Presented at the 2002 NSF-DMII Grantees Conference, San Juan, Puerto Rico January 7-11, 2002

A Model Predictive Control Framework For Robust Management Of Multi-Product, Multi-Echelon Demand Networks

Braun, Martin W.[1], Rivera, Daniel E.[1], Carlyle, W. Matt[2], Kempf, Karl G.[3]

[1]Department of Chemical and Materials Engineering,

Control Systems Engineering Laboratory, Arizona State University,

Tempe, AZ, USA 85287-6006

[2]Department of Industrial Engineering, Arizona State University,

Tempe, AZ, USA 85287-5906

[3]Decision Technologies, Intel Corporation,

5000 W. Chandler Blvd.,

Chandler, AZ, USA 85226

Abstract

Model Predictive Control (MPC) is shown to provide a robust, flexible framework to stabilize inventories yet still meet customer demand with minimal safety stock. The translation of available information in the supply chain problem into MPC variables is demonstrated with a two-node supply chain example. A six-node, two-product, three-echelon demand network simulation proposed by Intel Corp. is well managed by the framework under realistic demand and information inaccuracies.

1. INTRODUCTION

Recent literature suggests that billions of dollars in cost reductions can be achieved through improved management of semiconductor supply chains [4,8]. Semiconductor demand networks are particularly challenging since lead times for products often range on the order of months, and safety stock levels that cover as much as year's worth have demand have traditionally been held [5]. In this paper, an approach using Model Predictive Control is proposed as means to manage supply chains in a robust manner in the presence of inaccurate production leadtime and demand forecast estimates. The MPC framework is shown to track a qualitatively realistic demand pattern in spite of these inaccuracies, and make use of safety stock levels that are well below levels suggested by industry heuristics.

A supply chain (a.k.a. demand network, or value web) consists of the interconnected components required to transform ideas and raw materials into delivered products and services. The entire structure is organized and managed with the goal of maintaining a high level of customer service, while minimizing costs and maximizing profits. Companies no longer compete against other companies. Instead, supply chains compete against other supply chains. The supply chain that gains market share does so by providing customers the right product, in the right amount, at the right time, for the right price, and at the right place [1,4].

Recently, work utilizing Model Predictive Control has been found to provide an attractive alternative for inventory control [9], and supply chain management [2,6]. These approaches are conceptually different and require less detailed knowledge in comparison with cost-optimal stochastic programming solutions which require many “what-if” cases to be run and examined by highly skilled professionals [3]. Yet MPC offers the same flexibility in terms of the information sharing, network topology, and constraints that can be handled.

In the next section of this paper, a translation of the available information in a supply chain setting to process control variables is demonstrated to provide well behaved management of a two-node supply chain. In the third section of the paper, a six-node, two-product, three-echelon network developed to mimic the back end configuration of a semiconductor chain is robustly managed under realistic information inaccuracies. The paper is brought to a close with a summary of the ideas presented.

2. MPC FRAMEWORK

A. Model Predictive Control

Model Predictive Control has long been the preferred algorithm for robust, multivariable control in the process industries, with the number of implementations numbering in the thousands. The popularity of MPC stems from the relative ease with which it can be understood, and its ability to handle input and output constraints [7]. The objective function of an MPC controller can be written as

$$J = \sum_{l=1}^{p} Q_e(l)(\hat{y}(k+l \mid k) - r(k+l))^2$$

$$+ \sum_{l=1}^{m} Q_{\Delta u}(l)(\Delta u(k+l-1 \mid k))^2$$

$$+ \sum_{l=1}^{m} Q_u(l)(u(k+l-1 \mid k) - u_{\text{target}}(k+l-1 \mid k))^2$$

The three terms in the MPC cost function penalize predicted setpoint tracking error, excess movement of the manipulated variable, and deviation of the manipulated variable from a target value, respectively. The MPC optimization problem can be written

$$\min_{\Delta u(k|k)\ldots\Delta u(k+m-1|k)} J$$

s. t.

$$u_{\min} \le u(k+l-1 \mid k) \le u_{\max},$$
$$\Delta u_{\min} \le u(k+l-1 \mid k) \le \Delta u_{\max}$$

The problem can be modified to include output constraints as well. The optimization problem is readily solved by standard quadratic programming (QP) algorithms. Only the first control element of the solution is implemented. At the next time step the optimization problem is solved again with updated information from the system. This is referred to as the receding-horizon property of MPC as illustrated in Figure 1. Note that the MPC controller explicitly uses a model relating the inputs and measured disturbances to the outputs.

B. MPC Problem Definition

In this section, the available information in a supply chain setting is re-classified in a process control sense. This classification assigns aspects of the demand network problem to process control variables as shown in Table 1. Figure 2 illustrates the material flows from the Factory to the Retailer and on to the Customer. Both nodes are modeled with a mass balance

$$I_A(k+1) = P_A(k) - S_{2A}(k) + I_A(k),$$

where

$$P_A(k) = S_{1A}(k-\theta_p)$$

I_A is the inventory of A; S_{1A} is the incoming stream of A; S_{2A} is the outgoing stream of A; θ_p represents the processing delay; $P_A(k)$ is the material which has completed processing at time k of species A. Figure 3 provides a graphical representation of the information transfer between controllers and nodes of the two-node system for the MPC configuration.

Current customer demand is fed directly to the retailer and the retailer can immediately fill that demand that day. It is assumed that a demand forecast is known although it may be have biased or random error as dictated by the simulation conditions. The demand (measured disturbance) and demand forecasts (estimated future measured disturbances) are fed to the first echelon MPC controller. Using the current inventory (outputs) information from the retailer, the first echelon MPC controller decides what orders (inputs) for product A should be placed with the factory, and what the order forecast (estimated inputs) will look like. This order forecast is shared with the second echelon MPC controller.

