times and queuing delays. In practice, factories use scheduling (Lu *et al.* 1994) and work release policies (such as the CONWIP policy) to control the escalation and variation in flow time. Therefore, except for very short term, flow time could be treated as fixed. Achieving fixed flow time, however, requires endeavour in managing bottleneck machines, improvising alternative machines, and reducing variations in capacity availability and internal job arrival (Delp *et al.* 2005).

The second modelling approach is to view production as an operation of clearing up the WIP stock by processing them and sending them out of the factory. The resulting output models are mathematical functions that relate system output (such as throughput rate and other system performance measures) to system state variables such as workload, capacity and variety (Karmarkar 1989, Missbauer 2002). Output functions are also called clearing functions. The two approaches are actually not options of modelling. Instead, they represent different perspectives and their models describe different aspects of system behaviour. Adopting or integrating both views will provide a more complete description of system behaviour.

Figure 5 illustrates the concept of full-load production function that we propose. In the figure, the production function is made up of two segments. There is a regular segment for which workload is less than nominal capacity. This segment starts at the origin and ends at a saturation point. It relates the output rate to WIP level. It is this segment that is addressed in papers such as Karmarkar (1989) and Missbauer (2002). Beyond the saturation state, workload will be greater than the nominal capacity and the system is said to be in full-load states. These states of excessive workload could happen when there is a large shift in product mixes, an increase in workload due to rework, or simply an increase in demand orders. According to queuing analysis, as the arrival rate approaches the processing rate, flow time will grow exponentially to infinity. But this has never been allowed to happen in practice. When the workload exceeds capacity, instead of allowing flow time to escalate, contingent capacity, in the form of alternative routing, reactive scheduling, or postponement of maintenance, is activated to rein in the flow time. It is this phenomenon that is to be modelled by the second segment (the dotted curve), thereafter called the full-load production function, or full-load function for brevity.

In contrast to the regular production function, a full-load function describes the behaviour of a system when under active control. Therefore, its functional form will be determined by the control policy that is in effect. Secondly, a full-load function starts with a saturation state. A sound full-load function requires that this initial state be specified. In this paper, we consider complex manufacturing systems that have prominent

queuing phenomenon. Such systems will be called queuing manufacturing systems and wafer fabrication plants are a prominent example. (Some modern plants in Taiwan contain more than one thousand machines.) We will discuss the saturation state and control policy for this type of systems before describing a procedure for constructing full-load functions.

In queuing manufacturing systems, effective throughput and flow time are two of the most important measures of manufacturing performance. The saturation state is represented by the coordinates $(\tilde{w}, \tilde{d})$ in the w–y frame and, following the Little's formula, the average WIP is K times that of throughput rate:

$$\tilde{w} = \kappa \cdot \tilde{d}. \tag{1}$$

The point $(\tilde{w}, \tilde{d})$ represents an ideal state, its throughput $\tilde{w}$ can be considered as the system capacity, and K is, in fact, the average flow time. Since different factories might have different capacity level and different criteria for what the saturation state is, the K is a parameter to be instantiated. In addition, both flexible capacity and marginal throughput Ω are normalised by system capacity. Flexible capacity is represented by a fractional value parameter α, with domain $0 \leq \alpha \leq \bar{\alpha}$. The full-load function then is expressed as $\Omega = h(\alpha; \tilde{w}, \tilde{d})$, to explicitly show its dependence on the starting state $(\tilde{w}, \tilde{d})$.

Because flow time becomes a constraining factor in full-load states, the mode of manufacturing management should change to that of managing and controlling flow time or other congestion effects brought about by excessive work. For any α, $0 \leq \alpha \leq \bar{\alpha}$, if the relationship of Equation (1) was maintained, the WIP as a function of α would be:

$$w(\alpha) = \kappa \tilde{d}(1 + \alpha) = \tilde{w} + \kappa \alpha \tilde{d}.$$

This would imply that contingent capacity is of the same nature as regular capacity. The behaviour of the system at the full-load states would be a simple but erroneous extrapolation of that of the saturation state, as though 'active control' is absent. To rectify this deficiency, we stipulate the following objective for the control policy: the objective of control policy is to activate contingent resources so that $w(\alpha) \approx \tilde{w} + \alpha \tilde{d}$. We offer our argument as follows: contingent resources are not operational resources. They are tactical resources and they are surgically commissioned for the purpose of relieving a congested system. If $\alpha \tilde{d}$-equivalent of work is processed by contingent resources, it is desirable that the amount of WIP that is introduced is $\alpha \tilde{d}$, not a certain multiple of $\alpha \tilde{d}$, such as $\kappa \alpha \tilde{d}$. Thus, we stipulate that $w(\alpha) = \tilde{w} + \alpha \tilde{d}$ is a fitting, although aggressive, goal.

There are many possible contingent resources. In the remainder of this section, we will analyse alternative routing as a means to provide relief capacity. Alternative routing is also known as machine backup planning or machine grouping in practice (Johri 1993). In full-load states, products that have been assigned to the bottleneck machines could be re-routed to alternative machines to increase system throughput.

3.1 *Simulated data for alternative routing*

A set of industry data of wafer fabrication is used in this study on alternative routing. The data set contains five products and 73 machine types. The processing steps of the process routings range from three hundreds to five hundreds. Machine data includes such attributes as equipment type, availability and batch size of loading. This data set represents a case of multiple-product, multiple-stage manufacturing problems. The effect of alternative routing is dependent on product mixes. Since many product mixes must be

considered, product mixes are simulated in this study. The procedure used to generate simulated data is:

Step 1: For a given total volume of production $\tilde{d}$, generate product mix scenarios (indexed by s).

