

**Integration of Inventory and
Transportation Decisions in a
Logistics System: A Case
Study of the Coal and Petroleum
Industry in China**

Q.-H. Zhao, S. Chen,
S. C.H. Leung, K.K. Lai

G-2008-06

January 2008

Les textes publiés dans la série des rapports de recherche HEC n'engagent que la responsabilité de leurs auteurs. La publication de ces rapports de recherche bénéficie d'une subvention du Fonds québécois de la recherche sur la nature et les technologies.

**Integration of Inventory and Transportation
Decisions in a Logistics System:
A Case Study of the Coal and Petroleum
Industry in China**

Qiu-Hong Zhao

*GERAD and School of Economics and Management
Beihang University
Beijing, China
qiuhong.zhao@gerad.ca*

Shuang Chen

*School of Economics and Management
Beihang University
Beijing, China*

Stephen C.H. Leung

K.K. Lai

*Department of Management Sciences
City University of Hong Kong
Hong Kong*

January 2008

Les Cahiers du GERAD

G-2008-06

Copyright © 2008 GERAD

Abstract

Vendor managed inventory (VMI) takes on a new form in our case study. It occurs between a parent company and its several subsidiaries along the Yangtze River, and the parent company is responsible for the delivery of coal which is the raw material for production. Since coal is firstly supplied from the supplier to the port by train, and is then transferred to the subsidiaries by river using water carriers, we propose to establish a central warehouse at the port which should possess a fleet of vessels, making the ordering and delivery decisions under the VMI mode. Both decisions are based on Markov Decision Process (MDP), to optimize the replenishment strategy with the consideration of holding, shortage and transportation costs. An algorithm based on Modified Policy Iteration (MPI) with an action elimination procedure is applied to the MDP model, so an approximate optimal solution can be found within reasonable time. It is shown that the proposed strategy is able to reduce the overall costs of the system in contrast to the current mode, whilst also guaranteeing higher service level to the subsidiaries.

Key Words: Vendor managed inventory; Integrated logistics system; Inventory and transportation decisions; Markov Decision Process; Modified Policy Iteration algorithm.

Résumé

Le stock géré par le vendeur (SGV) prend une nouvelle forme dans notre étude de cas. Il prend place entre une compagnie parente et plusieurs sous-traitants le long de la rivière Yangtze et la compagnie parente est responsable pour l'approvisionnement en charbon qui est la matière première pour la production. Comme le charbon est d'abord livré depuis le fournisseur jusqu'au port par train, et est ensuite transféré au sous-traitant par la rivière à l'aide de barges, nous proposons d'établir un entrepôt central au port qui posséderait une flotte de vaisseaux et prendrait les décisions de commande et de livraison sous le mode SGV. Les deux décisions sont basées sur le processus de décision Markovien (PDM) afin d'optimiser la politique de renouvellement sous des considérations de disponibilité, rupture de stock et coûts de transport. Un algorithme basé sur l'itération d'une politique modifiée avec une procédure d'itération d'action est appliquée au modèle (PDM) de sorte qu'une solution approximativement optimale peut être trouvée en temps raisonnable. On montre que la stratégie proposée est capable de réduire les coûts totaux du système par rapport au mode d'opération courant tout en garantissant un plus haut niveau de service aux sous-traitants.

Mots clés : stock géré par le vendeur; système logistique intégré; décision de stockage et de transport; modèle de décision Markovien; algorithme d'itération de politique modifiée.

Acknowledgments: We are grateful to the anonymous referees for their many valuable and helpful comments. This work was partially supported by National Natural Science Foundation of China under Project No. 70301001.

1 Introduction

Vendor-managed inventory (VMI) is a supply-chain initiative where the supplier is authorized to determine and manage inventories of agreed-upon stock-keeping units at retail locations (Aviv and Federguen, 1998). It is shown that VMI can reduce inventories and shortages by using advanced online messages and data-retrieval systems (Angulo et al., 2004). In addition, as the vendor is guided by mutual agreements on inventory levels, fill rates and transaction costs, trading partners can maximize their benefit (Andel, 1996). However, as most of the systems under considered are complicated, the vendor is usually faced with the challenges of designing the integrated replenishment strategy (Silver et al., 1998).

In this paper, the VMI mode is applied to a two-echelon logistics system consisting of parent company — China Petroleum and Chemical Corporation (hereafter referred to as Sinopec) and its several subsidiaries along the Yangtze River, while the parent company is responsible for the delivery of coal which is the raw material for production, to its subsidiaries. The coal is firstly supplied from Huainan Mining Group Corporation (HMGC) to the Port of Wu Hu (WHP) by train, and is then transferred to the subsidiaries by river using water carriers.

In this paper, we propose to establish a central warehouse at WHP, own a fleet of vessels, and use the VMI mode at the central warehouse for ordering and delivery decisions. The objective is to integrate ordering and delivery decisions in the system such that the overall costs, including holding, shortage and transportation costs, can be minimized. Both the ordering and delivery decisions of the central warehouse are based on Markov Decision Process (MDP) with the consideration of holding, shortage, and transportation costs. An algorithm based on Modified Policy Iteration (MPI) with the action elimination procedure is designed to the MDP model, such that an approximate optimal solution can be found within reasonable time.

The rest of the paper is organized as follows. In Section 2, the background of the problem is given, followed by some strategies based on the VMI mode. Section 3 is the literature review. Section 4 addresses the delivery and the ordering decisions, and develops the MDP models. In Section 5, a MPI-based algorithm for solving the MDP models in Section 4 is presented. Comparative analysis of the proposed system with the current system is conducted in Section 6. Finally, the conclusion is given in Section 7.

2 The background of the problem

Sinopec is the largest producer and supplier of petrochemical products in China, where includes more than 80 subsidiaries, either wholly-owned, or with equity participation or majority control. Coal is one of the important sources for the petrochemical products such as synthetic resin, synthetic fiber monomers and polymers, synthetic fiber and chemical fertilizer. In 2004, Sinopec consumed 24.91 million tons of coal, of which 7.16 million tons was consumed by the subsidiaries in the Yangtze River region. One of the coal suppliers for this region is HMGC. Each year, about 0.6 million tons of coal is midway-transferred transshipped from WHP to Sinopecs subsidiaries. WHP is the largest port for coal transshipment along the Yangtze River and playing crucial role in coal distribution.