The second echelon MPC controller uses the order forecast (now an estimated future measured disturbance) from the first echelon MPC controller, and the inventory information (outputs) from the factory to decide on production starts (inputs) for the day. Both MPC controllers contain models that determine the effect that orders (measured disturbances) from downstream entities have on the future inventories (estimated outputs) in their node. This model also relates orders to the factory (orders for the first echelon MPC controller) and production starts (inputs for the second echelon controller) to the inventory levels (outputs). The inventory targets (setpoint trajectories) are a forward time shifted version of the estimated future measured disturbances (plus safety stock), for the first echelon controller. For the factory, the inventory targets are an exact replication (plus safety stock), of the estimated measured disturbances for the factory since there is no direct feed through (i.e. orders placed today are only on backorder if not filled tomorrow).

Table 1 Variable mapping for MPC controllers

Process Control Variable	Demand Network Information
setpoints r	inventory targets of species A
outputs y	inventories of species A minus cumulative outstanding backorders
estimated outputs $\hat{y}$	forecasted inventories of species A
measured disturbances u_d	demand or orders for species A being placed at the node
estimated future measured disturbances $\hat{u}_d$	forecasted demand or orders for species A being placed at the node
estimated inputs $\hat{u}$	forecasted orders for species A being placed at the upstream node
inputs u	orders for species A being placed at the upstream node

The MPC framework is evaluated with the two-node system under plant-model mismatch. The production leadtime in the Factory is actually 3 units, while the Factory MPC controller is implemented with a production leadtime of 2 units. The Retailer MPC controller is implemented as if there was no shipping delay between the Factory and the Retailer, however in the simulation there is 1 unit of delay. There is no production delay in the Retailer for either simulation or controller. First the controllers are implemented with no move suppression. The time series and metrics for the Retailer are shown in Figure 4. Inventory levels approach unthinkable levels, and customer service could be better. Using move suppression values of 150 for both controllers, the system is stabilized, and the 1500 units of safety stock are sufficient to eliminate backorders as shown in Figure 5 and Figure 6.

C. Performance Under Plant-Model Mismatch And Biased Forecast Error

The performance and robustness of the MPC control system will now be demonstrated on a six node network simulated using realistic plant-model mismatch for the Assembly/Test Nodes, biased forecast error, and a realistic demand profile. These conditions were recommended for evaluation by Intel Corp. on a six-node, two-product, three echelon simulation shown in Figure 7. The simulation mimics the packaging, distribution and retail sale of semiconductor products, in each of the three echelons.

Experience with the actual performance of Assembly/Test nodes and the estimated lead times by facility personnel, suggest that the facility personnel traditionally provide themselves a lead time buffer of one day. So for example, if in reality the process takes 9 days to complete, a 10 day lead time estimate will be quoted to others in the organization. Thus, a 9 day actual/10 day estimate plant-model mismatch is adopted for both Assembly/Test facilities. The Sales and Marketing personnel have been generally known to determine forecasts which are biased in an optimistic manner. As an example, the sales forecast for the next time period might be 11,000 units, when in fact the actual sales will be 10,000 units. To mimic this type of forecast bias, all demand forecasts passed to the Retailer level MPC controller are biased by +1000 units.

Lastly, products have been observed to follow demand patterns which may be correlated at times, and uncorrelated at other times. To mimic this type of behavior, the demand patterns for Product A and Product B follow correlated, deterministic steps up until time 110. The remainder of the time, the demand patterns remain uncorrelated. This behavior can be observed in Figure 9.

The MPC controllers for each echelon of the network, have to be modified slightly to account for the cross-shipment routes that can occur between echelons. The problem inputs now number eight, since there are two shipping lanes from each node and each lane can transport either or both products. To handle the extra degrees of freedom this brings to the problem, the penalties Q_u for the cross-shipment routes are purposefully set high, so these routes are not favored in the cost function. This penalty also allows for a unique solution to the MPC problem. The information flows for the six-node problem under MPC supervision are found in Figure 8.

The experiment is run with a value of 300 for all move suppression parameters $Q_{\Delta u}$, 0 for all penalties for values other than zero for direct shipment Q_u, and a penalty of 100 for all penalties for values other than zero for cross-shipments Q_u. Table 2 holds the prediction horizons N_p and control horizons N_u for all three MPC controllers. These horizon parameters are used for every six node management experiment in this document. Safety stock is set at 5,000 units per product, per node. All inventory control error weights Q_e are left at 1. All entities use a "pecking order" dispatch rule. The orders of Warehouse 1, Retailer 1, and Customer 1 take precedence over the orders of the corresponding counterparts. Figure 10 and Figure 11 demonstrate the performance of this approach with plots of inventories, demands, shipments, and factory starts.

Table 2 Prediction and control horizons used in six node controllers

Parameter	Echelon #1	Echelon #2	Echelon #3
N_p	49	41	33
N_u	42	34	12

At time 1, the Retailer MPC controller adjusts orders and order forecasts to the upstream nodes to start bringing in more product, since the demand forecast is now 11,000 even though the actual amounts being demanded are 10,000. Because of the move suppression, the increases in order amounts are less than 1,000. The increase in orders is evident in the increase in direct and cross shipments from the Warehouse echelon. Soon the Retailer MPC echelon realizes the actual amount supplied to customers is *not* increasing as suggested by the forecast and the Retailer MPC adjusts to account for the forecast error and reduce inventory.

In the first few time units, inventories in the Factory echelon and Warehouse echelon are drained below their target levels. The Factory echelon MPC controller now observes the effects of the plant-modelmismatch, since changes in starts show up sooner than expected. The inventory levels of the Assembly/Test nodes fluctuates, but the fluctuation remains at reasonable levels. No backorders take place throughout the entire experiment. This is rather impressive, since the general rule of thumb practiced for this network requires a safety stock level equivalent to between two and four times the expected demand for the next time period (e.g. if tomorrow's demand is expected to be 10,000 units, safety stock held today may range from 20,000 to 40,000 units).

Note that the cross-shipments in this example are used whenever there are rapid changes in the order/demand forecasts. The MPC controllers make use of the cross-shipments, since the costs associated with the move suppression weightings of the direct shipments become comparatively large at these times. This may make sense not only from an optimization standpoint, but in a realistic setting it may also allow nodes to hedge against uncertainties or disturbances in transportation links or nodes connected by the direct shipment lanes.

