Simulation-Based Optimal Tuning of Model Predictive Control Policies for Supply Chain Management using Simultaneous Perturbation Stochastic Approximation

Jay D. Schwartz and Daniel E. Rivera[1]

Control Systems Engineering Laboratory
Department of Chemical and Materials Engineering
Arizona State University, Tempe, Arizona 85287-6006

Abstract— Efficient management of inventory in supply chains is critical to the profitable operation of modern enterprises. The supply/demand networks characteristic of discrete-parts industries such as semiconductor manufacturing represent highly stochastic, nonlinear, and constrained dynamical systems whose study merits a control-oriented approach. Model Predictive Control (MPC) is presented in this paper as the basis for a novel inventory management policy for supply chains whose dynamic behavior can be adequately represented by fluid analogies. A Simultaneous Perturbation Stochastic Approximation (SPSA) optimization algorithm is presented as a means to obtain optimal tuning parameters for the proposed policies. The SPSA technique is capable of optimizing important system parameters, such as safety stock targets and/or controller tuning parameters. Two case studies are presented. The results of the optimization on a single-echelon system show that it is advantageous to act cautiously to forecasted information and gradually become more aggressive (with respect to factory starts) as more accurate demand information becomes available. For a three-echelon problem, the results of the optimization demonstrate that safety stock levels can be significantly reduced and financial benefit gained while maintaining robust operation in the supply chain.

I. INTRODUCTION

More effective operation of supply chains for manufactured goods is worth billions of dollars to our national economy. While a generic supply chain stretches from suppliers through manufacturing to customers, one of the most promising areas for improvement lies in the generation and execution of the plans for the factories that form the backbone of all such supply chains. In this context, improved inventory management plays a critical role in plans that effectively allocate factory capacity to make the right product at the right time, avoiding wasting capacity on products that may later be discarded, and 2) reduce activities that are undesirable to operations, such as excessive variability in factory starts (the input to the factory and a manipulated variable in the controller) and setups. Ultimately, improving inventory management leads to lower manufacturing costs while maximizing revenue and improving customer satisfaction.
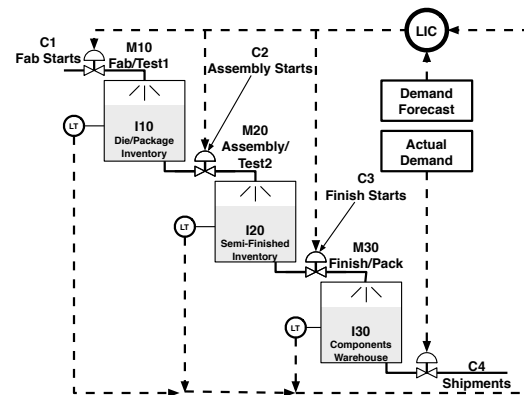

Fig. 1. Fluid representation for a representative three-echelon supply chain based on semiconductor manufacturing [4].

This paper is focused on how to effectively tune decision policies inspired from process control for supply chain inventory management problems in uncertain, stochastic environments, as is typically the case in industrial practice [4]. Material flows in manufacturing supply chains can be modeled using a fluid analogy. This analogy provides a basis for tactical decision policies based on process control principles. A fluid representation of a three-echelon supply chain as seen in semiconductor manufacturing and its corresponding inventory locations is shown in Figure 1. Here the manufacturing nodes are represented as "pipes", while the warehouse nodes are represented as "tanks". Material in these pipes and tanks correspond to Work-in-Progress and inventory, respectively.

Specifically, in this paper we consider Model Predictive Control (MPC) [1][3] as a decision policy that can provide improved performance in manufacturing systems with long throughput times and significant uncertainty, such as semiconductor manufacturing [4][8]. As control-oriented frameworks, MPC-based decision policies have the advantage that they can be tuned to provide acceptable performance in the presence of significant supply and demand variability and forecast error as well as constraints on production, inventory levels, and shipping capacity.

