Sensitivity Analysis of a Dynamic Vehicle Allocation Policy Using Approximate Dynamic Programming and Applications to Fleet Sizing

Huseyin Topaloglu* Warren B. Powell†

*School of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14853, USA
topaloglu@orie.cornell.edu

†Department of Operations Research and Financial Engineering
Princeton University
Princeton, NJ 08544, USA
powell@princeton.edu

1 Introduction

Given a vehicle fleet and a stochastic process characterizing the load arrivals in a transportation network, the primary objective of the fleet management models is to make the vehicle repositioning and vehicle-to-load assignment decisions so that some performance measure (profit, cost, deadhead miles, number of served loads, etc.) is optimized. However, besides making these vehicle allocation and assignment decisions, a very important question that is commonly overlooked by many fleet management models is how the performance measures would change in response to a change in certain model parameters. For example, freight carriers are interested in how much their profits would increase if they introduce an additional vehicle into the system or if they serve an additional load on a certain traffic lane. Railroad companies want to estimate the minimum number of railcars that is necessary to cover random shipper demands. Airlift Mobility Command is interested in the impact of limited airbase capacities on the delayed shipments. Answering such questions requires sensitivity analysis of the underlying fleet management model responsible for making the vehicle allocation decisions.

In this paper, we develop efficient sensitivity analysis methods for a stochastic fleet management model previously developed by Godfrey & Powell (2002). This model formulates the problem as a dynamic program, decomposing it into time-staged subproblems, and replaces the value functions with specially-structured approximations that are obtained through an iterative improvement scheme. Here, we develop methods to compute how much the profits would increase if an additional vehicle or an additional load is introduced into the system, and show how to apply these methods for fleet sizing.

2 Problem Formulation

We have a fleet of vehicles to serve the loads of different types occurring over time. At every time period, a certain number of loads enter the system, and we have to decide which loads to cover and to which locations we should reposition the empty vehicles. We are interested in maximizing the total profit over a finite horizon. For brevity, we assume that all travel times are single time period and the fleet is homogenous. We define the following:

$\mathcal{T}$ = Set of time periods in the planning horizon, $\mathcal{T} = \{1, \ldots, T\}$.

$\mathcal{I}$ = Set of locations in the transportation network.

$\mathcal{L}$ = Set of movement modes using which a vehicle can move from one location to another, $\mathcal{L} = \{0, \ldots, L\}$. Movement mode 0 always corresponds to empty repositioning, other modes correspond to carrying different types of loads.

x_{ijlt} = Number of vehicles dispatched from location i to j at time period t using movement mode l.

c_{ijlt} = Cost (negative profit) of dispatching one vehicle from location i to j at time period t using movement mode l.

D_{ijlt} = Random variable representing the number of loads that need to be carried from location i to j at time period t and correspond to movement mode l.

r_{it} = Number of vehicles at location i at time period t.

In practice, the movement modes in $\mathcal{L} \setminus \{0\}$ may correspond to different types of loads or different shippers, and usually $c_{ijlt} < 0$ when $l \in \mathcal{L} \setminus \{0\}$. Since the movement mode 0 corresponds to empty repositioning, we assume $c_{ij0t} \geq 0$ and $D_{ij0t} = \infty$ for all $i, j \in \mathcal{I}$, $t \in \mathcal{T}$. We use x_t, c_t, D_t, D and r_t to denote the vectors $\{x_{ijlt} : i, j \in \mathcal{I},\ l \in \mathcal{L}\}$, $\{c_{ijlt} : i, j \in \mathcal{I},\ l \in \mathcal{L}\}$, $\{D_{ijlt} : i, j \in \mathcal{I},\ l \in \mathcal{L}\}$, $\{D_t : t \in \mathcal{T}\}$ and $\{r_{it} : i \in \mathcal{I}\}$ respectively. We note that r_t completely defines the state of the vehicles necessary to make the decisions at time period t. Then, the set of feasible decision vectors at time period t is given by

$$\mathcal{X}(r_t, D_t) = \{\ x_t \in \mathbb{Z}_+^{|\mathcal{I}|^2|\mathcal{L}|} :\ \sum_{j \in \mathcal{I}} \sum_{l \in \mathcal{L}} x_{ijlt} = r_{it} \qquad \text{for all } i \in \mathcal{I} \tag{1}$$

$$x_{ijlt} \leq D_{ijlt} \qquad \text{for all } i, j \in \mathcal{I},\ l \in \mathcal{L}\ \}. \tag{2}$$