3. CONCLUSIONS

By using a systems/controls perspective, it is possible to stabilize supply chains in the presence of data inaccuracies or plant-model mismatch. The combination of move suppression and shared order forecasts allow nodes in the supply chain to appropriately buffer the inventory before large changes in demand take place at the retailer. An MPC framework can even handle demand networks of size and topology relevant to industrial needs as demonstrated through the successful management of the problem proposed by Intel Corp.

4. ACKNOWLEDGMENTS

This research has been funded through the DMI Award 0075682, “Scalable Enterprise Systems: Designing and Managing Dynamic Supply Chains Using Model-on-Demand Predictive Control” and an Intel Research Council grant entitled, “A Modular, Scalable Approach to Modeling and Analysis of Semiconductor Manufacturing Supply Networks.”

5. REFERENCES

1. Bodington, C. E. and D. E. Shobrys (1999). Optimize the supply chain. In: *Advanced Process Control and Information Systems for the Process Industries* (L. A. Kane, Ed.). pp. 236--240. Gulf Publishing Company. Houston, TX.

2. Flores, M. E., D. E. Rivera and V. Smith-Daniels (2000). Managing supply chains using model predictive control. In: *AIChE 2000 Annual Meeting, Nov. 12--17*. Los Angeles, CA. paper 262f..

3. Kafoglis, C. C. (1999). Maximize competitive advantage with a supply chain vision. *Hydrocarbon Processing* pp. 47--50.

4. Kempf, K., K. Knutson, B. Armbruster, P. Babu and B. Duarte (2001). Fast accurate simulation of physical flows in demand networks. In: *Advanced Simulation Technologies Conference*. The Society for Computer Simulation International. Seattle, WA.

5. Lee, H. L., V. Padmanabhan and S. Whang (1997). The bullwhip effect in supply chains. *Sloan Management Review* pp. 93--102..

6. Perea-Lopez, E., I. E. Grossman and B. E. Ydstie (2000). Application of mpc to decentralized dynamic models for supply chain management. In: *AIChE 2000 Annual Meeting, Nov. 12--17*. Los Angeles, CA. paper 262e.

7. Prett, C. E. Garcia D. M. and M. Morari (1989). Model predictive control: Theory and practice-a survey. *Automatica* **25**(3), 335--348.

8. PriceWaterhouseCoopers (2000). Silicon 2000: Millennium semiconductor survey on supply chain management. http://www.ebusiness.pwcglobal.com/external/genreg.nsf/semi?OpenForm\&Seq=1.

9. Tzafestas, S., G. Kapsiotis and E. Kyriannakis (1997). Model-based predictive control for generalized production planning problems. Computers in Industry **34**(2), 201--210.