[1]To whom all correspondence should be addressed. phone: (480) 965-9476; fax: (480) 965-0037; e-mail: daniel.rivera@asu.edu
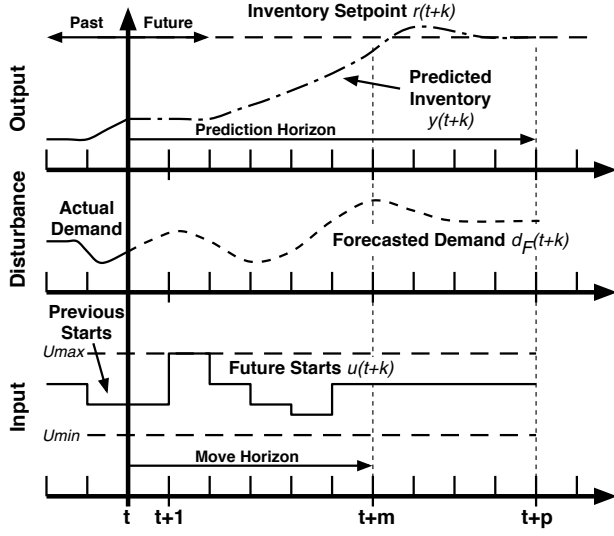
Fig. 2. Receding horizon representation of Model Predictive Control.

The ultimate objective of this paper is to present a simulation-based approach for optimally tuning these policies in a stochastic environment using the concept of Simultaneous Perturbation Stochastic Approximation (SPSA) [7]. The first scenario involves optimal tuning of a single node system under uncertain demand. In this case, the SPSA algorithm acts to maximize profit by optimizing over a range of move suppression values within the move calculation horizon. The second scenario involves determining both the optimal move suppression and inventory targets for a three echelon system involving stochastic yield, long throughput times, and uncertain demand. For this scenario, the algorithm maximizes profit by optimizing over the target value for each inventory node and the move suppression values that penalize changes in starts for each production node. In the second scenario move suppression values are constant across the move horizon, but there is a distinct weight assigned for each factory.

The paper is organized as follows. Section 2 is concerned with MPC-based tactical decision policies that are optimal with respect to linear time-invariant models derived using fluid analogies. Section 3 describes the SPSA optimization method that will seek optimal tuning and targets of the MPC policies when placed in a stochastic environment. Section 4 presents the results of applying SPSA for the previously described scenarios, with the results yielding some fundamental insights into the proper selection of inventory targets and tuning of the decision policies. A summary of the work and resulting conclusions are discussed in Section 5.

II. MPC AS A TACTICAL DECISION POLICY

Model Predictive Control (MPC) stands for a family of methods that select control actions based on on-line optimization of an objective function. MPC has gained wide acceptance in the chemical and other process industries as the basis for advanced multivariable control schemes [1][3]. In MPC, a system model and current and historical measurements of the process are used to predict the system behavior at future time instants. A control-relevant objective function is then optimized to calculate a sequence of future control moves that must satisfy system constraints. The first predicted control move is implemented and at the next sampling time the calculations are repeated using updated system states; this is referred to as a Moving or Receding Horizon strategy. Figure 2 is a useful visualization of the MPC approach. As shown, a demand forecast signal is used in the moving horizon calculation to anticipate future system behavior, which plays a significant role in the use of MPC for supply chain applications. There is significant flexibility in the form of the objective function that can be used in MPC; a meaningful formulation for the inventory management problems considered in this paper is as follows:

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

where the individual terms of J correspond to:

Keep Inventories at Inventory Planning Setpoints

$$J = \overbrace{\sum_{\ell=1}^{p} Q_e(\ell)(\hat{y}(k+\ell|k) - r(k+\ell))^2} \tag{2}$$

Penalize Changes in Starts

$$+ \overbrace{\sum_{\ell=1}^{m} Q_{\Delta u}(\ell)(\Delta u(k+\ell-1|k))^2}$$

subject to constraints on inventory capacity ($0 \leq y(k) \leq y_{max}$), factory inflow capacity ($0 \leq u(k) \leq u_{max}$), and changes in the quantity of factory starts ($\Delta u_{min} \leq \Delta u_k \leq \Delta u_{max}$). Equation 2 is a multi-objective expression that addresses the main operational objectives in the supply chain. The first term is a setpoint tracking term that is intended to maintain inventory levels at user-specified targets over time. The second term is a move suppression term that penalizes changes in the starts. Penalizing changes in the starts rate is not only desirable from the standpoint of factory operations, it also serves an important control-theoretic purpose, as a mechanism for introducing robustness in the control system in the face of uncertainty [3]. The emphasis given to each one of the sub-objectives in (2) (or to specific system variables within these objective terms) is achieved through the choice of weights ($Q_e(\ell)$ and $Q_{\Delta u}(\ell)$). These can potentially vary over the move and prediction horizons (m and p, respectively).