Given x_t and r_t, the state vector at time period $t+1$ is defined by

$$r_{j,t+1} = \sum_{i \in \mathcal{I}} \sum_{l \in \mathcal{L}} x_{ijlt} \qquad \text{for all } j \in \mathcal{I}. \tag{3}$$

We also define

$$\mathcal{Y}(r_t, D_t) = \{\ (x_t, r_{t+1})\ :\ x_t \in \mathcal{X}(r_t, D_t)$$
$$r_{j,t+1} = \sum_{i \in \mathcal{I}} \sum_{l \in \mathcal{L}} x_{ijlt} \text{ for all } j \in \mathcal{I}\ \}.$$

We are interested in finding a Markovian deterministic policy that minimizes the expected cost over the planning horizon. Such a policy π can be characterized by a sequence of *decision* *functions* $\{X_t^\pi(\cdot,\cdot) : t \in \mathcal{T}\}$, such that each $X_t^\pi(\cdot,\cdot)$ maps the state vector r_t and the realization of the loads D_t for time period t to a decision vector x_t. One can also define the *state transition functions* $\{R_{t+1}^\pi(\cdot,\cdot) : t \in \mathcal{T}\}$ of policy π, such that each $R_{t+1}^\pi(\cdot,\cdot)$ maps the state vector and the realization of the loads at time period t to a state vector for the next time period. (Given $X_t^\pi(\cdot,\cdot)$, $R_{t+1}^\pi(\cdot,\cdot)$ can easily be defined by noting (3).) Then, for a given state vector r_t and realization of future demands $\{D_t,\ldots,D_T\}$ at time period t, the *cumulative cost function* for policy π can be written recursively as

$$F_t^\pi(r_t, D_t, D_{t+1}, \ldots, D_T) = c_t \cdot X_t^\pi(r_t, D_t) + F_{t+1}^\pi\left(R_{t+1}^\pi(r_t, D_t), D_{t+1}, D_{t+2}, \ldots, D_T\right), \quad (4)$$

with $F_{T+1}(\cdot,\cdot) = 0$. The optimal policy π^* satisfies $\pi^* = \arg\min_\pi \; \mathbb{E}\{\; F_1^\pi(r_1, D_1, \ldots, D_T)\;\}$, and can be found by computing the *value functions* through the functional equation

$$V_t^{\pi^*}(r_t) = \mathbb{E}\left\{ \min_{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, D_t)} c_t \cdot x_t + V_{t+1}^{\pi^*}(r_{t+1}) \mid r_t \right\}. \quad (5)$$

In this case, the decision and transition function for the optimal policy become

$$\left(X_t^{\pi^*}(r_t, D_t),\; R_{t+1}^{\pi^*}(r_t, D_t)\right) = \operatorname*{arg\,min}_{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, D_t)} c_t \cdot x_t + V_{t+1}^{\pi^*}(r_{t+1}). \quad (6)$$

(6) also shows that one can obtain different (probably suboptimal) policies by replacing $V_{t+1}^{\pi^*}(\cdot)$ with different functions. In this paper, we follow a class of policies obtained by replacing $\{V_t^{\pi^*}(\cdot) : t \in \mathcal{T}\}$ with separable functions $\{V_t^\pi(\cdot) : t \in \mathcal{T}\}$ of the form $V_t^\pi(r_t) = \sum_{i \in \mathcal{I}} V_{it}^\pi(r_{it})$, where each of $V_{it}^\pi(\cdot)$ is a one-dimensional, piecewise-linear, convex function. Godfrey & Powell (2002) give an iterative, sampling-based algorithm that can be used to obtain a "good" set of approximations, and their experimental work indicates that these approximations yield very high quality solutions. In this paper, we are concerned with estimating the change in $F_1^\pi(r_1, D_1, \ldots, D_T)$ induced by changing an element of the state vector r_1 or the load availability vector D_1.