Step 2: For each scenario:

> **Step 2.1:** Determine a machine portfolio. The machine portfolio and its associated product mix are necessary information to define their saturation state.
>
> **Step 2.2:** Determine bottleneck machine groups and their alternative machines for the saturation state.
>
> **Step 2.3:** Re-balance the workload among bottleneck machines and their alternative machines to obtain an improved throughput rate.

In Step 1, a total volume is divided into 10 parts, and distributed to five products, resulting in 30 possible scenarios of product mixes. Machine portfolio planning is a complex task that involves combinatorial optimisation, economic analysis, and multiple criteria (Chou and Wu 2002). In this study, a static capacity model, as compared to queuing capacity models, is assumed in Step 2.1. Bottleneck machines are identified by ranking machine groups by their utilisation and the top 10% are selected. Alternative machines of the bottlenecks are selected from machines of the same types as the bottlenecks but with lower utilisation. In Step 2.3, the workload among machine groups and their alternative machines is rebalanced. This requires solving two mixed integer linear programs in sequence. Subscripts j, k, i and b are used to index product, machine, step and bottleneck machine, respectively. Let $w_{i,j,k}$ be the standard machine time, $q_{i,j,k}$ be the allocated production quanity, C_k be capacity, and N_k be the number of machines. The decision variable $X_{(b,k)}$ refers to the fraction of the available time of machine b that is used to back up machine k. Re-balancing of workload involves three constraints:

$$\sum_j \sum_i w_{i,j,k} \cdot q_{i,j,k} \leq C_k N_k + \sum_{(b,k) \in BR} X_{(b,k)} C_b N_b - \sum_{(k,b) \in BR} X_{(k,b)} C_k N_k \quad \forall k$$

$$X_{(b,k)} \leq Y_{(b,k)} \quad \forall (b,k)$$

$$\sum_k q_{i',j,k} \geq (1+\alpha)D_j \quad \forall j, (i' \text{ is particular step with } w_{i',j,k} \neq 0 \text{ for some } k)$$

$$N_k \in N; \quad q_{j,k}, \alpha \in R_+; \quad 0 \leq X \leq 1; \quad Y \in \{0, 1\}.$$

The first constraint is a balance equation for workload and capacity. The set BR contains pairs of machine (k) and backup machine (b). In the second constraint, the Y variable is a 0–1 indicator variable. The third constraint is used to model the excessive workload α. (Throughput and capacity are treated as equivalent at the saturation state and full-load states.) The purpose of this first formulation is to determine if alternative routing is effective and a solution can be found. Of the 30 scenarios, some of them are very extreme. For example, one scenario has all demands allocated to one product. Alternative routing has no effect on some degenerate cases. The following objective function is used in the first stage.

$$\min Z_1 = \sum_{(b,k) \in BR} Y_{(b,k)}.$$

Table 1. Simulated throughput data of alternative routing.

Scenario (mix ratio)	Excessive workload α				
	0.00	0.01	0.02	0.03	0.04
4:3:2:1:0	22314	22540	–	–	–
4:3:1:1:1	22605	22803	–	–	–
4:2:2:2:0	24850	25082	25265	25351	–
4:2:2:1:1	22025	22278	–	–	–
3:3:3:1:0	20807	21284	–	–	–
3:3:2:2:0	24071	24395	24559	24681	24753
3:3:2:1:1	22014	22454	22821	23088	23150
3:2:2:2:1	24415	24842	25158	25340	25411
2:2:2:2:2	23441	23849	–	–	–

If solutions exist for the first formulation, a second objective function is used to maximise the throughput rate:

$$\max Z_2 = \sum_j \sum_k q_{i',j,k}.$$

Using $\tilde{d} = 22{,}000$ and solving each of the 30 scenarios, the throughput rates for corresponding α are shown in Table 1. Scenarios which do not yield any solutions for the first formulation are not shown.

3.2 *Procedure for constructing a full-load function*

Fitting a regression function directly to the data of Table 1 will only produce an *ad hoc* formula with restricted applicability. Such an approach does not take the saturation state and the control policy into consideration. We constructed the full-load function as follows. Because $\Omega = h(\alpha)$ is hinged on the saturation state, which is a phenomenon in the w–y frame, the full-load function has to be constructed in the same frame.

Since contingent resources are not costless and are likely to have diminishing returns, full-load functions will be represented as a second order polynomial: $y = f(w) = aw^2 + bw + c$ with shape properties $f'(w) \geq 0$ and $f''(w) < 0$ and boundary conditions $f(\tilde{w}) = \tilde{d}$ and $h'(\bar{\alpha}) \approx 0$. Making use of the constraining equality $w(\alpha) = \tilde{w} + \alpha\tilde{d}$, it can be derived that, for $\tilde{w} \leq w \leq w_u$, the coefficients of $f(w)$ are:

$$a = \frac{d - \tilde{d}}{w^2 - 2w_u w - \tilde{w}^2 + 2w_u \tilde{w}}; \quad b = -2aw_u; \quad c = \tilde{d} - a\tilde{w}^2 + 2aw_u\tilde{w} \qquad (2)$$