MPC control of a fluid representation of the three-echelon semiconductor manufacturing supply chain described in the Introduction and shown in Figure 1 is as follows: controlled variables y for the problem in Figure 1 consist of the three inventory levels (I_{10}, I_{20} and I_{30}). The starts rates for the production nodes (C_1, C_2 and C_3) represent manipulated variables in this problem formulation, while demand is treated as a disturbance signal. The demand signal (which dictates the shipment flow in C_4) consists of two components: 1) actual demand (which is only fully known after the fact) and 2) forecasted demand, which is provided to the planning function by a separate organization. For the problem in Figure 1 the mass conservation relationship for I_{10} can be written as:

$$I_{10}(k+1) = I_{10}(k) + Y_1\, C_1(k - \theta_1) - C_2(k) \quad (3)$$

where θ_1 and Y_1 represent the nominal throughput time and yield for the first production node, respectively, while C_1 and C_2 represent the daily (or per-shift) starts that constitute inflow and outflow streams for I_{10} and M_{10}. Similar relationships to (3) can be written for the other inventory nodes, I_{20} and I_{30}.

III. SIMULATION-BASED OPTIMIZATION USING SPSA

The previous section showed the development of controllers that are based on nominal linear models, but will be implemented in uncertain, stochastic settings. To achieve optimality in a stochastic setting, it is necessary to add a second optimization layer. We propose a simulation-based optimization scheme that will seek optimality under the effects of stochastic yield, varying throughput times, and an erroneous, autocorrelated demand forecast signal. The Simultaneous Perturbation Stochastic Approximation (SPSA) method is a promising approach that has received considerable attention over the last decade. The method has been used for statistical parameter estimation, model fitting, adaptive control, and many other applications [6][7].

Simulation-based optimization algorithms are generally applied to problems where a closed-form relationship between the parameters being optimized and the objective function is unknown. This may be due to the presence of noise in the objective function evaluation, or the relationship between the parameters and the function is significantly complex. The lack of gradient information prompts interest in optimization algorithms that rely solely on measurements of the objective function. Historically, scientists and engineers have used a standard "two-sided" finite-difference approximation such as the Kiefer-Wolfowitz stochastic approximation algorithm [5]. These approach require $2p$ function evaluations to obtain an estimate of the gradient, where p is the number of parameters being optimized.

The SPSA technique represents a significant improvement over traditional finite-difference stochastic approximation (FDSA) methods. The basis of the method is an efficient and intuitive "simultaneous perturbation" estimate of the gradient. Only two measurements of the objective function are required at each iteration, regardless of the number of parameters p. SPSA realizes the same level of accuracy as comparable FDSA methods for a given number of iterations despite the fact that only two measurements are made to form an estimate, as opposed to $2p$ measurements. Therefore, SPSA requires p times fewer evaluations of the objective function to achieve an equivalent result. This approach has been shown to provide greater accuracy and efficiency than comparable stochastic approximation methods [2][7].

The underlying premise of SPSA is the minimization of an objective function, J. The objective function J takes a real-valued vector of search parameters $\vec{x}$ of dimension p and returns a scalar. The process begins with an initial guess of the input vector $\vec{x}$ and iterates using the simultaneous perturbation estimate of the gradient $g(\vec{x}) = \delta J / \delta \vec{x}$. Note that this formulation is similar to the FDSA algorithm discussed previously, but differs in the nature of the gradient estimate. The SPSA algorithm consists of the following steps:

1) *Initialize the Input Vector and Gain Sequences*: An initial guess of the optimal input vector is made ($\vec{x}_0$). At this stage, one must also select the coefficients of the gain sequences a_k and c_k. These sequences govern the step size at each iteration and the magnitude of the perturbation, respectively.
2) *Generate the Perturbation Vector*: A random perturbation vector (Δ_k) is generated. Each element of the vector is independently generated using a Bernoulli ± 1 distribution with a probability of $\frac{1}{2}$ for each possible outcome.
3) *Evaluate the Objective Function*: Two measurements of the objective function are obtained: $J(\vec{x}_k + c_k \Delta_k)$ and $J(\vec{x}_k - c_k \Delta_k)$.
4) *Approximate the Gradient*: The simultaneous perturbation approximation of the gradient, $\hat{g}(\vec{x}_k)$, is determined using (4).

$$\hat{g}(\vec{x}_k) = \frac{J(\vec{x}_k + c_k \Delta_k) - J(\vec{x}_k - c_k \Delta_k)}{2c_k \Delta_k} \quad (4)$$

5) *Update the Estimate*: The standard stochastic approximation form (5) is used to update $\vec{x}_k$ to $\vec{x}_{k+1}$.

$$\vec{x}_{k+1} = \vec{x}_k - a_k \hat{g}(\vec{x}_k) \quad (5)$$

Both practical and asymptotically optimal values of the coefficients a_k and c_k are available in the literature [6][7]. The values used in the case studies of the next section were obtained directly from Section III.B in [6].
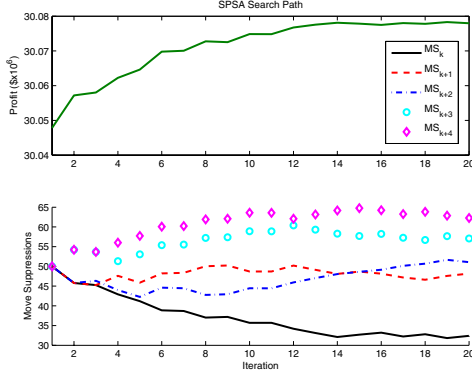
Fig. 3. SPSA optimization algorithm path for a 5-dimensional search space involving the selection of independent move suppression values over a horizon. Top: Profit, Bottom: Move Suppression Values.
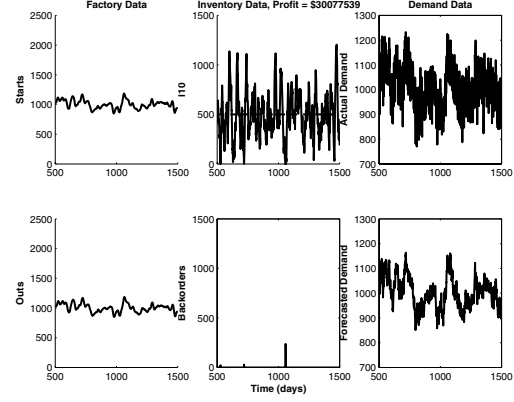

Fig. 4. Time series for a production/inventory node subject to uncertain demand. Move suppression values weights $\vec{Q}_{\Delta u} \sim [33\ 44\ 52\ 58\ 63]$ are obtained from the SPSA optimization shown in Fig. 3.

IV. CASE STUDIES

A. Financial Objective Function

It is desirable to financially optimize the MPC-based decision policy discussed earlier. Two cases studies are shown where the SPSA algorithm is used to determine optimal inventory targets (r_i) and/or move suppression weights ($Q_i^{\Delta u}$) according to the objective function shown in (6).

$$\max_{\vec{r},\vec{Q}^{\Delta u}} \text{Profit} = \text{Revenue} - \text{Production}_{cost} - \quad (6)$$

$$\text{Inventory}_{cost} - \text{Backorder}_{cost}$$

where

$$\text{Revenue} = \sum_{k=1}^{T_{final}} \gamma_R C_4(k) \quad (7)$$

$$\text{Production}_{cost} = \sum_{k=1}^{T_{final}} \sum_{j=1}^{N_{nodes}} \gamma_{C_j} C_j(k) \quad (8)$$

$$\text{Inventory}_{cost} = \sum_{k=1}^{T_{final}} \sum_{j=1}^{N_{nodes}} \gamma_I I_{10j}(k) \quad (9)$$

$$\text{Backorder}_{cost} = \sum_{k=1}^{T_{final}} \gamma_B \text{Backorders}(k) \quad (10)$$

This comprehensive objective function accounts for the production cost, inventory holding cost, backorder penalty, and revenue generation via the parameters γ_C, γ_I, γ_B, and γ_R, respectively. For the case studies presented here, the objective function weights are as follows: $\gamma_R = 40, \gamma_{C_1} = 10, \gamma_{C_2} = 8, \gamma_{C_3} = 2, \gamma_I = 0.1,$ and $\gamma_B = 5$.