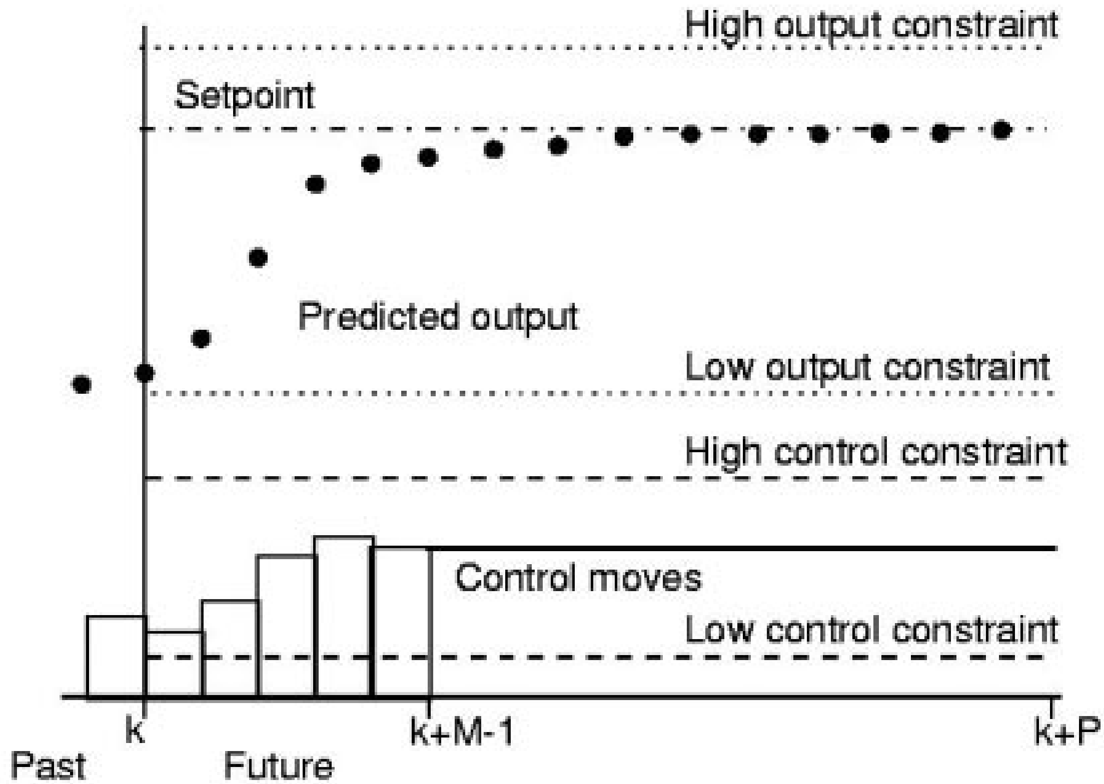

Figure 1 MPC receding horizon philosophy

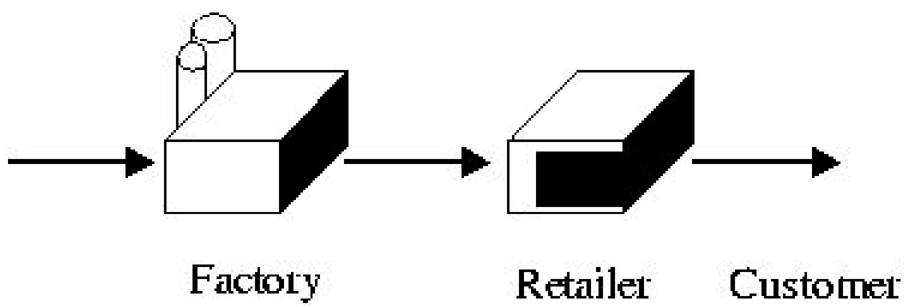

Figure 2 Two-node network material flows

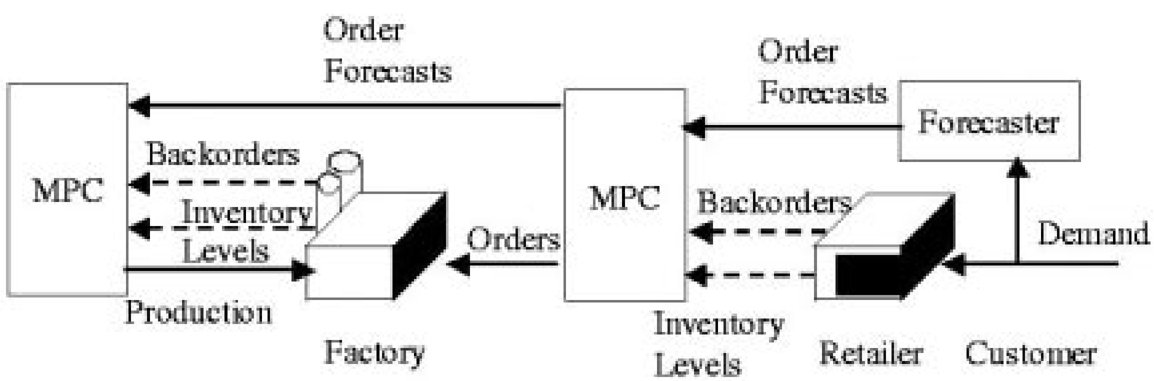

Figure 3 Two-node network MPC information flows