3 Decision and Transfer Function for Our Class of Policies

Letting π be a policy characterized by the set of separable, piecewise-linear, convex value function approximations $\{V_t^\pi(\cdot) : t \in \mathcal{T}\}$, we define the decision and state transition function for this policy as

$$\left(X_t^\pi(r_t, D_t),\; R_{t+1}^\pi(r_t, D_t)\right) = \operatorname*{arg\,min}_{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, D_t)} c_t \cdot x_t + V_{t+1}^\pi(r_{t+1}). \quad (7)$$

It can be shown that problem (7) is a min-cost network flow problem when the value function approximations are separable, piecewise-linear, convex functions with points of nondifferentiability being a subset of positive integers.

4 Policy Gradients with respect to Vehicle Availabilities

For a realization of loads $D = \{D_t : t \in \mathcal{T}\}$, we let $\left\{x_t^{\pi D} : t \in \mathcal{T}\right\}$ and $\left\{r_t^{\pi D} : t \in \mathcal{T}\right\}$ be the sequence of decisions and states visited by the system under policy π and load realization D.

That is, $\left\{x_t^{\pi D} : t \in \mathcal{T}\right\}$ and $\left\{r_t^{\pi D} : t \in \mathcal{T}\right\}$ are recursively computed by

$$x_t^{\pi D} = X_t^{\pi}\left(r_t^{\pi D}, D_t\right), \quad r_{t+1}^{\pi D} = R_{t+1}^{\pi}\left(r_t^{\pi D}, D_t\right), \text{ with } r_1^{\pi D} = r_1. \tag{8}$$

In this section, we develop an algorithm that computes how much the total cost under policy π would change if an additional vehicle is introduced into the system. That is, we are interested in computing

$$\Phi_t^{\pi}\left(e_i, D\right) = F_t^{\pi}\left(r_t^{\pi D} + e_i, D_t, \ldots, D_T\right) - F_t^{\pi}\left(r_t^{\pi D}, D_t, \ldots, D_T\right), \tag{9}$$

where e_i is the $|\mathcal{I}|$ dimensional unit vector with a 1 in the element corresponding to $i \in \mathcal{I}$.

We note that $\Phi_t^{\pi}(e_i, D)$ can be computed by two simulations of policy π under load realization D, one of which starts with the state vector $r_t^{\pi D}$ and the other with $r_t^{\pi D} + e_i$. However, in general, doing this for all $i \in \mathcal{I}$ and for multiple load realizations can get very time consuming. Our objective is to be able to compute $\Phi_t^{\pi}(e_i, D)$ for all $i \in \mathcal{I}$ from a single simulation. Using (4), (9) can be written as

$$\begin{aligned}\Phi_t^{\pi}\left(e_i, D\right) = c_t \cdot \{\ & X_t^{\pi}\left(r_t^{\pi D} + e_i, D_t\right) - x_t^{\pi D}\ \} \\ & + F_{t+1}^{\pi}\left(R_{t+1}^{\pi}\left(r_t^{\pi D} + e_i, D_t\right), D_{t+1}, \ldots, D_T\right) - F_{t+1}^{\pi}\left(r_{t+1}^{\pi D}, D_{t+1}, \ldots, D_T\right).\end{aligned}$$

Therefore, computing $\left\{X_t^{\pi}\left(r_t^{\pi D} + e_i, D_t\right) - x_t^{\pi D}\right\}$ and $\left\{R_{t+1}^{\pi}\left(r_t^{\pi D} + e_i, D_t\right) - r_{t+1}^{\pi D}\right\}$ is key to computing $\Phi_t^{\pi}(e_i, D)$ and these quantities are related to how the solution of the min-cost network flow problem (7) changes when the right side of constraints (1) is increased from $r_t^{\pi D}$ to $r_t^{\pi D} + e_i$. For this purpose, we use a well-known relationship between the perturbations of min-cost network flow problems and min-cost flow augmenting trees (see, for example, Powell (1989)).

Proposition 1 $X_t^{\pi}\left(r_t^{\pi D} + e_i, D_t\right) = x_t^{\pi D} + \xi_t^{\pi}(e_i, D)$, $R_{t+1}^{\pi}\left(r_t^{\pi D} + e_i, D_t\right) = r_{t+1}^{\pi D} + \delta_{t+1}^{\pi}(e_i, D)$, *where $\xi_t^{\pi}(e_i, D)$ and $\delta_{t+1}^{\pi}(e_i, D)$ can be computed for all $i \in \mathcal{I}$ by one min-cost flow augmenting tree computation. Furthermore, exactly one element of the vector $\delta_{t+1}^{\pi}(e_i, D)$ is equal to $+1$ and all the other elements of this vector are equal to 0.*

Then, the following result gives an efficient algorithm to compute $\Phi_t^{\pi}(e_i, D)$ for all $i \in \mathcal{I}$ and $t \in \mathcal{T}$.