where $w_u = \tilde{w} + \bar{\alpha}\tilde{d} = \kappa\tilde{d} + \bar{\alpha}\tilde{d}$. The data in Table 1 can be used to estimate the value of parameter a and the values of b and c by using Equation (2). The resultant function is $d = f(w) = \hat{a}w^2 + \hat{b}w + \hat{c}$, where $\hat{a}, \hat{b}$ and $\hat{c}$ are estimates for parameters a, b and c. Finally, the full-load function $\Omega = h(\alpha)$ can be derived by making use of the equalities: $d = (1 + \Omega)\tilde{d}$
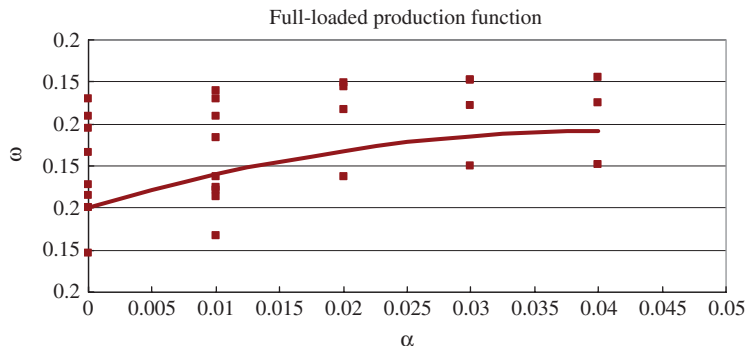
Figure 6. Full-load production function based on data of Table 1.

and $w = \tilde{w} + \alpha\tilde{d}$. It is obtained that

$$\tilde{d} + \tilde{d}\Omega = \hat{a} \cdot (\kappa\tilde{d} + \alpha\tilde{d})^2 - 2\hat{a}\hat{w}_u(\kappa\tilde{d} + \alpha\tilde{d}) + \tilde{d} - \hat{a}(\kappa\tilde{d})^2 + 2\hat{a}\hat{w}_u \cdot \kappa\tilde{d}$$
$$\Omega = h(\alpha) = \hat{a}\tilde{d}\alpha^2 - 2\hat{a}\tilde{d}\overline{\alpha}\alpha \quad 0 \le \alpha \le \bar{\alpha}. \tag{3}$$

The function dependent on $\tilde{d}$, $\bar{\alpha}$ and κ(via $\hat{a}$). Applying data in Table 1 results is $\Omega = -56.713\alpha^2 + 4.537\alpha$. The data of Table 1 and the full-load functional curve are plotted in Figure 6. It is worth noting that marginal throughput is greater than marginal capacity. This is a testimony to the effect of relieving the bottleneck.

4. A dynamic system model

A dynamic system model of Figure 4 can be represented by the following equations.

$$\dot{W}_1(t) = r_1(t) - \Omega_1(t) \tag{4}$$

$$\dot{W}_2(t) = \Omega_1(t) - \Omega_2(t) \tag{5}$$

$$r_1(t) = \bar{d}(t) + E(t) \tag{6}$$

$$\dot{I}(t) = \Omega_2(t) - \bar{d}(t) + \phi(t) \tag{7}$$

$$\Omega_k(t) = a_k\alpha_k^2 - e_k\alpha_k \quad k = 1, 2. \tag{8}$$

Equations (4)–(6) are rate balance equations. In Equation (7), the rate of change in channel inventory is equal to the difference between Ω_2 (t) and $\bar{d}(t)$, plus a noise $\phi(t)$. In nominal operation $\Omega_2(t)$ is driven by $\bar{d}(t)$; it should revert to $\bar{d}(t)$ in the long run. Equation (6) is a concave production function, in which the parameter $\alpha_k(t)$ is a control variable for activating flexible capacity.

Equation (6) can be substituted into Equation (4) and the production function (Equation (8)) can be substituted into Equations (4), (5) and (7). As discussed in Section 2,

the total throughput should equal to the total demand in the long run, that is, $\int (\Omega_2(t) - \bar{d}(t))dt$ should equal to 0. In the long run, excessive inventory should also be zero. Thus, $\int \dot{I}(t)dt = 0$ and $\int \phi(t)dt = 0$. In the short run, however, there will be excessive inventory, accumulated over certain duration $[0, t_0]$ so that $\int_0^{t_0} \phi(t)dt = s \neq 0$. This quantity s will enter the chain as an anticipatory input over a planning horizon of $[t_0, t_0 + T]$. After the substitution, the dynamic system (Equations (4–8)) can be reduced to:

$$\dot{W}_1(t) = \bar{d} + E(t) - [a_1\alpha_1(t)^2 - e_1\alpha_1(t)]. \tag{9}$$

$$\dot{W}_2(t) = [a_1\alpha_1(t)^2 - e_1\alpha_1(t)] - [a_2\alpha_2(t)^2 - 2e_2\alpha_2(t)]. \tag{10}$$

The anticipatory input s is a total amount and let $E(t)$ be the time path of the anticipated demand. Set the event time $t_0 = 0$ without loss of generality. $E(t)$ must satisfy the following equality integral constraint $\int_0^T E(t)dt = s$. Define a state variable $S(t)$ for the remaining amount of the anticipatory input. Therefore, $S(0) = s$, $S(T) = 0$ and

$$\dot{S}(t) = -E(t) \tag{11}$$

To manage dynamic events, an immediate objective is to solve this dynamic system for $\alpha_k(t)$ at the event time $t_0 = 0$, given $S(0) = s$. If a solution can be found, then it can be said that the impact of the dynamic event is manageable. If no solution can be found, then the dynamic event is disrupting and other means must be provisioned. This objective amounts to finding an admissible control trajectory of $\alpha_k(t)$ subject to some constraints on work-in-process or cumulative output.

The admissible question can be answered by solving an optimal control problem for the dynamic system. Equations (9) and (10) are equations of motion for W_k. Because activating flexible capacity will incur cost of machine setup, $\alpha_K(t)$ should not be a free variable.