2.1 The current system and the problem

Yangzi (YA), Yizheng (YI), Jinling (JI) and Anqing (AN) are four main Sinopecs subsidiaries along the Yangtze River. Observing from ordering data of past three years, it is estimated

the monthly demand of each subsidiary to closely follow Poisson distribution. In the current delivery system, each of subsidiaries submits its monthly order to the parent company where transfers the order to the supplier (HMGC), without making any modification. The coal is delivered to WHP by train and then transferred to the subsidiaries using water carriers. See Figure 1 for reference.

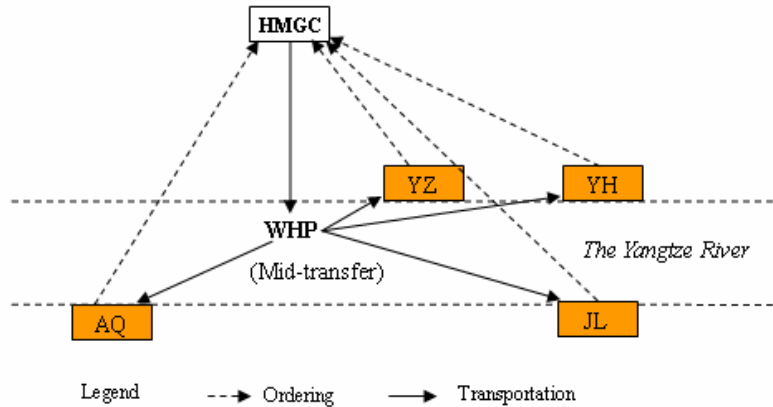


Figure 1: The current logistics system

Nowadays, in China, rail carriers are mainly owned by government, whilst water carriers are owned by government or private companies. In the current logistics system, Sinopec and subsidiaries hire trains and vessels for goods delivery and hence, they are out of control of Sinopec and the subsidiaries and they are often not on hand.

In recent years, the rapid development of China makes coal an important resource for industries, and the demand of railway and vessels keen, all of which make the lead time of ordering long and unstable. As a result, to guarantee stable supply of the coal, each subsidiary has to maintain huge inventory, resulting in complex delivery process and higher operation cost.

2.2 The general solution proposed

To solve the problem faced by the logistics system, a general solution based on VMI mode is proposed as follows.

Since HMGC, the rail carriers and the water carriers usually pay more attention to the large-scale order, in order to make the supply of the coal more stable, integration of the coal orders from the four subsidiaries is the first step in the general solution.

Further, as the unit inventory holding cost at WHP is lower than that at any subsidiary, the second step in the general solution is to establish a central warehouse in WHP so that most of coals can be stocked in the warehouse. Moreover, the central warehouse is responsible for ordering and delivery process and acts as the vendor of the subsidiaries in the VMI mode.

Finally, to reduce the inventory at the subsidiaries as low as possible, the Just-in-Time (JIT) mode is introduced in the delivery process from WHP to the subsidiaries. To achieve this, it is suggested that a fleet of vessels should be owned by the central warehouse. They act as backup in case the water carriers cannot handle and deliver the cargoes timely.

The proposed strategy can be summarized as follows. A central warehouse is proposed to establish at WHP, which will be leased and operated by the parent company. At the beginning of each decision period, the central warehouse sends the order to HMGC and makes delivery decision for the subsidiaries. The objective of both delivery and ordering decisions is to minimize the overall costs of the logistics system which implements JIT mode with considering self-owned fleet of vessels. See Figure 2 for reference.

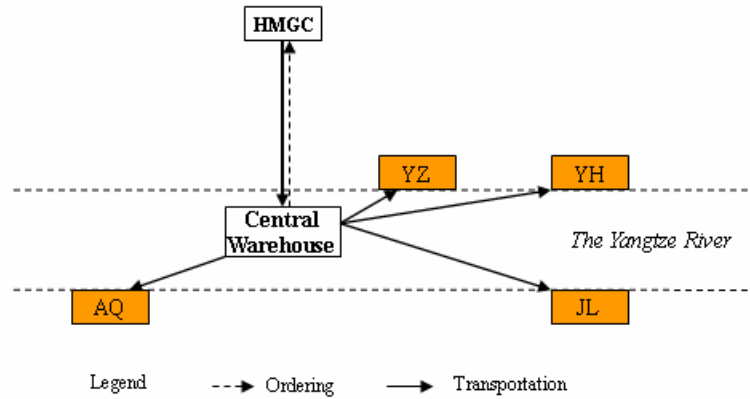


Figure 2: The proposed logistics system

In the proposed system, reasonable strategies should be proposed for ordering and delivery decisions. Here, the Markov Decision Process (MDP) is adopted for both decisions, followed by an algorithm based on Modified Policy Iteration (MPI) with an action elimination procedure. Further, since Sinopec should pay for setting up the central warehouse and owning vessels, more quantitative analysis and comparison of the proposed system with the current system should be conducted.

3 The literature

Managing inventories under uncertainty has received a lot of attention from academics and practitioners alike, and several kind of period review and reordering policies have been presented accordingly. Clark and Scarf (1960) showed that, in a serial multiechelon inventory system, when there is no fixed ordering cost and demand occurs at the lowest echelon, the optimal inventory control of the overall system follows a Base-Stock policy for every echelon. For the models with fixed ordering cost, the class of optimal policies are called as state dependent (s, \bar{s}) policies (Semi-Levi and Zhao, 2004), where s and \bar{s} are lower and upper thresholds that are determined by the given conditional distribution (at that period) on future demands. The rather simple forms of these policies do not always lead to efficient algorithms for computing the optimal policies, however. So the corresponding dynamic programs are relatively straightforward to solve, which is usually very time consuming when the state space is large. To avoid ‘curse of dimensionality’, some researchers have attempted to construct computationally efficient (but suboptimal) heuristics for these problems, such as myopic policies, which attempt to minimize the expected cost for one period, ignoring the potential effect on the cost in future periods (Levi et al. 2006).