B. Case Study 1: Tuning a Single Echelon Supply Chain

Figure 3 shows the optimization path for the SPSA algorithm applied to a single production/inventory node ($N_{nodes} = 1$) subjected to uncertain, autoregressive demand. The factory is characterized by a throughput time of 3 days and a 100% yield rate. The MPC prediction horizon p is 10 days and the move calculation horizon m is 5 days. Since the accuracy of a demand forecast decreases as projections are made further into the future, it is desirable to search for the optimal move suppression weights for each day in the move horizon (as opposed to using a single move suppression value over the entire move horizon).

SPSA is used to maximize profit by adjusting each move suppression value in the horizon. Lower move suppression values denote more aggressive factory starts changes and the optimizer shows that it is financially optimal to act most aggressively to information that is available at the current sampling instant and gradually be more detuned as the horizon extends into the future. The results corroborate with intuition as it is advantageous to act cautiously to far out forecasted information and gradually become more aggressive (with respect to factory starts) as more accurate demand information becomes available.

Fig. 4 shows the MPC simulation result when the optimal parameters obtained from SPSA are implemented. The most evident effect of move suppression is the "smoothing" effect it has on factory starts. Abrupt changes in starts are penalized in the MPC objective function, which leads to a factory starts response that is significantly less noisy than the actual demand. This is appealing from a factory manager's perspective, as substantial changes in the plant output may lead to disruptions from production plans and may be financially detrimental.
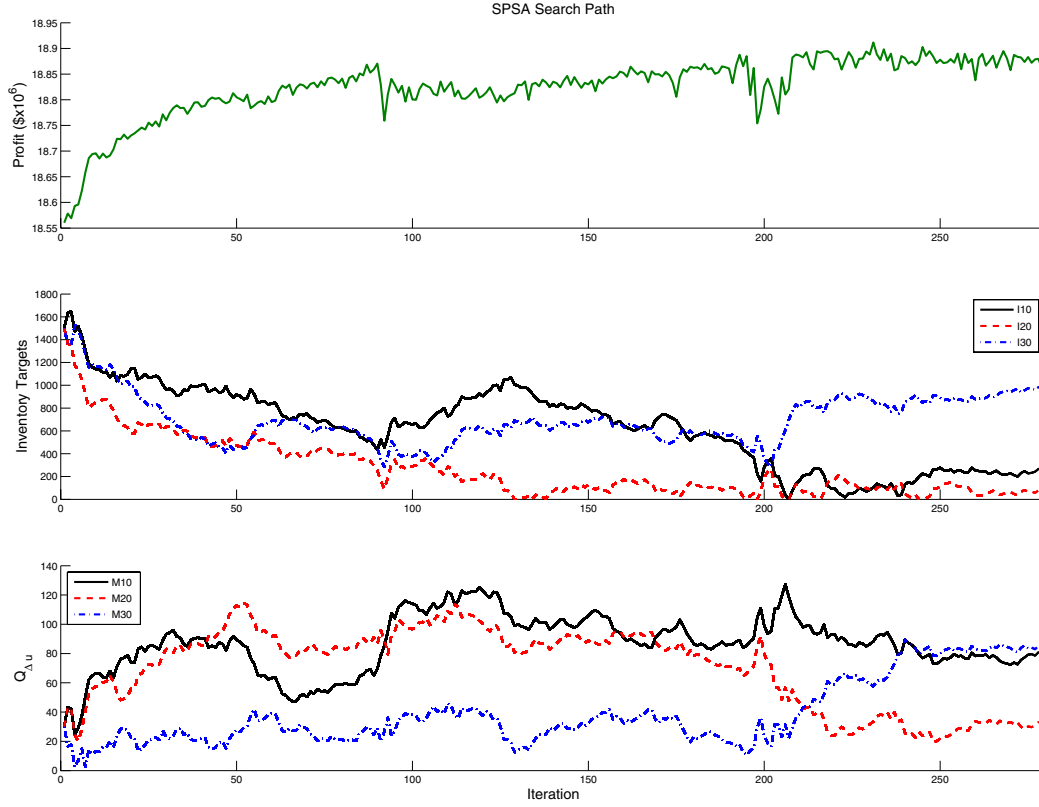
Fig. 5. SPSA optimization algorithm path for a 6-dimensional search space involving the selection of inventory targets and move suppression values for a semiconductor supply/demand network characterized by long throughput times, stochastic yield, and autocorrelated demand. Top: Profit, Middle: Inventory Targets, Bottom: Move Suppression Values.