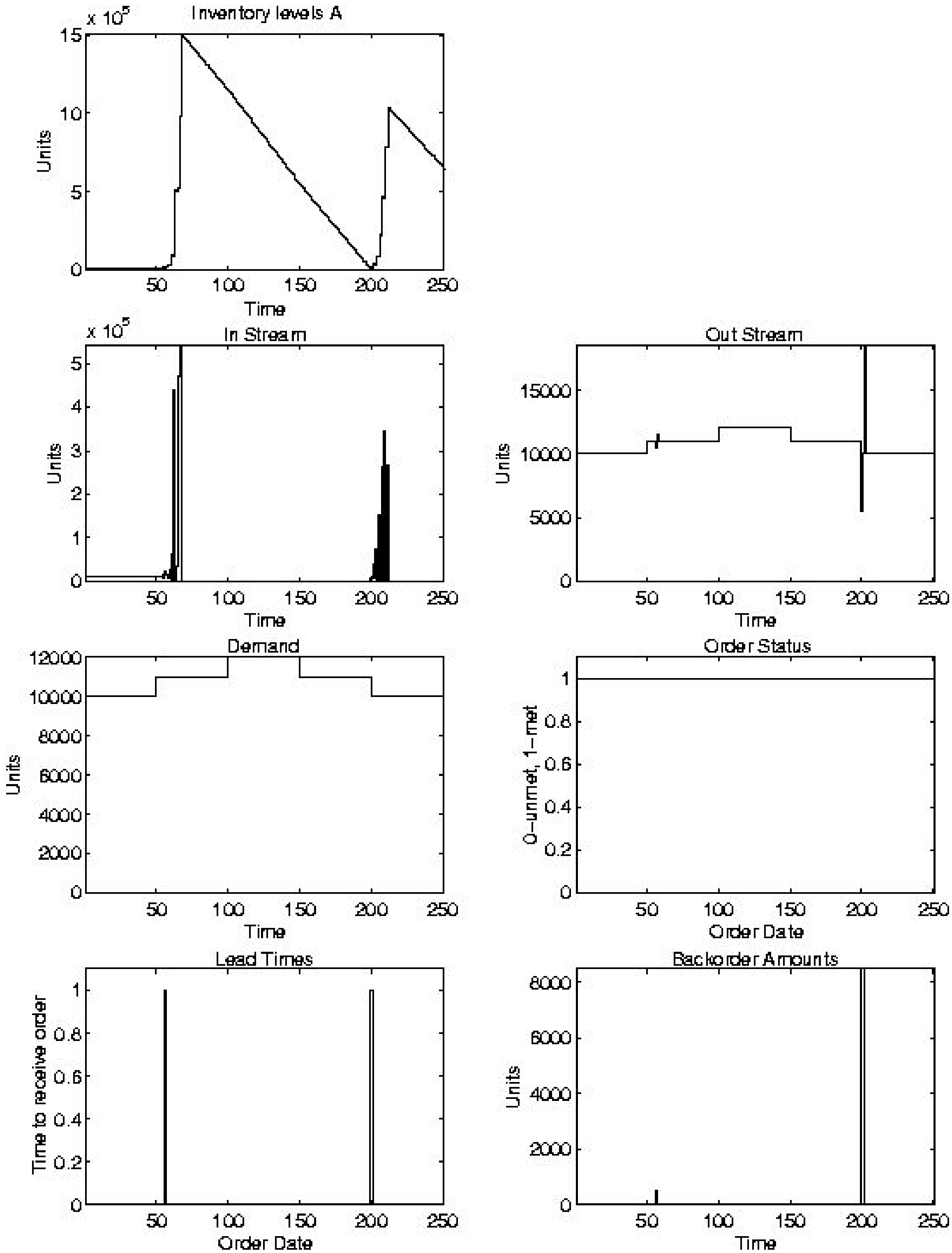

Figure 4 Retailer Responses and Metrics, Two-Node Example, No Move Suppression

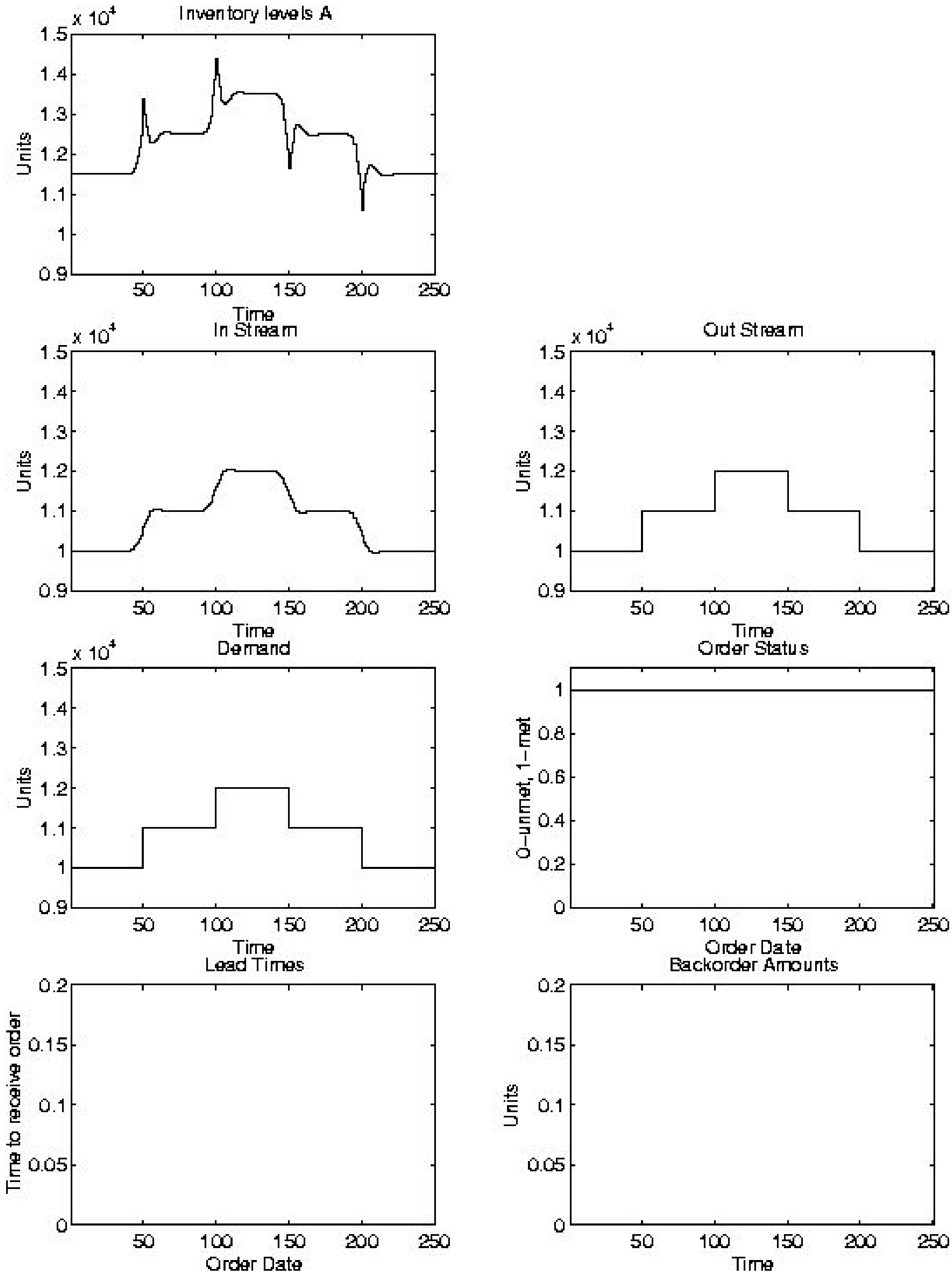

Figure 5 Retailer Responses and Metrics, Two-Node Example, With Move Suppression

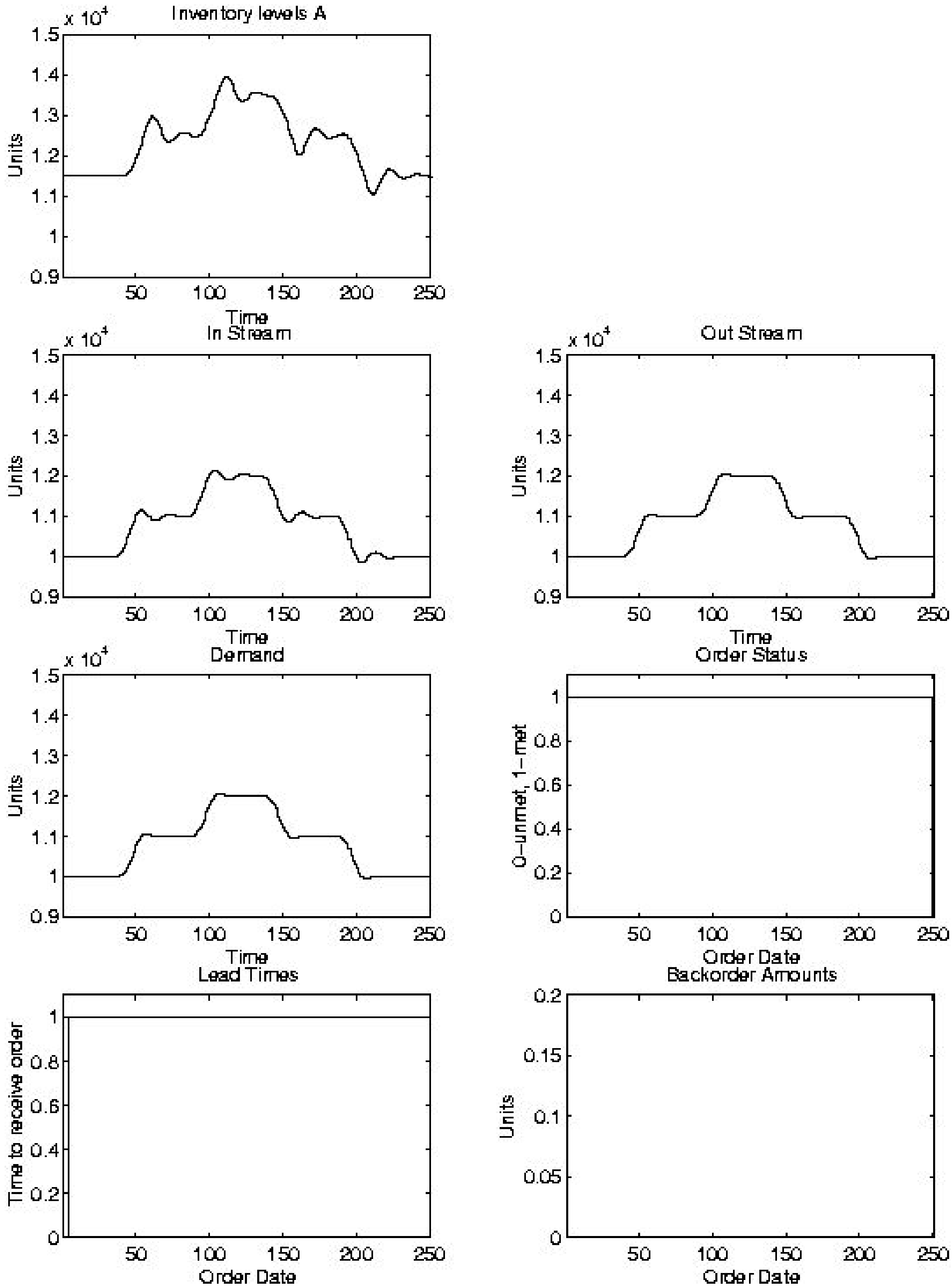