4.1 *A formulation of optimal control*

A manufacturing chain of two nodes is used to describe a basic model of optimal control in this section. Let the state variable x be $[W_1, W_2, S]^T$ and the control variable u be $[\alpha_1, \alpha_2, E]^T$. We will start with the objective function of the basic model by considering the interest of both customers and manufacturers. Customers would be concerned with feasible $E(t)$ and the output of the chain. Whether an $E(t)$ path is feasible is determined by existence of solution $\alpha_k(t)$. Since output is determined by the total capacity in the full-load state, it can also be derived from $\alpha_K(t)$. Manufacturers are concerned with utilising flexible capacity and its associated cost. Thus, a general form of the objective function can be expressed as

$$V^* = \max_u \int \left[\sum_k -g_k(\alpha_k(t)) - h_k(W_k(t)) \right] - m_0 S dt$$

where g_k and h_k are continuous and differentiable cost functions and m_0 is a parameter. The mission of S is to track the depletion of s. It is assumed that the cost function g_k is increasing and convex and that the cost is zero when $\alpha_K(t) = 0$. Although high level of work-in-process is not desirable, its associated cost is not as severe as flexible capacity.

The cost function h_k is chosen to be linear in excessive work-in-process W_k. There is another justification for this design. As the objective functional contain two cost functions, h_k can be considered as a reference base for g_k. In this paper, g_k and h_k are represented by polynomial functions:

$$g_k(\alpha_k(t)) = b_k\alpha_k^2(t) + c_k\alpha_k(t) \tag{12}$$

$$h_k(W_k(t)) = m_k W_k(t) \tag{13}$$

where b_k, c_k, and m_k are coefficients, $g_k' \geq 0$ and $g_k'' \geq 0$. This implies that g_k is a convex function with a positive slope. Numerically, $2b_k\alpha_k + c_k \geq 0$ and $b_k \geq 0$. Expressing Equations (9–11) in a concise form, the equations of motion for state variables are:

$$\dot{W}_1(t) = f_1(x, u) \equiv \bar{d}(t) + E(t) - \left[a_1\alpha_1^2(t) - e_1\alpha_1(t)\right]$$
$$\dot{W}_2(t) = f_2(x, u) \equiv \left[a_1\alpha_1^2(t) - e_1\alpha_1(t)\right] - \left[a_2\alpha_2^2(t) - e_2\alpha_2(t)\right]$$
$$\dot{S}(t) = f_3(x, u) \equiv -E(t).$$

The Hamiltonian function is

$$H(x, u, \lambda, \mu) = \sum_k [-g(\alpha_k(t)) - h(W_k(t))] - m_0 S + \sum_{i=1}^{2} \lambda_i(t)f_i(x, u) + \mu(t)f_3(x, u)$$

where $\lambda_t(t)$ and $\mu(t)$ are co-state variables. The initial state of the system can be expressed as $[\tilde{w}_1, \tilde{w}_2, s]$ for given $\tilde{w}_1$ and $\tilde{w}_2$. The state variable $S(t)$ has a fixed terminal state; its transversality is replaced by the condition $S(T) = 0$ (Chiang 1992, p. 182). The equations of motion for co-state variables are

$$\dot{\lambda}_1(t) = -\nabla_{W_1(t)}H = m_1 \tag{14}$$

$$\dot{\lambda}_2(t) = -\nabla_{W_2(t)}H = m_2 \tag{15}$$

$$\dot{\mu}(t) = -\nabla_{S(t)}H = m_0. \tag{16}$$

Since the terminal conditions for W_k is unconstrained, the transversality condition $\lambda_i(T) = 0$ should hold. Solve for $\lambda_1(t)$, $\lambda_2(t)$, and $\mu(t)$ to obtain:

$$\lambda_1(t) = m_1(t - T)$$
$$\lambda_2(t) = m_2(t - T)$$
$$\mu(t) = m_0 t + k$$

where k is a constant. The transversality condition of $\mu(t)$ is no longer $\mu(T) = 0$ but is replaced by the fixed terminal condition of $S(T) = 0$.

We now turn to the maximum principle condition $\text{Max}_u\{H\}$. The function H has four terms. The first term is concave in α_K. The second and fourth terms are not

functions of α_K. The functional form of the third term depends on the relative magnitude of λ_1 and λ_2.

$$\frac{\partial H}{\partial \alpha_1} = -2b_1\alpha_1 - c_1 + (\lambda_2(t) - \lambda_1(t))(2a_1\alpha_1 - e_1)$$

$$\frac{\partial^2 H}{\partial \alpha_1^2} = -2b_1 + 2(\lambda_2(t) - \lambda_1(t))a_1 = -2b_1 + 2a_1(m_2 - m_1)(t - T).$$

By definition, $b_1 \geq 0$ and $a_1 \leq 0$. It is desirable to move the work-in-process downstream. If $m_1 \geq m_2$, $\partial^2 H/\partial\alpha_1^2 \leq 0$ and H is concave in α_K. (If $m_1 < m_2$, the convexity of H has to be further determined by the values of related coefficients.) To find the optimum, let $\partial H/\partial\alpha_1 = 0$.

$$-2b_1\alpha_1 + 2a_1(\lambda_2(t) - \lambda_1(t))\alpha_1 = c_1 + (\lambda_2(t) - \lambda_1(t))e_1$$

$$\alpha_1^*(t) = \frac{e_1(m_1 - m_2)(t - T) - c_1}{2a_1(m_1 - m_2)(t - T) + 2b_1}. \tag{17}$$