Proposition 2 *$\Phi_t^{\pi}(e_i, D)$ can be computed for all $i \in \mathcal{I}$ and $t \in \mathcal{T}$ by the backward recursion*

$$\Phi_t^{\pi}(e_i, D) = c_t \cdot \xi_t^{\pi}(e_i, D) + \Phi_{t+1}^{\pi}\left(\delta_{t+1}^{\pi}(e_i, D), D\right).$$

We also remark that using the properties of min-cost flow decreasing trees, a similar procedure can be developed to compute $\Phi_t^{\pi}\left(-e_i, D\right) = F_t^{\pi}\left(r_t^{\pi D} - e_i, D_t, \ldots, D_T\right) - F_t^{\pi}\left(r_t^{\pi D}, D_t, \ldots, D_T\right)$.

5 Policy Gradients with respect to Load Availabilities

In this section, we construct an algorithm that is useful to assess how much the total cost under policy π would change if an additional load is introduced into the system. More precisely, letting e_{ijl} be the $|\mathcal{I}|^2|\mathcal{L}|$ dimensional unit vector with a 1 in the element corresponding to $i, j \in \mathcal{I}$, $l \in \mathcal{L}$, we want to compute

$$\begin{aligned} \Psi_t^\pi(e_{ijl}, D) &= F_t^\pi \left(r_t^{\pi D}, D_t + e_{ijl}, \ldots, D_T\right) - F_t^\pi \left(r_t^{\pi D}, D_t, \ldots, D_T\right) \\ &= c_t \cdot \left\{ X_t^\pi \left(r_t^{\pi D}, D_t + e_{ijl}\right) - x_t^{\pi D}\right\} \\ &\quad + F_{t+1}^\pi \left(R_{t+1}^\pi \left(r_t^{\pi D}, D_t + e_{ijl}\right), D_{t+1}, \ldots, D_T\right) - F_{t+1}^\pi \left(r_{t+1}^{\pi D}, D_{t+1}, \ldots, D_T\right). \end{aligned} \tag{10}$$

In order to compute the quantity above, we now need to characterize $\left\{X_t^\pi \left(r_t^{\pi D}, D_t + e_{ijl}\right) - x_t^{\pi D}\right\}$ and $\left\{R_{t+1}^\pi \left(r_t^{\pi D}, D_t + e_{ijl}\right) - r_{t+1}^{\pi D}\right\}$. These quantities are related to how the solution of the min-cost network flow problem (7) changes when the upper bounds on the variable x_t is increased from D_t to $D_t + e_{ijl}$. The following result characterizes this change.

Proposition 3 $X_t^\pi \left(r_t^{\pi D}, D_t + e_{ijl}\right) = x_t^{\pi D} + \zeta_t^\pi(e_{ijl}, D)$, $R_{t+1}^\pi \left(r_t^{\pi D}, D_t + e_{ijl}\right) = r_{t+1}^{\pi D} + \eta_{t+1}^\pi(e_{ijl}, D)$, *where $\zeta_t^\pi(e_{ijl}, D)$ and $\eta_{t+1}^\pi(e_{ijl}, D)$ can be computed for all $j \in \mathcal{I}$, $l \in \mathcal{L}$ by one min-cost flow augmenting tree computation. Furthermore, $\eta_{t+1}^\pi(e_{ijl}, D)$ can be written as $\eta_{t+1}^\pi(e_{ijl}, D) = \eta_{t+1}^{\pi+}(e_{ijl}, D) - \eta_{t+1}^{\pi-}(e_{ijl}, D)$, where in both of the vectors on right, exactly one element is equal to +1 and all the other elements are equal to 0.*

Then, the following result gives an efficient algorithm to compute $\Psi_t^\pi(e_{ijl}, D)$ for all $i, j \in \mathcal{I}$, $l \in \mathcal{L}$ and $t \in \mathcal{T}$.