Figure 6 Factory Responses and Metrics, Two-Node Example, With Move Suppression

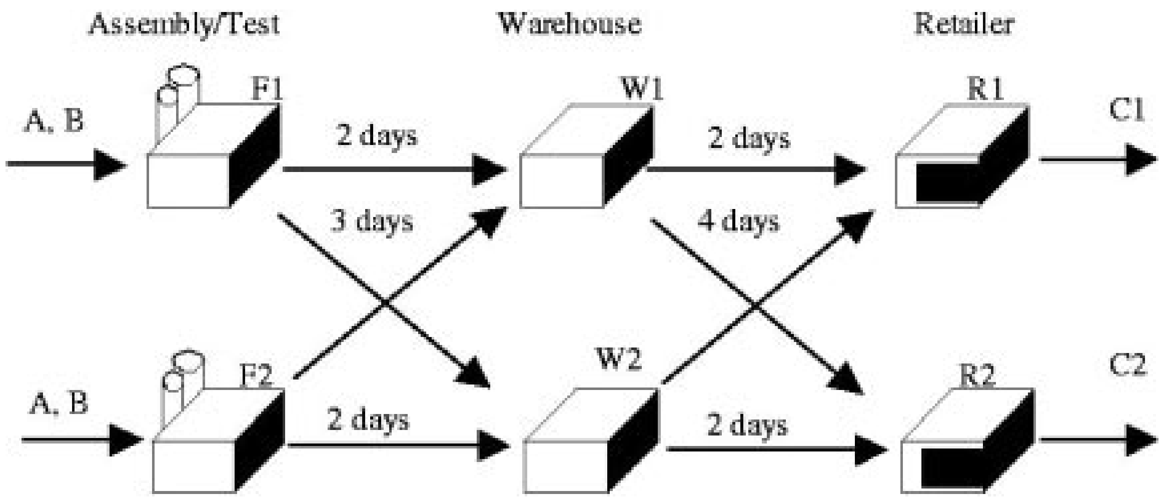

Figure 7 Six-node network material flow

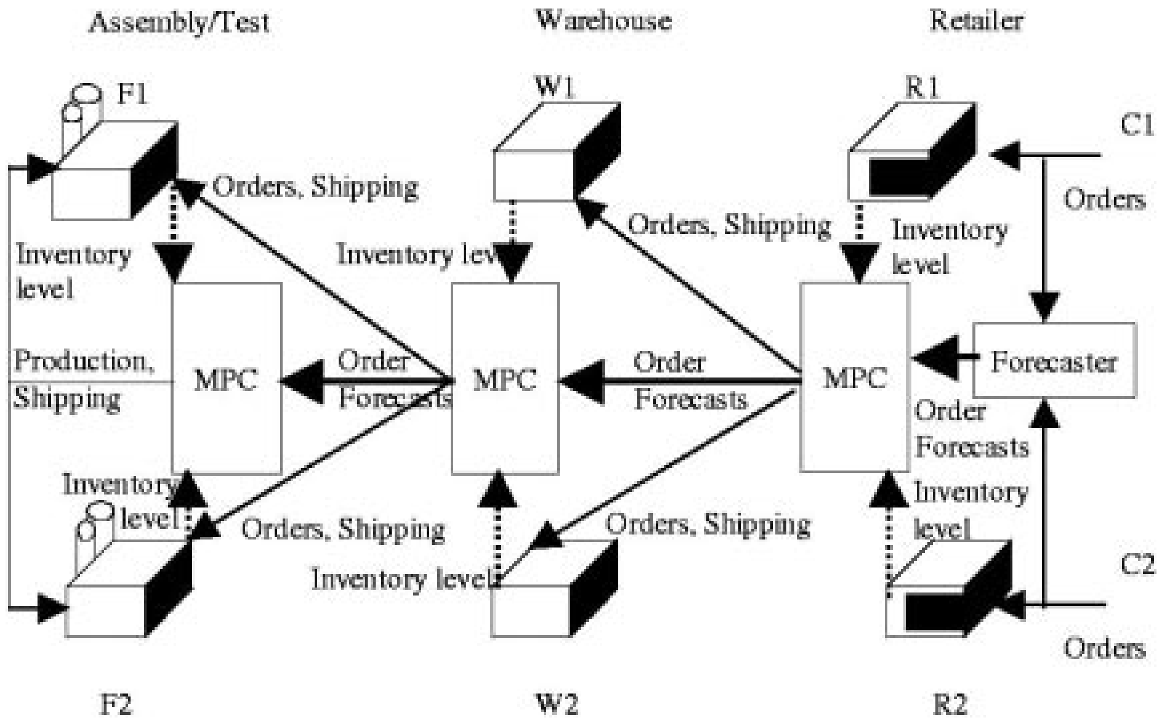

Figure 8 Information flow for management of the six-node Intel network