C. Case Study 2: Optimizing Inventory Targets and Controller Tuning in a Three Echelon Supply Chain

Figure 5 shows the optimization path for the SPSA algorithm applied to the larger network topology ($N_{nodes} = 3$) shown in Figure 1. For this case study, both demand and supply (factory output) are uncertain. Measurements are repeated at each iteration to obtain more accurate estimates of the gradient, therefore the controller is being optimized with respect to the expected value of the overall profit for a given set of targets and tunings. The throughput times in M_{10} vary according to a triangular distribution, with 80% of the output produced after 35 days and the remaining 20% evenly split between days 34 and 36. Throughput times vary similarly in a similar for M_{20} and M_{30}, with ranges between 5 to 7 and 1 to 3 days, respectively. Yield rates vary uniformly for each manufacturing node (95% ± 2% for M_{10}, 98% ± 2% for M_{20}, and 99% ± 1% for M_{30}).

Initially, the optimizer drives the inventory targets towards zero. As baseline inventory levels are decreased, the risk of backorders increases.

Eventually, the optimization algorithm converges to an optimum where the inventory targets are approximately 400, 0, and 1000 units for I_{10}, I_{20}, and I_{30}, respectively. Given supply and demand uncertainty, it makes physical sense to keep a buffer in the final inventory stage (I_{30}, the inventory closest to the demand), but seek to minimize the amount of excess intermediate products stored in earlier stages. Note that the optimal target for I_{10} is greater than that of I_{20}, this agrees with intuition as the factory M_{10} has the longest throughput time and most stochastic behavior of all the production nodes. As demand variability and stochasticity increases, it will become necessary to keep larger inventories at all levels of the supply/demand network. For the cost values used in this study and the characteristics of the demand signal, profitability of the supply chain seems somewhat less sensitive to changes in the move suppression values than to changes in setpoint targets. Inventory and backorder costs are significantly greater than costs incurred by detuning the MPC decision policy.
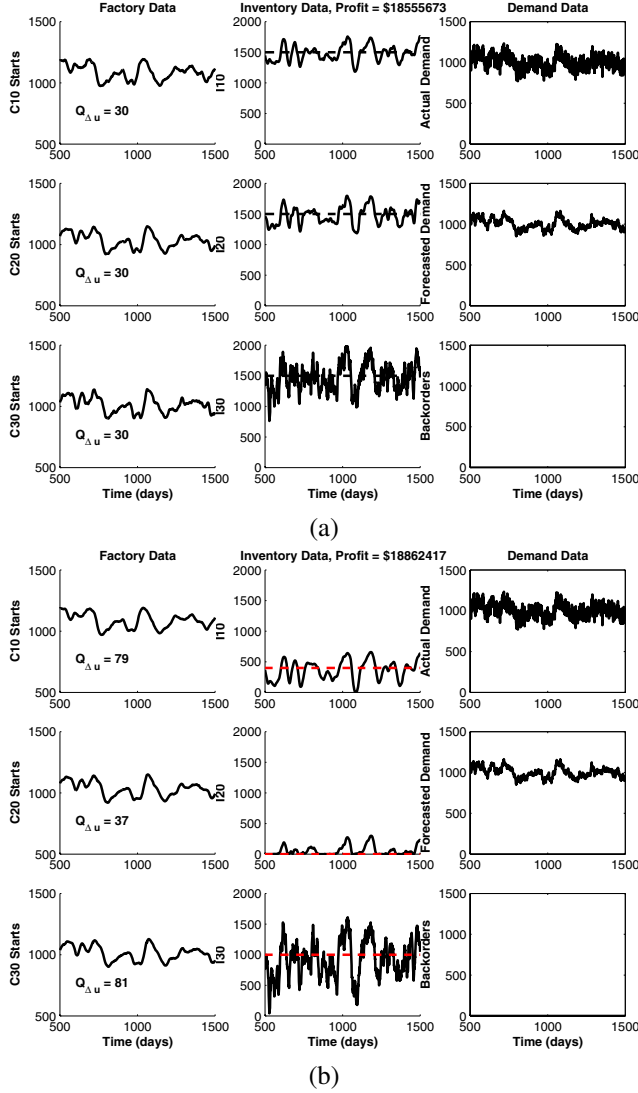
Fig. 6. Time series for the three echelon supply chain shown in Fig. 1. Top, initial guess: $\vec{Q}_{\Delta u} \sim [30\ 30\ 30]$ and $\vec{r} \sim [1500\ 1500\ 1500]$. Bottom, optimization results: $\vec{Q}_{\Delta u} \sim [79\ 37\ 81]$ and $\vec{r} \sim [400\ 0\ 1000]$.