It can be seen that the optimal control path is dependent on the parameters of the production function, capacity cost and WIP cost. Similarly,

$$\frac{\partial H}{\partial \alpha_2} = -2b_2\alpha_2 - c_2 - \lambda_2(t)(2a_2\alpha_2 - e_2)$$

$$\frac{\partial^2 H}{\partial \alpha_2^2} = -2b_2 - 2a_2\lambda_2(t) = -2b_2 - 2a_2m_2(t - T)$$

As $a_2 < 0$, $b_2 \geq 0$ and $m_2 \geq 0$, it can be seen that $\partial^2 H/\partial\alpha_2^2 \leq 0$. Let $\partial H/\partial\alpha_2 = 0$.

$$[2b_2 + 2a_2\lambda_2(t)]\alpha_2 = 2e_2\lambda_2(t) - c_2$$

$$\alpha_2^*(t) = \frac{e_2m_2(t - T) - c_2}{2a_2m_2(t - T) + 2b_2}. \tag{18}$$

To find the optimal E path, take partial derivative of H.

$$\frac{\partial H}{\partial E} = \lambda_1(t) - \mu(t) = m_1(t - T) - (m_0t + k) = (m_1 - m_0)t - (m_1T + k) \quad t \in [0\ T].$$

The sign of $\partial H/\partial E$ is dependent on the relative magnitude of m_0 and m_1. Because m_0 and m_1 are inventory cost and it is desirable to move inventory downstream in manufacturing chains, we will assume $m_0 > m_1$. (If $m_0 < m_1$ on the contrary, the analysis will be similar but opposite results will be obtained.) The switching function of $\partial H/\partial E$ is determined by $t = (m_1T + \text{k})/(m_1 - m_0)$. When $t < (m_1T + \text{k})/(m_1 - m_0)$, it can be seen $(m_1 - m_0)t - (m_1T + k) > 0$ and H is linear in E with an upward-sloping curve. When $t > (m_1T + \text{k})/(m_1 - m_0)$, H is a downward-sloping curve. Since $S(0) = s$, $S(T) = 0$ and $\int_0^T E(t)dt = -[S(0) - S(T)] = s$, the value of k can be obtained by solving the equality (Appendix C):

$$\int_0^{(m_1T+k)/(m_1-m_0)} \bar{E}dt + \int_{(m_1T+k)/(m_1-m_0)}^T \underline{E}dt = \bar{E}\frac{(m_1T+k)}{(m_1-m_0)} + \underline{E}T - \underline{E}\frac{(m_1T+k)}{(m_1-m_0)} = s$$

in which $(m_1 T + k)/(m_1 - m_0) \le T$. Therefore,

$$k = \frac{(s - \underline{E}T)(m_1 - m_0)}{(\bar{E} - \underline{E})} - m_1 T, \text{ and by substituting } k, \ s \le \bar{E}T.$$

$$E^*(t) = \begin{cases} \bar{E} & t < \dfrac{(m_1 T + k)}{(m_1 - m_0)} = \dfrac{s - \underline{E}T}{\bar{E} - \underline{E}} \\[3mm] \underline{E} & t > \dfrac{(m_1 T + k)}{(m_1 - m_0)} = \dfrac{s - \underline{E}T}{\bar{E} - \underline{E}} \end{cases} \tag{19}$$

where $\underline{E}$ and $\bar{E}$ denotes the lower and upper limits of the domain of $E(t)$.

Substituting the control variables (Equations (17–19)) into the Hamiltonian function, the resultant new Hamiltonian function, renamed H^0, contains λ_i and W_k alone, with no control variables. It is straightforward to show that the new function H^0 has the following form (after substituting α_k^* and E^*):

$$H^0 = \tilde{g}(t) - m_1 W_1(t) - m_2 W_2(t) - m_0 S(t) + \lambda_1(t)\tilde{E}(t) + \lambda_2(t)\tilde{f}(t) + \mu(t)\tilde{E}(t).$$

where $\tilde{g}(t)$, $\tilde{E}(t)$, and $\tilde{f}(t)$ are functions of t only.

Since H^0 is concave in state variables for all t for given λ and μ, the maximum principle conditions are sufficient for the global maximisation of V, according to the Arrow sufficiency theorem (Chiang 1992, p. 182).

4.2 Numerical examples

Two manufacturing shops have a capacity to meet nominal demand $\tilde{d}_1 = 22{,}000$ and $\tilde{d}_2 = 22{,}500$, respectively. All demands are normalised and demand $\bar{d} = 0.0$. The full-load production functions, in which the throughput rate has been normalised by capacity, are

$$\Omega_1 = -50.562 \cdot \alpha_1^2 + 4.045 \cdot \alpha_1,$$
$$\Omega_2 = -51.711 \cdot \alpha_2^2 + 4.137 \cdot \alpha_2.$$

The upper limit of flexible capacity is $\bar{\alpha}_k = 0.04$ for $k = 1, 2$.