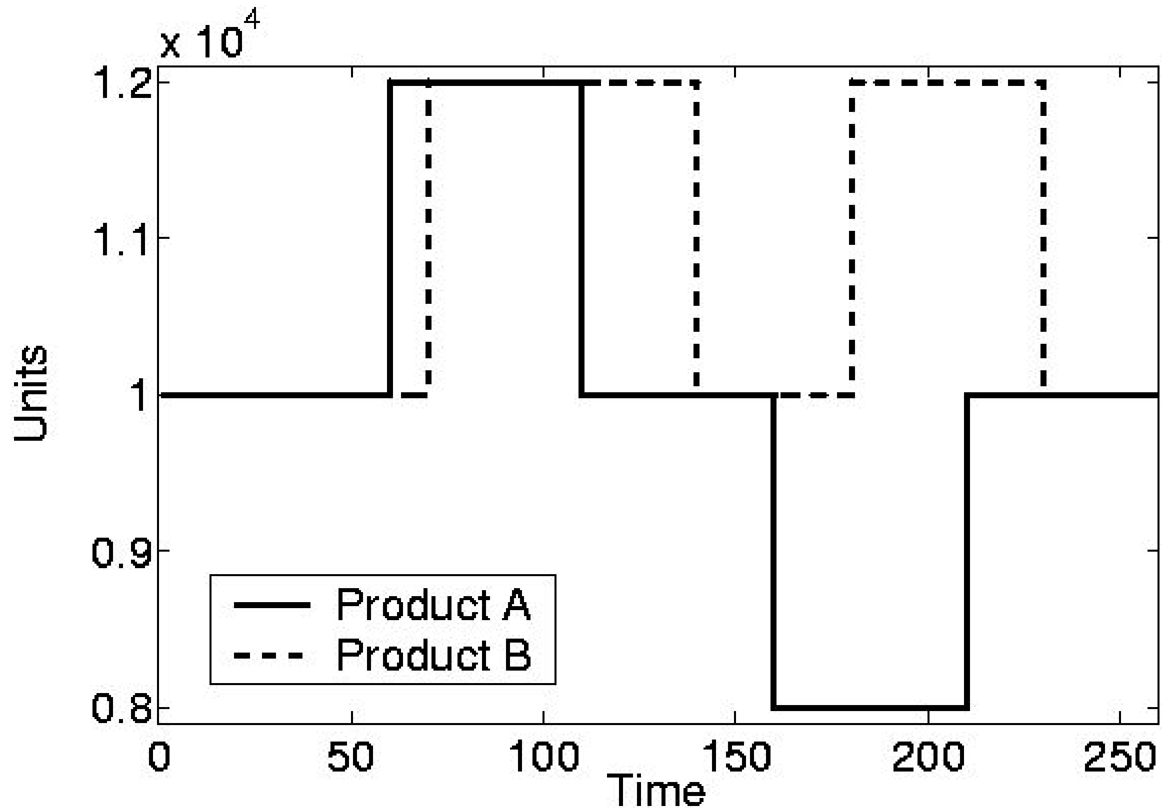

Figure 9 Qualitatively realistic demand profile for products A and B

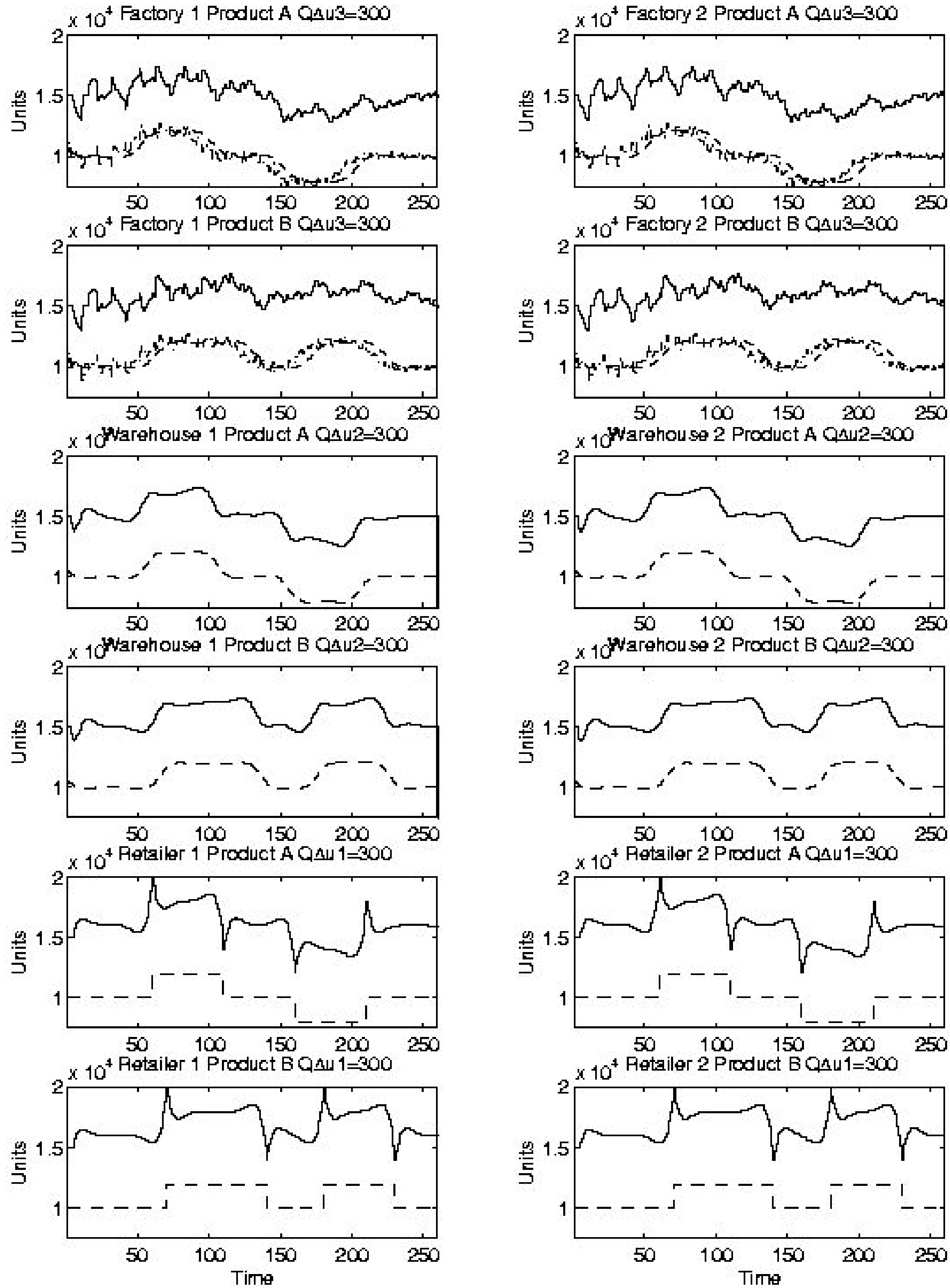

Figure 10 Inventories (solid) and demand (dashed) by facility and product; Assembly/Test