Approximations for non-stationary stochastic lot-sizing under (s,Q)-type policy

Citation for published version:
Ma, X, Rossi, R & Archibald, TW 2022, 'Approximations for non-stationary stochastic lot-sizing under (s,Q)-type policy', *European Journal of Operational Research*, vol. 298, no. 2, pp. 573-584. https://doi.org/10.1016/j.ejor.2021.06.013

Digital Object Identifier (DOI):
10.1016/j.ejor.2021.06.013

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
European Journal of Operational Research

Approximations for non-stationary stochastic lot-sizing under (s, Q)-type policy

Xiyuan Ma*[a], Roberto Rossi[a], and Thomas Welsh Archibald[a]

[a] *Business School, University of Edinburgh, Edinburgh, United Kingdom*

Abstract

This paper addresses the single-item single-stocking location non-stationary stochastic lot-sizing problem under a reorder point – order quantity control strategy. The reorder points and order quantities are chosen at the beginning of the planning horizon. The reorder points are allowed to vary with time and we consider order quantities either to be a series of time-dependent constants or a fixed value; this leads to two variants of the policy: the (s_t, Q_t) and the (s_t, Q) policies, respectively. For both policies, we present stochastic dynamic programs (SDP) to determine optimal policy parameters and introduce mixed integer non-linear programming (MINLP) heuristics that leverage piecewise-linear approximations of the cost function. Numerical experiments demonstrate that our solution method efficiently computes near-optimal parameters for a broad class of problem instances.

Keywords Inventory, (s,Q) policy, stochastic lot-sizing, non-stationary demand

1 Introduction

The non-stationary stochastic lot-sing problem is an extension of the well-known dynamic lot-sizing problem (Wagner and Whitin, 1958). In this problem, one considers a single-item single-stocking-location inventory system under a finite planning horizon and periodic review; the demand is stochastic and non-stationary. To deal with the uncertainty inherent in a stochastic lot-sizing problem, Bookbinder and Tan (1988) introduce three control strategies: the "static

*Corresponding author: Xiyuan.Ma@ed.ac.uk

uncertainty", the "static-dynamic uncertainty", and the "dynamic uncertainty", which represent different approaches for determining the timing and size of orders.

Bookbinder and Tan's control strategies are captured by various policies. The (R, Q) policy determines the inventory review schedule R and the order quantity Q before the system operates; this is the static uncertainty strategy. The (s, S) policy is the dynamic uncertainty strategy, in which the timing and size of orders are decided as late as possible, in a wait-and-see fashion, by leveraging the reorder point s, and the order-up-to level S. Scarf (1960) shows that if the holding and shortage costs are convex, the optimal policy in each period is of (s, S) type. In a static-dynamic uncertainty strategy one either fixes the order schedule at the outset, and computes the exact order quantity only when orders are issued, via suitable order-up-to-levels; or fixes the order quantities at the outset, and decides when orders are issued in a wait-and-see fashion, by relying on a reorder threshold. This leads to the (R, S) policy and (s, Q) policy (also referred to as the (r, Q) policy), respectively.

Compared to stationary demand, there are relatively few studies in the literature that consider non-stationary demand. However, in the majority of practical circumstances, demand is not only stochastic but also non-stationary.

The following works investigate the static uncertainty strategy under non-stationary demand. Sox (1997) proposes a mixed integer non-linear programming (MINLP) formulation of the dynamic lot-sizing problem with dynamic costs and develops a solution algorithm that resembles the Wagner-Whitin algorithm. This strategy is also investigated by Vargas (2009), who develops a stochastic dynamic programming model which is equivalent to a shortest path problem in a specified acyclic network. Vargas also provides a rolling horizon optimisation algorithm comprising two stages: (1) to determine optimal replenishment quantities for any sequence of replenishment points, and (2) to identify the optimal sequence of replenishment points.

For the static-dynamic uncertainty strategy, research under non-stationary demand mostly considers the (R, S) policy. Tarim and Kingsman (2004) formulates the problem as a mixed integer program (MIP). They model the total expected cost by minimising the sum of holding and ordering costs under a constraint on the probability of the closing inventory being non-negative in each time period. A method to solve this model efficiently is introduced in (Tarim et al., 2011), where the relaxation of the original MIP model is converted to a shortest path problem and implemented by branch-and-bound procedures. Tarim and Kingsman (2006) provide another MIP formulation where the objective function is obtained by the mean of a piecewise linearisation. The accuracy of the approximation can be adjusted ad libitum by introducing new breakpoints.

Özen et al. (2012) consider both penalty cost and service level and prove that the optimal policy is a base stock policy for both penalty and service-level constrained models, and also for capacity limitations and minimum order quantity requirements. More recently, Rossi et al. (2015) consider several service level measures — α service level on each period, β^{cyc} service level independently for each replenishment cycle, and the classic β service level — by adding suitable constraints that leverage the loss function and its complementary function to describe the expected total holding and penalty cost. A piecewise linearisation approach is utilized to convert the cost function from non-linear to linear form. Tunc et al. (2018) present an efficient MIP reformulation along with a dynamic cut generation approach that progressively refines the piecewise linearisation to achieve a prescribed linearisation error.