Contrasting to the models considering fixed ordering costs, those integrating both transportation and inventory decisions make the optimal policies more intractable. Ganeshan (1999) presents a (s,Q)-type inventory policy for a network with multiple suppliers by replenishing a central depot, which in turn distributes to a large number of retailers; his paper considers transportation costs, but only as a function of the shipment size. Qu et al. (1999) deal with an inbound material-collection problem so that decisions for both inventory and transportation are made simultaneously; however, vehicle capacity is assumed to be unlimited so that it is solved as a traveling salesman problem (TSP). Cheung and Lee (2002) present a shipment coordination and stock rebalancing strategy in ordering to enjoy economies of scale in shipment. Cardos and Garcia-Sabater (2006) study an outbound delivery problem considering multiple items; they integrate the inventory and transportation decisions into one model and transportation costs are calculated upon the basis of detailed capacitated delivery schedules by VRP heuristic.

Despite the large number of inventory related models developed, there is still a wide gap between theory and practice. Markovian formulations (see Giannoccaro and Pontrandolfo, 2002; Hartanto et al., 2005) are useful in solving a number of real-world problems under uncertainties such as determining the inventory levels for retailers, maintenance scheduling for manufacturers, and scheduling and planning in production management. It is shown the MDP system results in lower average inventory level and requires fewer reorders to be placed (Yin et al., 2004). These advantages are more pronounced for products with highly variable demands, for which most of the other methods do not perform well.

A Markov Decision Process is a sequential decision-making stochastic process characterized by five elements: decision epochs, states, actions, transition probabilities, and rewards (Puterman, 1994). At each decision epoch, the system occupies a decision-making state. As a result of taking an action in a state, the decision-maker receives a reward (which may be positive or negative) and the system goes to the next state with a certain probability which is called the transition probability. A decision rule is a function for selecting an action in each state, while a policy is a collection of such decision rules over the state-space. Implementing a policy generates a sequence of rewards. The MDP problem is to choose a policy to optimize a function of this reward sequence.

Value iteration and *policy iteration* are two fundamental dynamic programming algorithms for solving MDPs (Howard, 1960), which are sometimes inefficient. Puterman and Shin (1978) proposed a Modified Policy Iteration algorithm, which seeks a trade-off between cheap and effective iterations and is preferred by some practitioners. In a sense, it is just policy iteration where the policy evaluation step is carried out via value iteration. This can be shown to produce an approximation to substantial speedups. In our paper, the algorithm is adopted to solve the MDP models. To improve the efficient of the Modified Policy Iteration algorithm for our case, we add one of action elimination procedures in Puterman and Shin (1982) at each iteration.

4 The delivery and the ordering decisions

Wuhu port (WHP) is the largest transshipment point of coal along the Yangtze River. There are some coal suppliers, but their prices are higher than HMG. So the delivery decision to the subsidiaries can be guaranteed if the quantities on hand in the central warehouse are not enough.

Due to the realization of the supply situation mentioned above, the delivery decision and the ordering decision can be made separately, and the optimal solution of each stage is also the optimal decision for the overall logistics system (Lee et al., 2000).

4.1 The delivery decision to the subsidiary companies

In the delivery process, the subsidiaries check their inventory levels periodically, T , in discrete time point tT , $t = 1, 2, \dots$, and report it to the central warehouse. Then, in order to minimize the expected long-run costs including holding, shortage and transportation costs, the central warehouse delivers the optimal quantity to each subsidiary under JIT mode, and the cargoes delivered is to meet the demand in period $[tT, (t+1)T)$. The unit weight of coal is 1,000 tons in our case.

To determine the optimal decision, the MDP model is employed. Considering the practical situation that coal in one vessel must be shipped to one subsidiary company due to expensive transportation and unloading costs, the MDP model is formulated and solved for each subsidiary i respectively.

Before the illustration of the delivery strategy, notations are defined as follows.

- i : index of the subsidiary, where 1 for YA, 2 for YI, 3 for JI and 4 for AN.
- T : length of decision period. In our case the inventory is checked monthly.
- v : capacity of the vessel.
- c_v : vessel's unit variable transportation cost (in *Yuan*, 1US\$=7.8*Yuan*).
- C_v : vessel's fixed transportation cost per mission (in *Yuan*).
- d_i : distance from the central warehouse to subsidiary i (in *Kilometer*).
- h_i : unit holding cost at subsidiary i per decision period (in *Yuan*).
- q_i : unit shortage cost at subsidiary i per decision period (in *Yuan*).
- μ_i : mean of demand subsidiary i faces per decision period (in 1,000 tons). Then $p(x_i = r) = e^{-\mu_i} \frac{(\mu_i)^r}{r!}$ is the probability when the demand of subsidiary i is 1,000 r , $r = 1, \dots, \infty$. In our case, the demands of the subsidiaries can be regarded to be independent of each other and independent in different decision periods.

The two items c_v and C_v refer to those of the water carriers. The vessels owned act as the supplement in case the vessels from the water carrier cannot handle and deliver the cargoes timely, so the quantities delivered by them occupy a little part of the total. For the purpose of simplification, we do not differentiate these two items between the vessels owned and rented.

As the unit time is one month in our case study, and the longest return trip to the subsidiaries can be finished within one day, it is estimated by the department manager of Sinopec that two vessels with the same size as those owned by the ship company (1,500 tons) is enough to meet the requirement of JIT delivery. Both the price and the buoyancy of the river are the reasons for selecting the vessels with size of 1,500 tons, which are beyond the consideration of our case study. In addition, since the vessels are bought at first decision period, the cost, denoted as C_V , is considered in the computational analysis section other than in the MDP model.

For any subsidiary i , to all the possible quantities on hand (the inventory) in different decision periods, a delivery policy refers to the quantities delivered accordingly. Denote π as a stable delivery policy, which means the decision is stable at different decision periods, our objective is to find the optimal policy π^* over an infinite time horizon. To determine

the optimal inventory-delivery decision, we formulate the above problem as a discrete time Markov Decision Process (MDP) with finite state and action space. It contains the following five components (for the sake of notational convenience, we omit the sub-symbol i in the following expression).

State space S . S is the set of X^t which is the possible on-hand inventory in the subsidiary at time point tT (before the occurrence of the demand). An upper bound D^{\max} and a lower bound D^{\min} can be given to the quantity of demand in the reality, by ignoring those the probabilities are very small (lower than 0.01 in our case). Therefore, we have $X^t \in S = \{0, 1, \dots, D^{\max} - D^{\min}\}$.