Fig. 6(a) shows the MPC simulation result when the initial guess (corresponding to the zeroth iteration in Fig. 5) is used to parameterize the system. The high safety stock levels ensure that there are no backorders. However, the targets can be reduced further to eliminate the cost of holding excess inventory, thus increasing the profit. Fig. 6(b) shows the MPC simulation where controller tuning and inventory targets are determined from the SPSA optimization algorithm (the final iteration shown in Fig. 5). Safety stock levels are reduced to a level where inventory targets are as low as possible without incurring backorders. This minimizes the inventory holding costs and increases profit.

V. SUMMARY AND CONCLUSIONS

Model Predictive Control (MPC) has been demonstrated to be capable of managing inventory in uncertain multi-echelon supply/demand networks under constraints. The use of a simulation based optimization strategy allows for the determination of controller tunings and operating targets that lead to optimal results from either an operational or financial standpoint. The results of the optimization on a single node example show that it is advantageous to act cautiously to forecasted information and gradually become more aggressive (with respect to factory starts) as more accurate demand information becomes available. For the three echelon problem, the use of the simulation-based optimization method led to insights concerning the proper parameterization and tuning of the tactical MPC decision policy. The amount of safety stock necessary for optimal profitability is a function of the accuracy and magnitude of the demand forecast. SPSA provides a way of systematically determining the financially optimal inventory targets and the move suppression values present in the MPC objective function simultaneously. In practice, the optimization could be performed at regular time intervals or whenever updated demand forecasts become available. For the semiconductor manufacturing problem posed here, it was found that the optimization problem was more sensitive to changes in inventory targets, and less sensitive to changes in move suppression. This allows for flexibility when tuning the decision policy, as robustness considerations do not have to be cast aside in favor of increased profitability.

VI. ACKNOWLEDGEMENTS

REFERENCES

[1] E. F. Camacho and C. Bordons. *Model Predictive Control*. Springer-Verlag, London, 1999.

[2] D. C. Chin. Comparative study of stochastic algorithms for system optimization based on gradient approximations. *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, 27:244–249, 1997.

[3] C. E. García, D. M. Prett, and M. Morari. Model predictive control: theory and practice-a survey. *Automatica*, 25(3):335–348, 1989.

[4] K. G. Kempf. Control-oriented approaches to supply chain management in semiconductor manufacturing. In *Proceedings of the 2004 American Control Conference*, pages 4563–4576, Boston, MA, 2004.

[5] J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, 23(3):462–266, 1952.

[6] J. C. Spall. Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3):817–823, 1998.

[7] J. C. Spall. *Introduction to Stochastic Search and Optimization Estimation, Simulation, and Control*. John Wiley and Sons, Inc., Hoboken, New Jersey, 2003.

[8] W. Wang, D. E. Rivera, K. G. Kempf, and K.D. Smith. A model predictive control strategy for supply chain management in semiconductor manufacturing under uncertainty. In *Proceedings of the American Control Conference*, pages 4577–4582, Boston, MA, 2004.