plots include starts (dash-dot)

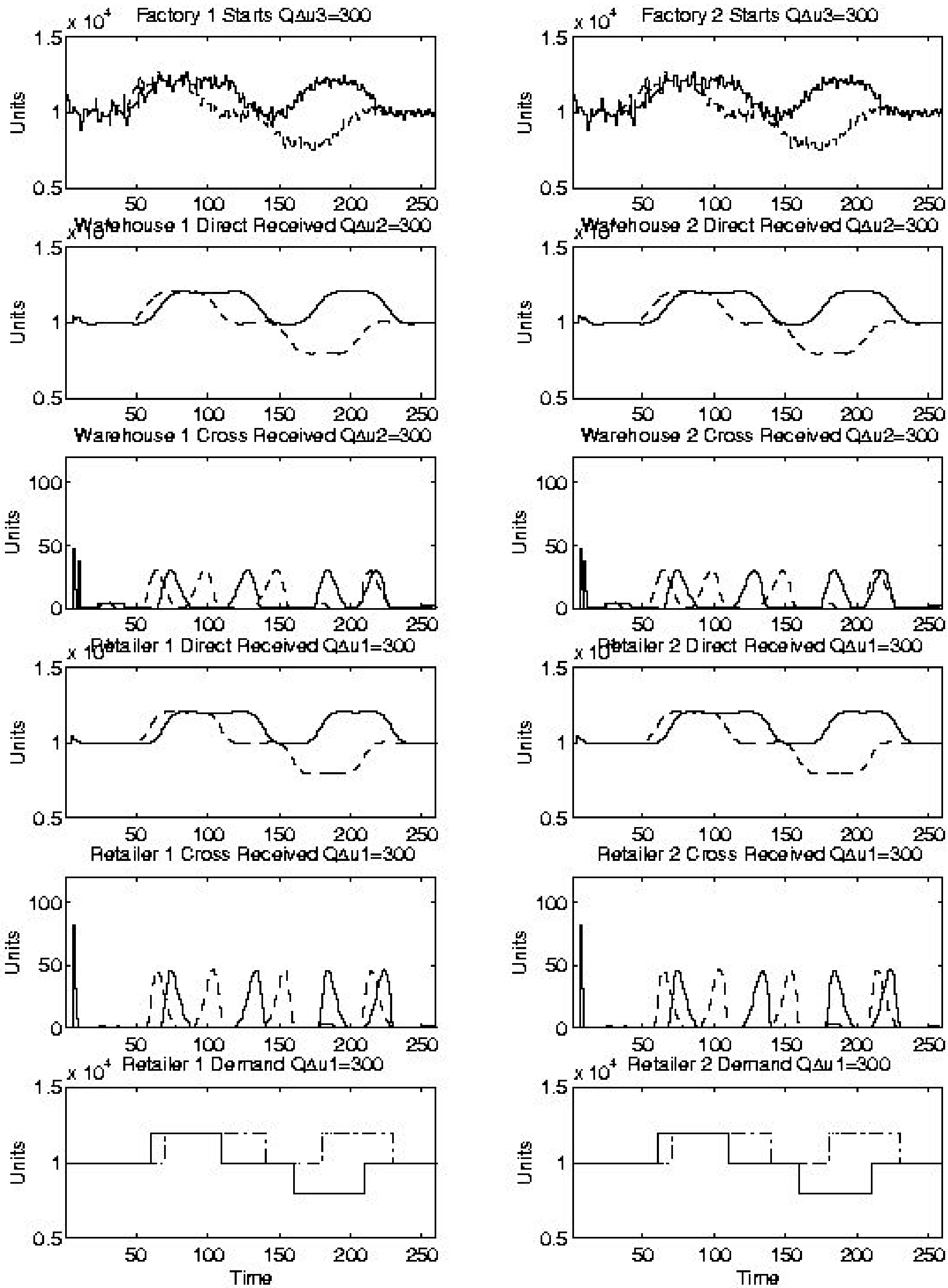

Figure 11 Material flows Total (solid) and Product A (dashed); Demand Product A (dotted) and Total (dash-dot) and corresponding customer receipts Product A (dashed) and Total (solid)