Action space A . For any state $X^t \in S$, the action space $A(X^t)$ is the set of delivery decisions. Our objective is to find the optimal policy which is stable over an infinite time horizon, so the delivery decision is based on the state only, having none relation with the time period the decision is made on. And the delivery decision variable $a \in A(X^t) = \{0, 1, \dots, D^{\max}\}$.

Transfer matrix G . Denote D^t as the demand at time period $[tT, (t+1)T)$, so the quantity consumed during this period is $\min\{X^t + a, D^t\}$, $a \in A(X^t)$, and the inventory at time point $(t+1)T$ is $X^{t+1} = \max\{X^t + a - D^t, 0\}$. Denote $P(X^{t+1} | X^t, a)$ as the probability the inventory state transfers from X^t to X^{t+1} under decision $a \in A(X^t)$. Let $X^t = k, X^{t+1} = j$, then

$$P(X^{t+1} | X^t, a) = P(j | k, a) = P_{kj}(a) = \begin{cases} p(D^t \geq k + a), & j = 0 \\ p(D^t = k + a - j), & j \neq 0 \end{cases} \quad (1)$$

Denote G as the state-transfer matrix. G can be a three-dimension matrix $G(k, j, a)$ or a two-dimension matrix $G(k, j)$ or $G(j, a)$, depending on the variables considered.

Instantaneous reward g . Denote $g(X^t, a)$ as the single decision period expected cost with decision $a \in A(X^t)$ under state X^t , then,

$$g(X^t, a) = n_t(C_v + 2dc_v) + E(I^t) + E(s^t), \quad (2)$$

where $n_t(C_v + 2dc_v)$ is the transportation cost, in which $n_t = \lceil \frac{a}{v} \rceil$ is the transportation times of the vessels, C_v is the fixed transportation cost of each mission, and $2dc_v$ is the return variable transportation cost; $E(I^t)$ and $E(s^t)$ are expected inventory cost and expected shortage cost, respectively. Denote $\bar{D} = D^{\max} - D^{\min}$, we have,

$$E(I^t) = \sum_{I^t \in \{0, \bar{D}\}} p(D^t = X^t + a - I^t) I^t h,$$

$$E(s^t) = \sum_{s^t \in \{0, \bar{D}\}} p(D^t = X^t + a + s^t) s^t q,$$

where $D^t \in [D^{\min}, D^{\max}]$.

Long-run expected discounted cost V . Denote $\pi = \{\pi(k) | k = 0, 1, \dots, D^{\max} - D^{\min}\}$, and $\beta \in [0, 1)$ be the economics discount factor. With an initial state k , the objective of the delivery decision is to find a policy π such that the expected discounted cost $V(\pi, k)$ over an infinite time horizon is minimal, and we have

$$\begin{aligned}
V(\pi, k) &= \sum_{t=1}^{\infty} \beta^{t-1} \sum_{\substack{j \in S \\ a \in A(k)}} P(X^t = j \mid X^1 = k, a) g(j, a) \\
&= g(k, \pi(k)) + \beta \sum_{j \in S} P(j \mid k, \pi(k)) V(\pi, j).
\end{aligned}$$

Here $k \in S, \pi \in \Pi$, Π is the set of all Markov policies. Denote $V(\pi^*, k)$ as the optimal expected cost with initial state k , π^* is the optimal policy accordingly, then

$$V(\pi^*, k) = \min_{\pi \in \Pi} V(\pi, k) = \min_{\pi \in \Pi} \{g(k, \pi(k)) + \beta \sum_{j \in S} P(j \mid k, \pi(k)) V(\pi, j)\}. \quad (3)$$

4.2 The central warehouse's ordering strategy

In the proposed system, in addition to fulfill the demands and makes delivery decisions to the subsidiaries, the central warehouse makes ordering decisions for itself as well. The quantities ordered will be arrived after a lead time period. For the ordering strategy, we can also follow the MDP to find the optimal solution.

Before the illustration of the MDP model, notations used are listed as bellow.

- c_n : unit transportation cost from HMGC to the central warehouse (in *Yuan*).
- C_w : fixed cost for hiring and operating the central warehouse per decision period, which is supposed to be independent of the quantities on hand (in *Yuan*).
- C_f : fixed cost for holding the vessels per decision period (in *Yuan*).
- q_w : unit penalty cost for obtaining items from the alternative source instead of HMGC. It equals to unit shortage cost and happens only when the quantity on hand is less than that delivered to the subsidiaries (in *Yuan*).
- h_w : unit holding cost per decision period at the central warehouse (in *Yuan*).
- l : lead time of the central warehouse for ordering decision (in *month*), which is regarded to be a constant in the new strategy.

Similar to the delivery decision, we aim at finding the optimal ordering policy which is stable over an infinite time horizon. However, the existence of the lead time and the attribution that the lead time is longer than the decision time period make the solution be complicated. To solve the problem by MDP model, the decision variable L made at time point tT is defined as the sum of the quantities ordered but still not yet arrived, that is, the total quantities ordered in time periods $\sum_{i=1}^l (t-i+1)T$.

To formulate the transfer matrix, it is assumed that the demand faced by the central warehouse during the lead time period follows Poisson distribution, with mean $U = \sum_{i=1}^4 u_i l$ ($l = 2$ in our case). Explanation for such assumption is listed in Appendix 1.

State space S_w . S_w is the set of Y^t which is the possible inventory level in the central warehouse at time point tT (before the realization of demand).

Lemma 1. $Y^t \in \{0, 1, \dots, \sum_{i=1}^4 D_i^{\max} l - \sum_{i=1}^4 D_i^{\min} l\}$.

Proof. At time point tT , the decision variable L is defined as the sum of the quantities ordered but still not yet arrived. It can be deduced that the quantities on hand plus L are to meet the demand faced by the central warehouse during the lead time period, which is $L + Y_t \in [\sum_{i=1}^4 D_i^{\min}l, \sum_{i=1}^4 D_i^{\max}l]$. It is obvious $Y^t \in \{0, 1, \dots, \sum_{i=1}^4 D_i^{\max}l - \sum_{i=1}^4 D_i^{\min}l\}$.