Computing (s, S) policy parameters under non-stationary demand is a challenging task. The classic Silver and Meal heuristic algorithm (Silver and Meal, 1973) for deterministic demand has been extended by Silver (1978) and Askin (1981). Silver's algorithm uses a deterministic model to calculate the number of periods that each order must cover; when this replenishment plan is known, the associated safety stocks are then determined myopically. Askin (1981) explicitly includes the cost effects of probabilistic demand in the choice of the number of periods in which to order. Bollapragada and Morton (1999) approximate the non-stationary problem via a series of stationary problems based on the method developed by Zheng and Federgruen (1991). Parameters are determined by equating the cumulative mean demand of stationary and non-stationary problems over the expected reorder cycle. Xiang et al. (2018) introduce a MINLP formulation for an (s, S) policy by applying the piecewise linearisation approximation proposed by Rossi et al. (2015). Xiang et al. (2018) also derives a heuristic algorithm with binary search. Both solution methods outperform the previous heuristics in computational efficiency in tests involving short and long planning horizons. The comparison of the two proposed algorithms shows that binary search requires significantly less time than the MINLP. Visentin et al. (2021) propose a hybrid of branch-and-bound and stochastic dynamic programming model to compute optimal (R, s, S) policy parameters.

Based on this literature survey, we note a gap in the study of non-stationary demand: no literature discusses or investigates the static-dynamic uncertainty strategy in the form of an (s, Q) policy. In this paper, we propose a new control strategy for the stochastic lot-sizing problem under non-stationary demand. Under this strategy, the reorder points s_t vary with time, and we consider two cases for the order quantities: one in which the order quantity varies with time (Q_t) and another in which the order quantity is constant (Q). This leads to two (s, Q)-type policies: the (s_t, Q_t) policy and the (s_t, Q) policy. These policies require values for s_t and Q_t (or Q) to be determined at the beginning of the planning horizon. Compared to the optimal policy introduced by Scarf (1960), which allows the order quantity to vary with inventory level and time period, the order quantity in an (s_t, Q_t) policy is only affected by the time period and applies to all inventory levels, while the order quantity in an (s_t, Q) policy is a constant value for the entire planning horizon, and does not vary with inventory level or time period.

We make the following contributions to the stochastic lot-sizing literature.

- We model the non-stationary stochastic lot-sizing problem under a static-dynamic uncertainty policy in which order quantities are determined "statically", at the beginning of the planning horizon, while reordering decisions are determined "dynamically", in a wait-and-see-fashion. *We prove that the resulting optimal policy takes the non-stationary (s, Q) form.*
- We develop a new heuristic algorithm to efficiently determine near-optimal policy parameters of the proposed (s_t, Q_t) and (s_t, Q) policies. The algorithm is composed of two steps. The first-step uses the (s, S)-policy heuristic introduced in Xiang et al. (2018) to determine the order quantities, and the second-step is based on a newly developed MILP model that applies the piecewise linearisation approach discussed in Rossi et al. (2014) to determine the order-up-to levels.
- In a comprehensive numerical study, we show that optimality gaps for the (s_t, Q_t) policy obtained via our heuristic are tighter than those of a near-optimal (R_t, S_t) policy obtained via the approach in Rossi et al. (2015). We also observe that an (s_t, Q) policy lacks flexibility and leads to substantial optimality gaps.

The rest of this paper is structured as follows. In Section 2 we introduce the problem settings and present a stochastic dynamic programming (SDP) formulation. Section 3 discusses the stochastic dynamic programming formulation of the (s_t, Q_t) and (s_t, Q) policies. We also show that the resulting optimal policies take the non-stationary (s_t, Q_t) and (s_t, Q) forms through the uniqueness of reorder points. In Section 4, we develop a heuristic algorithm to compute near-optimal policy parameters for the (s_t, Q_t) policy and discuss the application of this algorithm to the (s_t, Q) policy. A computational analysis is presented in Section 5; finally, we draw conclusions in Section 6.

2 Problem description

We consider a single-item single-location non-stationary stochastic lot-sizing problem over a planning horizon of T periods. Replenishment orders are placed and instantaneously delivered at the beginning of each time period. Each replenishment order incurs an ordering cost $c(\cdot)$ comprising a fixed ordering cost K and a linear ordering cost z proportional to the non-negative order quantity Q, where

$$c(Q) \triangleq \begin{cases} K + z \cdot Q, & Q > 0; \\ 0, & Q = 0. \end{cases} \tag{1}$$

The periods' demands d_t, for $t = 1, \cdots, T$, are independent random variables with known probability density functions $g_t(\cdot)$. Any unmet demand at the end of the period is back-ordered. At the end of each period, a linear holding cost h is incurred for each unit carried from one period to the next, and a linear penalty cost b is charged on each unit back-ordered. The expected immediate holding and penalty cost at the end of period t is expressed as

$$L_t(y) \triangleq \mathbb{E}[h \max(y - d_t) + b \max(d_t - y)], \tag{2}$$

where y denotes the inventory level after receiving the replenishment and $\mathbb{E}[\cdot]$ denotes the expectation operator.

