Online Supplements: The Stochastic Two-Echelon Production Routing Problem with Adaptive Routing

Ali Kermani, Jean-François Cordeau, Raf Jans
HEC Montréal, CIRRELT, and GERAD, 3000 chemin de la Côte-Sainte-Catherine, Montréal, H3T 2A7, Canada
ali.kermani@hec.ca, jean-francois.cordeau@hec.ca, raf.jans@hec.ca

This file contains supplementary tables for the paper titled "The Stochastic Production Routing Problem with Adaptive Routing and Service Level Constraints." These materials comprise detailed tables presenting results on various service level measures across different granularity levels. Specifically, Section **??**, Section **??**, and Section **??** present the results for Customer Level-Global, Plant Level-Single Period, and Plant Level-Global, respectively. For a comprehensive overview of the results, please visit `https://github.com/AliK094/online_supplements_sprpar_sl/tree/main/results_tables`.

---

1 Instance Generation

In this section, we provide details on how we generated datasets for the two-echelon production routing problem. All instances are available in the repository at (link). To generate the new instances, we used the existing benchmarks from the PRP introduced by [**?**] and further modified by [**?**]. For the stochastic instances, we also employed the logic presented by [**?**]; however, some values were adjusted to reflect the differences between the problem addressed in this study and those in previous works.

Given the large number of instances provided by [**?**], we selected a subset to generate the deterministic 2EPRP instances. We considered four classes of instances: the standard cost setting, high production costs, high transportation costs, and no customer inventory holding cost. For each class, we used the first instance with 50 customers from the original dataset. Consistent with the original setup, the number of periods is set to 6 in all instances. We generated instances with 2 to 5 warehouses and with 10 to 50 customers, in increments of 10, resulting in $4 \times 16 = 64$ instances.

We also considered three different vehicle configurations for the plant and warehouses: (i) one vehicle at the plant and one per warehouse, (ii) two vehicles at the plant and two per warehouse, and (iii) two vehicles at the plant and three per warehouse. This leads to a total of $3 \times 64 = 192$ instances for the deterministic problem.

The production capacity at the plant was set as $C = 2\sum_{i\in\mathcal{N}_c}\sum_{t\in\mathcal{T}} d_{it}/T$, and the storage capacity at the plant was defined as $L_0 = C/2$, following the settings in [**?**]. Since the original dataset does not provide information on warehouse nodes, we generated them synthetically. We first used the K-means algorithm to cluster customer nodes into w groups, where w corresponds to the number of warehouses in the instance. For each cluster, we identified the centroid and placed the corresponding warehouse randomly within a circular sector, located between 20% and 80% of the distance from the centroid to the farthest customer in that cluster.

After generating the warehouse nodes, we virtually assigned each customer to its nearest warehouse. That is, the set of customers virtually assigned to warehouse w is defined as

$$\mathcal{C}^w = \left\{ i \in \mathcal{N}_c \;\middle|\; w = \arg\min_{w'}(c_{w',i}) \right\}.$$

Based on this assignment, the warehouse storage capacity was sampled from a uniform distribution:

$$U\left[\frac{1}{T}\sum_{i\in\mathcal{C}^w}\sum_{t\in\mathcal{T}} d_{it},\; \frac{1.5}{T}\sum_{i\in\mathcal{C}^w}\sum_{t\in\mathcal{T}} d_{it}\right].$$

The inventory holding cost for warehouses was drawn from an integer uniform distribution $U[3,6]$.

Finally, the capacity of vehicles at the plant was determined using the following formula:

$$Q^p = \left\lfloor \frac{\gamma\overline{\overline{L}}}{K_p} \right\rfloor \tag{1}$$

where $\gamma = \left\lfloor \frac{N_c}{10} + 1 \right\rfloor$ and $\overline{\overline{L}} = \max_{i\in\mathcal{N}_c}\{L_i\}$. The capacity of each vehicle at the warehouses was calculated as:

$$Q^w = \max\left\{\overline{\overline{L}},\; \frac{\gamma\overline{\overline{L}}}{N_w K_w}\right\} \tag{2}$$

In this section we provide details on how we generated datasets for the two-echelon production routing problem. All instances can be found in (link) repository. To generate the new instances we used the existing benchmarks from the PRP that were introduced by cite(archetti2011) and further modified by cite(adulyasak2014). For the stochastic instances we also emplyed the logic presented by cite(adulyasak2015), however, some of the values were adjusted as the problem addressed in this study is different than the one addressed in previous studies.

As the number of instances presented by [?] is too large we used a subset of these benchmarks to generate the instances for the determinsitic 2EPRP. We used the four class of instances including the standarrd setting, with higher production costs, higher transportation costs and no cusstomer inventory holding cost. For each calss we used the first instance with 50 customers in the original dataset and similar to the original dataset the number of periods is considered to be 6 in all instances. We generated instances with 2 to 5 warehouses and 10 to 50 customer with increment of 10. This resulted in $4 \times 16 = 64$ instances. We also considered 3 different vehicle configuration for plant and warehouses. The first one was considering 1 vehicle at the plant and 1 vehicle per each warehouse, the second one was to consider 2 vehicle at the plant and 2 vehicles per each warehouse and the last configuration was 2 vehicles at the plant and 3 vehicles at each wrehouse. Thus, the total number of instances would be $3 \times 64 = 192$ for the determinstic problem.

The production capaity at the plant was considered $C = 2\sum_{i\in\mathcal{N}_c}\sum_{t\in\mathcal{T}} d_{it}/T$ and the storage capacity at the plant was set to $L_0 = C/2$ as in [?].Since the original dataset lacks the information for the warheouse nodes we need to generate these nodes. To do so, we use a K-means algorithm to cluster customer nodes into w clusters where w is the number of warehouses in that instance. Then we used the obtained centroids of each cluster and added randomness to the warehouse location by randomly locating it within a circle pie between the 20% and 80% of the distance between the center and the firthest node in the cluster. After generating the warehouse nodes, we use consider the customers in that cluster to be virtually assigned to the corresponding warehouse. Thus, assuming $\mathcal{C}^w = \{i \in \mathcal{N}_c \mid w = \arg\min_{w'}(c_{w',i})\}$ as the set of customers that are virtually assgined to warehouse w, to calculate the warehouse storage capacity we use a uniform distribution: $U[\sum_{i\in\mathcal{C}^w}\sum_{t\in\mathcal{T}} d_{it}/T, 1.5\sum_{i\in\mathcal{N}_c}\sum_{t\in\mathcal{T}} d_{it}/T]$. The inventory cost for warehouses is also set using a integer uniform distribution $U[3,6]$.

Finally to set the plant vehicle capacity we used the following equation:

$$Q^p = \left\lfloor \gamma\overline{\overline{L}}/K_p \right\rfloor \tag{3}$$

where $\gamma = \left\lfloor \frac{N_c}{10} + 1 \right\rfloor$ and $\overline{\overline{L}} = \max_{i\in\mathcal{N}_c}\{L_i\}$. The capacity of each vehicle at the warehouses is calculated as follows:

$$Q^w = \max\left\{\overline{\overline{L}}, \gamma\overline{\overline{L}}/N_w K_w\right\} \tag{4}$$