Similar to the analysis to the subsidiary, an upper bound D_w^{\max} and a lower bound D_w^{\min} are given to the quantity of the demand faced by the central warehouse during the lead time period, with probability less than 0.01. So,

$$Y^t \in S_w = \{0, 1, \dots, D_w^{\max} - D_w^{\min}\} \in \{0, 1, \dots, \sum_{i=1}^4 D_i^{\max}l - \sum_{i=1}^4 D_i^{\min}l\}.$$

Action space A_w . For any state $Y^t \in S_w$, the action space $A_w(Y^t)$ is the set of L defined as the sum of the quantities ordered but still not yet arrived. It is easily deduced $L \in A_w(Y^t) = \{0, 1, \dots, D_w^{\max}\}$.

Transfer matrix G_w . Denote D_w^t as the total demand faced by the central warehouse during time periods $\sum_{i=0}^{l-1} (t+i)T$, the quantity consumed (from what provided by HMGC) during the duration is $\min\{L + Y^t, D_w^t\}$, and the inventory at the end of this duration is $Y^{t+l} = \max\{L + Y^t - D_w^t, 0\}$. Denote $F(Y^{t+l} | Y^t, L)$ as the probability the inventory state transfers from Y^t to Y^{t+l} under decision $L \in A_w(Y^t)$. Let $Y^t = k, Y^{t+l} = j$, then

$$F(Y^{t+l} | Y^t, L) = F(j | k, L) = F_{kj}(L) = \begin{cases} p(D_w^t \geq k + L), & j = 0 \\ p(D_w^t = k + L - j), & j \neq 0 \end{cases} \quad (4)$$

where D_w^t follows Poisson distribution with mean $\sum_{i=1}^4 \mu_i l$. State-transfer matrix G_w can be a three-dimension matrix $G_w(k, j, L)$ or a two-dimension matrix $G_w(k, j)$ or $G_w(j, L)$, depending on the variables considered.

Instantaneous reward g_w . Denote $g_w(Y^t, L)$ as the single period expected cost with decision variable $L \in A_w(Y^t)$, then,

$$g_w(Y^t, L) = C_w + C_f + E(R_n^t) + E(I_w^t) + E(s_w^t), \quad (5)$$

where $E(R_n^t)$, $E(I_w^t)$ and $E(s_w^t)$ are respectively the expected variable transportation cost, expected inventory cost and expected shortage cost. As $E(R_n^t)$ is related to the quantity ordered at time period tT , it is determined by L and difficult to calculate. Here we estimate $E(R_n^t)$ as

$$E(R_n^t) = \frac{c_n L}{l}.$$

Denote $\bar{U} = (\sum_{i=1}^4 D_i^{\max}l - \sum_{i=1}^4 D_i^{\min}l)$, $E(I_w^t)$ and $E(s_w^t)$ can be calculated as follows.

$$\begin{aligned} E(I_w^t) &= \sum_{I_w^t \in \{0, \bar{U}\}} p(U^t = Y^t + L - I_w^t) I_w^t h_w, \\ E(s_w^t) &= \sum_{s_w^t \in \{0, \bar{U}\}} p(U^t = Y^t + L + s_w^t) s_w^t q_w, \end{aligned}$$

where $U^t \in [D_w^{\min}, D_w^{\max}]$.

Long-run expected discounted cost V_w . Denote $\pi_w = \{\pi_w(k) \mid k = 0, 1, \dots, D_w^{max} - D_w^{min}\}$. With an initial state k , the objective of the ordering decision is to find a policy π_w with minimal expected discounted cost $V_w(\pi_w, k)$ over an infinite time horizon, and we have

$$\begin{aligned} V_w(\pi_w, k) &= \sum_{t=1}^{\infty} \beta^{t-1} \sum_{\substack{j \in S_w \\ L \in A_w(k)}} F(X^{t+l} = j \mid X^1 = k, L) g_w(j, L) \\ &= g_w(k, \pi_w(k)) + \beta \sum_{j \in S_w} F(j \mid k, \pi_w(k)) V_w(\pi_w, j). \end{aligned}$$

Here $k \in S_w, \pi_w \in \Pi_w, \Pi_w$ is the set of all Markov policies. Denote $V_w(\pi_w^*, k)$ as the optimal expected value with initial state k, π_w^* is the optimal policy accordingly, then

$$V_w(\pi_w^*, k) = \min_{\pi_w \in \Pi_w} V_w(\pi_w, k) = \min_{\pi_w \in \Pi_w} \{g_w(k, \pi_w(k)) + \beta \sum_{j \in S_w} F(j \mid k, \pi_w(k)) V_w(\pi_w, j)\}. \quad (6)$$

5 Modified Policy Iteration algorithm

To find the optimal solutions of equations (3) and (6), we apply the Modified Policy Iteration algorithms with action elimination procedures which can obtain the approximation optimal action for each state (Puterman and Shin, 1982). The algorithm is represented by Figure 3 and the detail is presented in Appendix 2.

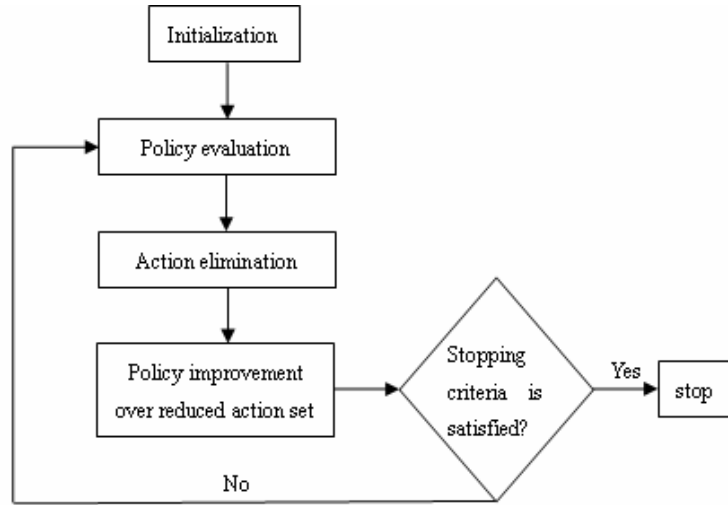


Figure 3: Modified Policy Iteration Algorithms with action elimination

6 Comparative analysis

6.1 Advantage of the proposed system

In the proposed system, the ordering decision and the delivery process are integrated and performed by the central warehouse, so the subsidiaries can concentrate on production, without

paying more attention on coals ordering, transportation and delivery decisions. In addition, as the JIT mode is introduced, the subsidiaries need not to maintain huge inventory, and the corresponding inventory cost can also be reduced. So the benefits of the proposed system to the subsidiaries are obvious.