Proposition 4 *If $F_{t+1}^\pi(\cdot, D_{t+1}, \ldots, D_T)$ is a separable function, then*

$$\Psi_t^\pi \left(e_{ijl}, D\right) = c_t \cdot \zeta_t^\pi(e_{ijl}, D) + \Phi_{t+1}^\pi \left(\eta_{t+1}^{\pi+} \left(-e_i, e_j, D\right), D\right)\right) + \Phi_{t+1}^\pi \left(-\eta_{t+1}^{\pi-} \left(-e_i, e_j, D\right), D\right)\right), \tag{11}$$

where $\Phi_{t+1}^\pi(\mp e_i, D)$ is as defined in (9).

In general $F_{t+1}^\pi(\cdot, D_{t+1}, \ldots, D_T)$ is not a separable function. However, our numerical experiments show that (11) is an accurate approximation even when the separability assumption does not hold.

6 Preliminary Numerical Experiments

Our first set of experiments show that (11) is an accurate approximation even when $F_{t+1}^\pi(\cdot, D_{t+1}, \ldots, D_T)$ is not a separable function. Figure 1 shows the results for one particular problem, where we compare the approximation to $\Psi_1^\pi(e_{ijl}, D)$ (obtained through (11)) with the exact value of $\Psi_1^\pi(e_{ijl}, D)$ (obtained in brute force fashion by physically adding a load of type l on lane (i, j) and simulating the behavior of policy π under load realization D). Different data points correspond to different values of $i, j \in \mathcal{I}$ and $l \in \mathcal{L}$. Figure 1 indicates that (11) is an accurate approximation when the separability assumption is not satisfied.

Our second set of experiments investigates using (9) for fleet sizing. Our general approach is along the lines of a gradient search method, where we increment the number of vehicles at location i if $\mathbb{E}\{\Phi_1^\pi(e_i, D)\}$ is less than the cost of "leasing" a vehicle at location i over the

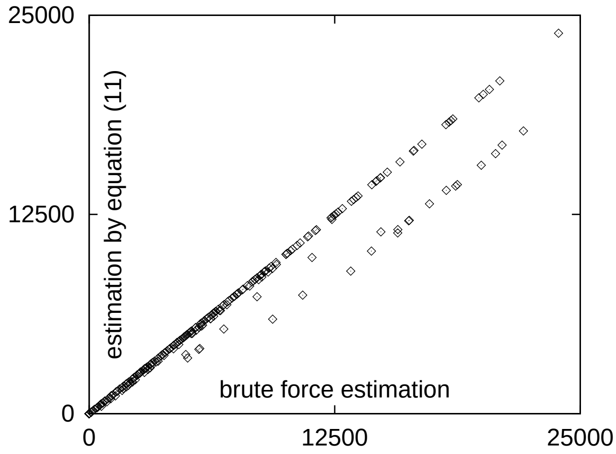

Figure 1: (11) approximates $\Psi_1^\pi(e_{ijl}, D)$ quite well when $F_2^\pi(\cdot, D_2, \ldots, D_t)$ is not separable.

Problem	1	2	3	4	5	6
% diff.	3.7	3.1	3.2	4.0	2.9	3.1

Figure 2: A gradient search-based algorithm that uses $\mathbb{E}\{\Phi_1^\pi(e_i, D)\}$ yields 3.1-4.0% better results than deterministic state-time network-based models.

planning horizon. We approximate the expectation $\mathbb{E}\{\Phi_1^\pi(e_i, D)\}$ by using 10 demand realizations and averaging $\Phi_1^\pi(e_i, D)$ over 10 realizations. As a benchmark, we use a deterministic model based on the state-time network formulation of the problem (see, for example, Sherali & Tuncbilek (1997)). Figure 2 shows that our approach yields significantly better solutions than the benchmark method on all six test problems.

References

G. A. Godfrey. and W. B. Powell, "An Adaptive, Dynamic Programming Algorithm for Stochastic Resource Allocation Problems I: Single Period Travel Times", *Transportation Science* 36, 21–39 (2002).

W. B. Powell, "A Review of Sensitivity Results for Linear Networks and a New Approximation to Reduce the Effects of Degeneracy", *Transportation Science* 23, 231–243 (1989).

H. D. Sherali and C. H. Tuncbilek, "Static and Dynamic Time-Space Strategic Models and Algorithms for Multilevel Rail-Car Fleet Management", *Management Science* 43, 235–250 (1997).