Assume initial inventory $[\tilde{w}_1, \tilde{w}_2] = [0.13, 0.08]$. Let $s = 0.12$, the planning horizon $T = 3$, and the domain of $E(t)$ is set as $[\underline{E}, \bar{E}] = [0.035, 0.045]$. The objective function is formulated as

$$\text{Max} \int_0^T \sum_{k=1}^2 [\Omega_k(t) - \alpha_k(t)] - 4W_1(t) - W_2(t) - 6S dt.$$
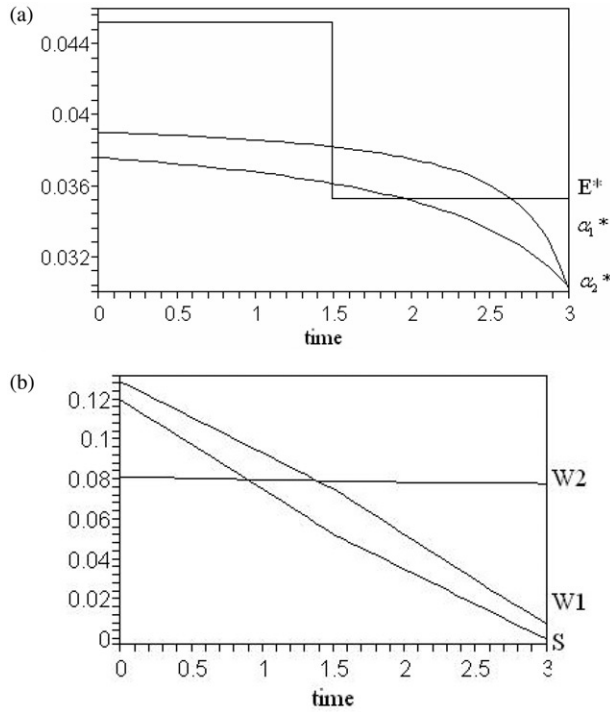
Figure 7. Contingent capacity and state paths (legends on the right axis). (a) Control variables; (b) state variables.

Thus, the optimal control paths and resultant state path are:

$$\alpha_1^*(t) = 0.04 - \frac{1}{1011.240 - 303.372t}$$

$$\alpha_2^*(t) = 0.04 - \frac{1.000}{413.688 - 103.422t}$$

$$E^*(t) = \begin{cases} 0.045 & t < 1.5 \\ 0.035 & t > 1.5 \end{cases}$$

$$S(t) = 0.12 - 0.045 \min(t, 1.5) - 0.035 \max(t - 1.5, 0).$$

The optimal control paths and the resultant state variables are plotted in Figure 7. Since an optimal path exists, it can be concluded that the channel inventory correction will not degrade the quality of supply services.

4.3 *Application scenario of the model*

In the era of globalisation and outsourcing, the manufacture of product and provision of services are increasingly carried out by chains of nodes. The contribution of this paper is on theoretical development of a model for risk management in such chains of production.

The model can be applied to many situations. In this section, an application scenario is described and procedure of application is outlined. The application scenario is based on Meirehofer and Steiner (2004) who describe an overload problem that arises in outsourcing projects. In many cases, the outsourcers are large enterprises and the insourcers are small and medium enterprises. A relatively small variation in workload at the outsourcers can easily overwhelm the insourcers, resulting in a breach of service level agreement. (This phenomenon is similar to what is described in this paper as demand shocks.) In overload situations, the insourcers must activate contingent capacity, which is normally more expensive, otherwise jobs will be delayed. (In this paper, alternative machines are modelled as contingent resources.) For outsourcers, there are business risk and financial consequences which usually dwarf the value of the outsourced jobs. Meirehofer and Steiner (2004) used a workload smoothing method to level out the phases of peak load and used a simulation program to evaluate the dynamic aspects of delay fluctuations. (The dynamic system model described in this paper has the same capability as the workload smoothing method and simulation program.)

The dynamic system model described in this paper can be applied to solving similar problems in sourcing projects at small production shops and more complex factories. There are four key points to follow in application:

(1) *Construct a production function model (for* Equation (8)): the production function is an aggregate model of the plant, not a detailed scheduling model. Its input is capacity (including base and contingent capacity) and its output is any performance measure of concern.
(2) *Determine an objective function:* the function is for maximising a chosen performance measure and can contain resource-use and/or job-related terms, by modelling them as control or state variables.
(3) The demand shock is modelled as the shock input $S(0)$.
(4) Solutions are given by algebraic Equations (17–19).

5. Discussion and conclusions

In this paper, a dynamic system model of supply chains is described which integrates the full-load production functions of production units with a shock demand input process. It addresses a type of supply chain problems in which dynamic events can not be satisfactorily modelled by probability distribution. It can be applied to proactively manage dynamic, disrupting events by assessing the impacts of the events and, if necessary, by activating contingency plans for their mitigation. Given an unordinary event of demand shock, this model can be used to determine if the shock can be absorbed by the manufacturing chains, without degrading its fulfilment services. It should be noted, however, that the model is not intended to be used to control the supply chains automatically. The intended use is to enhance the capability of managing disruption events so that they become deviation events. If the effect of a class of dynamic events can be mitigated, they will no longer be dynamic events. Instead, their remnant uncertainty will be at the level of deviations. If that happened, then we could say that supply chain control has been improved. In this paper, a novel production function is also proposed for analysing the full-load states of complex manufacturing systems which have prominent queuing phenomenon.

For different applications, the production function might have different input and output variables. Application of this model will require the production function to be first validated. In addition, there are usually multiple types of contingent resources, the saturation point will need to be validated for each application problem. Relevant validation for various applications is suitable future research work.

Acknowledgement

References

Blackhurst, J., Craighead, C.W., Elkins, D. and Hangfield, R.B., 2005. An empirically derived agenda of critical research issues for managing supply-chain disruptions. *International journal of production research*, 43, 4067–4081.

Bollapragada, R., Rao, U.S. and Zhang, J., 2004. Managing two-stage serial inventory systems under demand and supply uncertainty and customer service level requirements. *IIE transactions*, 36, 73–85.

Chiang, A.C., 1992, *Elements of Dynamic Optimisation,* New York, NY: McGraw-Hill.