Besides, the advantages of the new plan also include: (1) The order is bigger due to the integration of orders of the four subsidiaries, making the supply of the coal more stable. (2) All orders and deliveries for the subsidiaries will be taken into integrated consideration, realizing the scale of economy both in transportation and ordering processes. (3) As the unit inventory cost at the warehouse is lower than that at the subsidiaries, and the integration of coal can reduce the quantities of safety inventory, the transfer of the inventory from the subsidiaries to the central warehouse can reduce the inventory cost greatly. (4) The Just-in-Time mode is introduced in the delivery process from WHP to the subsidiaries, effectively lowering the uncertainty faced by the subsidiaries.

However, as establishment of the central warehouse and the ownership of vessels fleet are costly, quantity analysis between the current and the proposed systems are needed. In this section, by operating the algorithms for MDP models, we compare the costs between the two systems.

6.2 The costs in the proposed system

Tables 1 and 2 list the cost and demand parameters regard to the subsidiaries and the central warehouse respectively (some of them are estimated due to the unaccessibility to the real data). By using the MDP model and the MPI algorithm, we determine the optimal delivery decisions for each subsidiary and the optimal ordering decision of the central warehouse, along with the long-run discounted costs. The algorithm is coded in MATLAB software and run on a Pentium 4 computer with 1400MHz processor and 256 MB of RAM. The parameters for computation are taken as: $m = 5, \varepsilon = 0.5, \rho = 0.8$ (See explanation of these parameters in Appendix 2). The computational results are listed in Table 3 and Table 4.

Table 1: Parameters of the subsidiary companies

i	h_i	q_i	d_i	D_i^{max}	D_i^{min}	μ_i	c_v	C_v
1	32,000	86,000	56	35	15	25	90	975
2	32,000	86,000	81	28	10	19	90	975
3	32,000	86,000	81	6	1	4	90	975
4	32,000	86,000	90	7	1	5	90	975

$i = 1$ for YA; 2 for YI; 3 for JI; 4 for AN

Table 2: Parameters of the central warehouse

h_w	q_w	l	D_w^{max}	D_w^{min}	μ_w	c_n	C_w	C_f	C_V
15,000	50,000	2	123	89	106	8,000	1,000,000	100,000	3,600,000

In Table 3, optimal delivery quantity for any state of each subsidiary company is given, along with the long run discounted cost when the state is taken as the initial value. It can be seen from the computational results that, for each subsidiary company, the sum of the inventory level and the delivery quantity vary around some area, which is [27,28] for YA; [20,22] for YI; [5] for JI and [6,7] for AN. The sum of the inventory level and the delivery quantity is not a constant indicates that, the constant Base-Stock level policy is not suitable

Table 3: Optimal MDP delivering policies and the long-run discounted costs ($\times 10^6$)

Inventory level	Delivery amount				Long-run discounted cost			
	YA	YI	JI	AN	YA	YI	JI	AN
0	27	21	5	6	1.786	1.752	0.621	0.725
1	27	21	4	6	1.780	1.747	0.605	0.723
2	25	18	3	4	1.775	1.734	0.510	0.708
3	24	18	2	3	1.764	1.721	0.589	0.691
4	24	18	1	3	1.758	1.716	0.574	0.689
5	22	15	0	1	1.753	1.703	0.559	0.674
6	21	15	-	0	1.742	1.689	-	0.657
7	21	15	-	-	1.736	1.685	-	-
8	19	12	-	-	1.730	1.672	-	-
9	18	12	-	-	1.719	1.659	-	-
10	18	12	-	-	1.714	1.654	-	-
11	16	9	-	-	1.708	1.641	-	-
12	15	9	-	-	1.697	1.626	-	-
13	15	9	-	-	1.691	1.623	-	-
14	13	6	-	-	1.686	1.610	-	-
15	12	6	-	-	1.675	1.596	-	-
16	12	6	-	-	1.669	1.592	-	-
17	10	3	-	-	1.664	1.579	-	-
18	9	3	-	-	1.653	1.565	-	-
19	9	-	-	-	1.647	-	-	-
20	7	-	-	-	1.642	-	-	-

for our problem in which transportation cost is limited by the capacity of the vessel and is none-linearly changed.

In Table 4, optimal ordering quantity for any state is given, along with the long run discounted cost when the state is taken as the initial value. Different from the computational results in Table 3, the sum of the inventory level and the delivery variable is a constant value, which is 113 here. As the fixed costs ($C_w + C_f$) happen at each decision period no matter whether coal is supplied by the HMGC or not, and the variable transportation cost is linearly changed along with the increment of the decision variable L , the optimal decisions follow the constant Base-Stock level strategy.

6.3 Comparison of costs of the two systems

Currently, each subsidiary just orders the coal according to predication and experience. In addition, without the central warehouse and the self-owned vessels, the lead time from the supplier HMGC to every subsidiary is quite long. Therefore, each subsidiary has to order and hold a larger stock to avoid shortages. Based on the yearly supply contract made between the subsidiaries and HMGC, the estimated cost structure can be found as given in Table 5. Here the approximation of long-run discounted cost is computed by summing up the geometric sequence of the period total cost with discount rate $\beta = 0.8$.