Chou, Y.-C., Cheng, C.-T., Yang, F.-C. and Liang, Y.-Y., 2007. Evaluating alternative capacity strategies in semiconductor manufacturing under uncertain demand, price scenarios. *International journal of production economics*, 105, 591–606.

Connors, D.P. and Yao, D.D., 1996. Methods for job configuration in semiconductor manufacturing. *IEEE transactions semiconductor manufacturing*, 9, 401–411.

Delp, D., Si, J., Hwang, Y., Pei, B. and Fowler, J., 2005. Availability-adjusted X-factor. *International journal of production research*, 43, 3933–3953.

Gaonkar R., and Viswanadham, N., 2004. A conceptual and analytical framework for the management of risk in supply chains, in *Proceedings of 2004 IEEE International Conference on Robotics & Automation*, pp. 2699–2704.

Graves, S.C., Kletter, D.B. and Hetzel, W.B., 1998. A dynamic model for requirement planning with application to supply chain optimization. *Operational research*, 46, S35–S49.

Holton, G.A., 2004. Defining risk. *Financial analysis journal*, 60, 19–25.

Johri, P.K., 1993. Practical issues in scheduling and dispatching in semiconductor wafer fabrication. *Jounal manufacturing system*, 12, 474–485.

Jüttner, U., Peck, H. and Christopher, M., 2003. Supply chain risk management: outlining an agenda for future research. *International journal logistics: research and application*, 6, 197–210.

Jüttner, U., 2005. Supply chain risk management: understanding the business requirements from a practitioner perspective. *International journal logistics management*, 16, 120–141.

Karmarkar, U.S., 1989. Capacity loading and release planning with work-in-progress (WIP) and leadtimes. *Journal manufacturing operational management*, 2, 105–123.

Kleinforfer, P.R. and Saad, G.H., 2005. Managing disruption risks in supply chains. *Production operational management*, 14, 53–68.

Knight, F.H., 1921, *Risk, Uncertainty and profit,* Boston, MA: Houghton Mifflin.

Lee, H.L. and Yano, C.A., 1988. Production control in multistage systems with variable yield losses. *Operational research*, 36, 269–278.

LeRoy, S.F. and Singell Jr, L.D., 1987. Knight on risk and uncertainty. *Journal of political economy*, 95, 394–406.

Lu, S.C.H., Ramaswamy, D. and Kumar, P.R., 1994. Efficient scheduling policies to reduce mean and variance of cycle-time in semiconductor manufacturing plants. *IEEE transactions semiconductor manufacturing*, 7, 373–388.

Meierhofer, J. and Steiner, A., 2004. Quantitative assessment and capacity planning of business processes in the service sector. *Proceedings of the IPLnet 2004 Workshop*, Murten, Switzerland, 7–8 September 2004.

Missbauer, H., 2002. Aggregate order release planning for time-varying demand. *International journal of production research*, 40, 699–718.

Ni, M., Luh, P.B. and Xiong, B., 2003. An inventory control policy for maintenance networks with advanced demand information, *2003 Proceedings of IT-MLS*, pp. 15–16.

Runde, J., 1998. Clarifying Frank Knight's discussion of the meaning of risk and uncertainty. *Cambridge journal of economic*, 22, 539–546.

Sheffi, Y., 2001. Supply chain management under the threat of international terrorism. *International journal logistics management*, 12, 1–11.

Tang, C.S., 2006. Robust strategies for mitigating supply chain disruptions. *International journal logistics: research and application*, 9, 33–45.

Vorst, J.G.A.J. and Beulens, A.J.M., 2002. Identifying sources of uncertainty to generate supply chain redesign strategies. *International journal of physics distributed logistics management*, 32, 409–430.

Appendix A: channel inventory fluctuation

In the semiconductor supply chains, channel inventory includes those stockpiles of microchips in foundry manufacturers, IC product companies, electronic manufacturing service companies and distributors. Market intelligence firms actively monitor the channel inventory. Figure 8 shows an example of the excessive inventory data published by iSuppli® over a duration of two years. (Excess inventory is the inventory in excess of the normal seasonal inventory level.) A closer examination of the data will reveal the impact of channel inventory information on the supply chain production. The average excess inventory was US\$1.6 billion in Q3 of 2004 and it was predicted at that time that the inventory will decrease to US\$1.5 billion in the following quarter. However, in Q4 of 2004 the excess inventory turned out to be US\$1.0 billion. Given that the total revenue was approximately US\$200 billion a year, if the inventory adjustment of US\$1.5 billion were to be completed in one month, the impact on the supply chain would amount to a capacity impact of 9% ($12/200*1.5$) at the aggregate
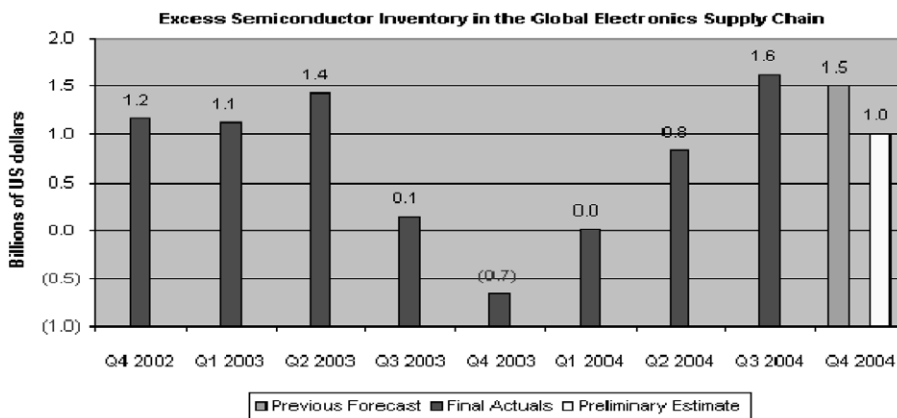

Figure 8.   Worldwide channel inventories (compiled in this study).

level. Considering that aggregate demand is less volatile than individual demands, the figure provides some evidence on the potentially disrupting effect of channel inventory fluctuation to supply chain production.