Table 6 lists the comparison of the costs between the current policy and proposed policy. Under the proposed policy, the long-run cost of each subsidiary is the mean value of that at every initial inventory level in Table 3. It can be seen from Table 6 that, for each subsidiary,

Table 4: Optimal MDP ordering policy and its long-run discounted cost ($\times 10^6$)

k	L	$V_w(\pi_w^*, k)$	k	L	$V_w(\pi_w^*, k)$	k	L	$V_w(\pi_w^*, k)$
0	113	8.648	24	89	8.552	48	65	8.456
1	112	8.644	25	88	8.548	49	64	8.452
2	111	8.640	26	87	8.544	50	63	8.448
3	110	8.636	27	86	8.540	51	62	8.444
4	109	8.632	28	85	8.536	52	61	8.440
5	108	8.628	29	84	8.532	53	60	8.436
6	107	8.624	30	83	8.528	54	59	8.432
7	106	8.620	31	82	8.524	55	58	8.428
8	105	8.616	32	81	8.520	56	57	8.424
9	104	8.612	33	80	8.516	57	56	8.420
10	103	8.608	34	79	8.512	58	55	8.416
11	102	8.604	35	78	8.508	59	54	8.412
12	101	8.600	36	77	8.504	60	53	8.408
13	100	8.596	37	76	8.500	61	52	8.404
14	99	8.592	38	75	8.496	62	51	8.400
15	98	8.588	39	74	8.492	63	50	8.396
16	97	8.584	40	73	8.488	64	49	8.392
17	96	8.580	41	72	8.484	65	48	8.388
18	95	8.576	42	71	8.480	66	47	8.384
19	94	8.572	43	70	8.476	67	46	8.380
20	93	8.568	44	69	8.472	68	45	8.376
21	92	8.564	45	68	8.468	69	44	8.372
22	91	8.560	46	67	8.464	70	43	8.368
23	90	8.556	47	66	8.460	—	—	—

the cost under the current policy is much higher than that under the proposed policy. It is because each subsidiary should incur all kinds of cost under the current policy. On the other hand, the transportation cost from HMGC to WHP as well as most of the inventory cost for all subsidiaries are accounted from the point of the central warehouse under the proposed policy. As the standard deviation of the demand faced by the central warehouse is lower than the sum of that of the four subsidiary companies, and also the decision made by the central warehouse can achieve scale of economics and integrated optimization, the transportation and the inventory costs can be reduced greatly. The total cost for the whole system can be reduced after establishing the central warehouse and operating a fleet of self-owned vessels.

7 Conclusion

This paper presents a practical logistics problem confronted by Sinopec, a state-owned petroleum and chemical corporation in China, and its subsidiaries along the Yangtze River. We recommend VMI concept to the parent company by establishing a central warehouse at WHP to restrain the upstream wave in the supply chain and serve as a buffer for stabilizing the coal provision. The central warehouse is responsible for delivery decisions for the subsidiaries and ordering decision of itself. By operating a fleet of self-owned vessels, the JIT mode is adopted to the delivery processes.

Table 5: Current periodic cost structure of the subsidiary companies ($\times 10^5$)

i	Holding	Transportation	Shortage	Total	Approximated long-run cost
1	10.6	5.072	0	15.672	78.361
2	8.5	4.644	2	15.144	75.722
3	2.1	1.102	0	3.162	15.811
4	2.0	1.167	1	4.117	20.585

Table 6: System cost comparison between the two policies ($\times 10^5$)

i	Under the current policy	Under the proposed policy
1	78.361	17.137
2	75.722	16.614
3	15.811	5.882
4	20.590	6.917
CW+Vessels	-	85.050+36.0
Total	190.479	167.617

To find the optimal delivery and ordering decisions, the MDP models are developed, in which inventory and transportation decisions are integrated concurrently to minimize the long-run expected cost under discount criterion including transportation cost, expected inventory cost and expected shortage cost. We apply the Modified Policy Iteration algorithm with an action elimination procedure to solve the MDP models and approximately obtain the optimal solutions.

With a set of data, the long-run expected cost of the whole system under the proposed strategy with the current one is compared. It is shown that the proposed VMI strategy overcomes the current one by reducing the cost.

Even though the computational results are gained under some assumption such as the demand characters and the cost structures, it can be deduced from qualitative and quantitative analysis as shown in this paper that, the advantage of utilizing VMI concepts and scientific management decisions are obvious. It is concluded that this case study provides a good guide not only for Sinopec, also for the companies faced similar problems as well.

Appendix 1. Discussion of the distribution of the demand faced by the central warehouse

In the optimal delivery decision, denote $Z_i^t, i = 1, \dots, 4$, as the sum of the quantities received and on hand of subsidiary i at time point tT . If there are not fixed costs, based on Heyman and Sobel (1984), $Z_i^t = Z_i, t = 1, \dots, \infty$.

In the delivery decision process, the transportation cost depends on the capacity of the vessel and the transportation times. So, for achieving the scale of economics in transportation, the optimal Z_i^t may not be a stable value and varies around some area.

Denote a_i^t as the quantity delivered to subsidiary i at time period $[tT, (t+1)T)$, according to Heyman and Sobel (1984), we have

$$\begin{aligned} a_i^t &= Z_i^t - Z_i^{t-1} + D_i^t, \\ a_i^{t-1} &= Z_i^{t-1} - Z_i^{t-2} + D_i^{t-1}. \end{aligned}$$

So,

$$a_i^t + a_i^{t-1} = Z_i^t - Z_i^{t-2} + D_i^t + D_i^{t-1}. \quad (7)$$

It can be deduced from qualitative analysis the mean of $(Z_i^t - Z_i^{t-2})$ is 0. The reasons are as follows: (i) In our case study, the final delivery strategy is taken as stable, that is, the quantity delivered depends only on the inventory level, having none relationship with the time period when it is made; (ii) Z_i^t relies on the demand faced by subsidiary i , which is *i.i.d.* over an infinite time horizon. The second item also indicates that, Z_i^t is correlative with D_i^t .

Based on above analysis, it can be concluded from formulation (7) the distribution of $(a^t + a^{t-1})$ approximately follows $(D_i^t + D_i^{t-1})$, and the conclusion is suitable to each subsidiary. As $\sum_{i=1}^4 \sum_{t=1}^2 D_i^t$ follows Poisson distribution with mean $\sum_{i=1}^4 2\mu_i$, it is reasonable to assume that $\sum_{i=1}^4 \sum_{t=1}^2 a_i^t$ also follows Poisson distribution with mean $\sum_{i=1}^4 2\mu_i$.

Appendix 2. Modified Policy Iteration algorithm

The algorithm is described to solve MDP model for delivery decision. The algorithm of MDP model for ordering decision is the same as this one except the difference of the parameters given.

Step 1. Initialization. Select $\varepsilon > 0$, to be used as a stopping criterion. Denote integers n and m be the iteration stage and the order of the algorithm, respectively. Set $n = 0$ and integer $m > 0$. Define policy π_0 and single period reward $g_0(ki)$ by $g_0(k) = g(i, \pi_0(k)) = \min_{a \in A(k)} g(k, a)$, $k \in S$. Define $E_n(k)$ to be the set of actions that have been eliminated in state k at iteration stage n and set $E_0(k) = \emptyset$. Set $V_0 = \frac{1}{1-\beta} [\max_{k \in S} g_0(k)] \times e$, where e denotes a column vector of ones.

Step 2. Evaluation Phase. Calculate $V_{n+1}(k) = V_n(k) + \sum_{t=0}^m \beta^t P^t(k, \pi_n(k))(TV_n(k) - V_n(K))$, $k \in S$, where $V_n(k)$ represents $V(\pi_n, k)$, which is the total expected value at iteration stage n with initial state k (Here π_n is composed by $a \in A(k) - E_n(k)$, $k \in S$). $TV_n(k)$ and $T_{\pi_n}(K)$ are two operators defined as:

$$T_{\pi_n}(k) = g(k, \pi_n) + \beta \sum_{j \in S} P_{kj}(\pi_n) V_n(j), \quad k \in S, \pi_n \in \Pi$$

$$TV_n(k) = \min_{\pi_n \in \Pi} T_{\pi_n} V(k) = \min_{a \in A(k) - E_n(k)} \{g(k, a) + \beta \sum_{j \in S} P_{kj}(a) V_n(j)\}, \quad k \in S$$

Increment n by one, i.e. $n = n + 1$.

Step 3. Action Elimination. Suppose that at iteration $n + 1$ and for $\forall a \in A(i) - E_n(i)$,

$$g(k, a) + \beta \sum_{j \in S} P_{kj}(a) V_n(j) + \beta U(DV_{n,1}) - V_{n+1}(k) > \beta L(PHV_n),$$

then $a \in E_{n+1}(k)$. Here:

$$\begin{aligned} DV_{n,1}(k) &= V_{n+1}(k) - V_n(k), \\ U(DV_{n,1}) &= \min_{k \in S} DV_{n,1}(k), \\ L(PHV_n) &= \max_{k \in S} \{(\beta P_n)^m [TV_n(k) - V_n(k)]\}, \end{aligned}$$

where P_n denotes the state-transfer probability matrix used at iteration n , that is $G(k, \pi_n)$.

Step 4. Improvement Phase. Find policy π_n composed of $a \in \{A(k) - E_{n-1}(k)\}, k \in S$:

$$TV_n = g_{\pi_n} + \beta P(\pi_n)V_n = \min_{\pi \in \Pi} [g_{\pi} + \beta P(\pi)V_n].$$

If $\rho(V_n - TV_n) = \max(V_n - TV_n) < \varepsilon$, go to Step 5. Otherwise, return to Step 2.

Step 5. Final Extrapolation. Set $V' = TV_n + \beta(1 - \beta)^{-1} \max_{i \in S} [TV_n(i) - V_n(i)]e$ as approximation of V_{β}^* , π_n is $\beta(1 - \beta)^{-1}$ optimal.

References

- [1] Andel, T., 1996. Manage inventory, own information. *Transportation and Distribution* 37, 54–58.
- [2] Angulo, A., Nachtmann, H., Waller, M.A., 2004. Supply chain information sharing in a vendor managed inventory partnership. *Journal of Business Logistics* 25, 101–125.
- [3] Aviv, Y., Federguen, A., 1998. The operational benefits of information sharing and vendor managed inventory (VMI) programs, Technical Report, Columbia University, New York.
- [4] Cardos, M., Garcia-Sabater J.P., 2006. Designing a consumer products retail chain inventory replenishment policy with the consideration of transportation costs. *International Journal of Production Economics* 104, 525–535.
- [5] Cheung, K.L., Lee, H.L., 2002. The inventory benefit of shipment coordination and stock rebalancing in a supply chain. *Management Science* 48, 300–306.
- [6] Clark, A., Scarf, H., 1960. Optimal policies for a multi-echelon inventory problem. *Management Science* 6, 475–490.
- [7] Ganeshan, R., 1999. Managing supply chain inventories: a multiple retailer, one warehouse, multiple supplier model. *International Journal of Production Economics* 59, 341–354.
- [8] Giannoccaro, I., Pontrandolfo, P., 2002. Inventory management in supply chains: a reinforcement learning approach. *International Journal of Production Economics* 78, 153–161.
- [9] Hartanto, W., Dirk, C., Dirk V.O., 2005. Stocking decisions for repairable spare parts pooling in a multi-hub system. *International Journal of Production Economics* 93-94, 309–317.
- [10] Heyman, D., Sobel. M.J., 1984. *Stochastic Models in Operations Research, Vol II*. McGraw-Hill, New York.
- [11] Howard, R.A., 1960. *Dynamic programming and Markov Processes*. MIT Press, Cambridge, Massachusetts.
- [12] Lee, H.L., So, K.C., Tang, C.S., 2000. The value of information sharing in a two-level supply chain. *Management Science* 46, 626–643.
- [13] Levi, R., Pal, M., Roundy, R.O., Shmoys, D.B., 2007. Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research* 32, 284–302.
- [14] Puterman, M.L., Shin, M.C., 1978. Modified policy iteration algorithms for discounted Markov decision problems. *Management Science* 24, 1127–1137.

- [15] Puterman, M.L., Shin, M.C., 1982. Action elimination procedures for modified policy iteration algorithms. *Operations Research* 30, 301–318.
- [16] Puterman, M.L., 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley, New York.
- [17] Qu, W.W., Bookbinder, J.H., Iyogun, P., 1999. An integrated inventory-transportation system with modified periodic policy for multiple products. *European Journal of Operational Research* 115, 254–269.
- [18] Silver, E.A., Pyke, D.F. and Peterson, R., 1998. *Inventory Management and Production Planning and Scheduling*, 3rd Edition, John Wiley and Sons, Chichester.
- [19] Simchi-Levi, D., Zhao, Y., 2004. The value of information sharing in a two-stage supply chain with production capacity constraints: the infinite horizon case. *Probability in the Engineering and Informational Sciences* 18, 247–274.
- [20] Yin, K.K., Yin, G.G. and Liu, H., 2004. Stochastic modeling for inventory and production planning in the paper industry. *AIChE Journal* 50, 2877–2890.