Appendix B: supply chain configuration altered by industrial accident

(For brevity, acronyms of firms are used. Their identity can be easily discerned.) In the semiconductor IC packaging and testing industry, ASE and SPIL are two competing conglomerates, both ranked in the top four globally. Both firms have many subsidiaries that supply key parts or subsystems to other plants and/or firms. PPT is a subsidiary of SPIL. The Chairperson of the Board (COB) of SPIL also acts as the COB of PPT. On 1 May 2005, a large fire burned down a major facility of an ASE subsidiary which manufactures several key parts (substrate PCB boards) for IC packaging. For certain part types, complete capacity of the conglomerate was lost. Ensuing negotiation between ASE and SPIL resulted in PPT becoming an independent firm. The agreement called for the COB of SPIL to resign from the position of COB of PPT. PPT then became a supplier of ASE, beside SPIL. This industry accident not only disrupted normal production and supply at many firms in the supply chains in the short-medium term but also altered long-term capacity expansion plans, chain organisation and several other joint venture relationships.

Appendix C: dynamic optimisation

Key steps of the optimisation are described in this appendix to supplement section 4.1.

$$\frac{\partial H}{\partial \alpha_1} = -\frac{\partial g(\alpha_1(t))}{\partial \alpha_1} - (\lambda_1(t) - \lambda_2(t))\frac{\partial \Omega_1(t)}{\partial \alpha_1} = -2b_1\alpha_1 - c_1 - (\lambda_1(t) - \lambda_2(t))(2a_1\alpha_1 - e_1)$$

To find the optimum, let $\partial H / \partial \alpha_1 = 0$.

$$[2b_1 + 2a_1(\lambda_1(t) - \lambda_2(t))]\alpha_1 = (\lambda_1(t) - \lambda_2(t))e_1 - c_1,$$

By substituting the solution of $\lambda_1(t)$ and $\lambda_2(t)$,

$$\alpha_1^*(t) = \frac{e_1(m_1 - m_2)(t - T) - c_1}{2a_1(m_1 - m_2)(t - T) + 2b_1}.$$

Similarly,

$$\frac{\partial H}{\partial \alpha_2} = -\frac{\partial g(\alpha_2(t))}{\partial \alpha_2} - \lambda_2(t)\frac{\partial \Omega_2(t)}{\partial \alpha_2} = -2b_2\alpha_2 - c_2 - \lambda_2(t)(2a_2\alpha_2 - e_2) = 0$$

$$[2b_2 + 2a_2\lambda_2(t)]\alpha_2 = 2e_2\lambda_2(t) - c_2 \text{ where } \lambda_2(t) = m_2(t - T)$$

$$\alpha_2^*(t) = \frac{e_2 m_2(t - T) - c_2}{2a_2 m_2(t - T) + 2b_2}.$$

Since the switching function of $\partial H/\partial E$ is $t = (m_1 T + k)/(m_1 - m_0)$, and $\int_0^T E(t)dt = s$, the value of k can be obtained by solving the equality:

$$\int_0^{(m_1 T + k)/(m_1 - m_0)} \bar{E} dt + \int_{(m_1 T + k)/(m_1 - m_0)}^T \underline{E} dt = \bar{E}\frac{(m_1 T + k)}{(m_1 - m_0)} + \underline{E}T - \underline{E}\frac{(m_1 T + k)}{(m_1 - m_0)}$$

$$= \frac{(m_1 T + k)\bar{E} + \underline{E}T(m_1 - m_0) - \underline{E}(m_1 T + k)}{(m_1 - m_0)}$$

$$= \frac{m_1 T\bar{E} - \underline{E}Tm_0 + k\bar{E} - k\underline{E}}{(m_1 - m_0)}$$

$$= \frac{m_1 T\bar{E} - \underline{E}Tm_0}{(m_1 - m_0)} + \frac{k(\bar{E} - \underline{E})}{(m_1 - m_0)} = s$$

$$k = (s - \frac{m_1 T\bar{E} - \underline{E}Tm_0}{(m_1 - m_0)})\frac{(m_1 - m_0)}{(\bar{E} - \underline{E})}$$

$$= \frac{s(m_1 - m_0) - m_1 T\bar{E} + \underline{E}Tm_0 - \underline{E}Tm_1 + \underline{E}Tm_1}{(m_1 - m_0)}\frac{(m_1 - m_0)}{(\bar{E} - \underline{E})}$$

$$= \frac{s(m_1 - m_0) - \underline{E}T(m_1 - m_0) - m_1 T(\bar{E} - \underline{E})}{(\bar{E} - \underline{E})} = \frac{(s - \underline{E}T)(m_1 - m_0)}{(\bar{E} - \underline{E})} - m_1 T.$$

The switching point for $E^*(t)$ is given by:

$$\frac{(m_1 T + k)}{(m_1 - m_0)} = \frac{1}{(m_1 - m_0)}\left(m_1 T + \frac{(s - \underline{E}T)(m_1 - m_0)}{(\bar{E} - \underline{E})} - m_1 T\right) = \frac{(s - \underline{E}T)}{(\bar{E} - \underline{E})}.$$