Let $C_t(x)$ represent the expected total cost of an optimal policy over periods $t, \ldots, T$ with opening inventory level x; then the problem can be modelled as a stochastic dynamic program (Bellman, 1957)

$$C_t(x) \triangleq \min_{y \geq x} \{c(y - x) + L_t(y) + \mathbb{E}[C_{t+1}(y - d_t)]\}, \tag{3}$$

where $C_{T+1}(x) \triangleq 0$, is the boundary condition.

Scarf (1960) showed that, if $L_t(y)$ is convex, the optimal policy of the dynamic inventory problem is of an (s, S) type, where the inventory system places a replenishment to reach the order-up-to level S when the stock is found to be below the reorder point at a review point. This conclusion is based on a study of the function $G_t(y) + zy$, where

$$G_t(y) \triangleq L_t(y) + \mathbb{E}[C_{t+1}(y - d_t)], \tag{4}$$

and $G_t(y)$ represents the expected total cost over periods t to T when the opening inventory is y and no order is placed in period t. Table A1 in Appendix A summarises the notation used in this paper.

In the rest of this paper, we conduct the discussion assuming $L_t(y)$ is convex. In fact, as the holding and penalty costs used in this paper are linear, $L_t(y)$ is a weighted sum of two convex functions and hence convex. A detailed proof can be found in (Rossi et al., 2014, page 490).

Example 1. Consider a 4–period stochastic lot-sizing problem under Poisson-distributed demand with rates $d_t = \langle 20, 40, 60, 40 \rangle$. The cost parameters are $K = 100$, $z = 0$, $h = 1$ and $b = 10$. Fig. 1 illustrates the variation of $G_t(I_0)$ with $I_0 \in [0, 200]$ and no replenishment order placed in period 1, where $G_1(0) = 481$.

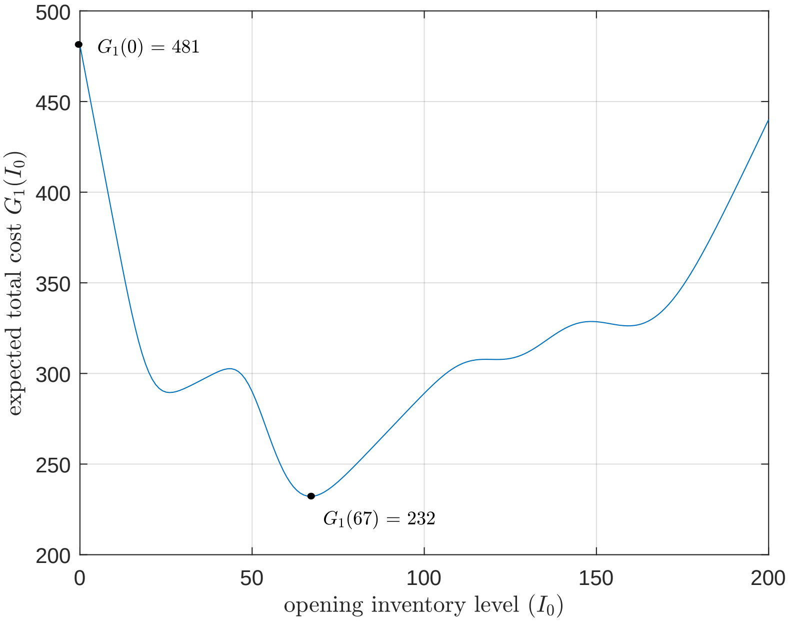

Figure 1: Plot of $G_1(I_0)$

3 Stochastic dynamic programs for the (s_t,Q_t) and (s_t,Q) policies

This section introduces the stochastic dynamic programming formulations of the stochastic lot-sizing problem under the (s_t, Q_t) policy and the (s_t, Q) policy in Section 3.1 and Section 3.2, respectively.

3.1 A stochastic dynamic program for the (s_t, Q_t) policy

An (s_t, Q_t) policy places a replenishment order of size Q_t at the beginning of period t if the inventory level is below the reorder point s_t, and does not place any order otherwise (Silver et al., 1998). The optimal expected total cost of the system controlled under an (s_t, Q_t) policy can be determined by computing all feasible combinations of reorder quantities Q_t, for $t = 1, \ldots, T$. Let $\boldsymbol{q}_t = \langle Q_t, \ldots, Q_T \rangle$ denote a $(T - t + 1)$-dimensional vector representing order quantities $Q_t, \ldots, Q_T$ and $\mathcal{Q}_t$ be the vector space representing all combinations of order quantities $\boldsymbol{q}_t$. For any $\boldsymbol{q}_t \in \mathcal{Q}_t$, the expected total cost over periods t to T when the opening inventory level is x is denoted as

$$V_t(x, \boldsymbol{q}_t) \triangleq \min_{\delta \in \{0,1\}} \{c(\delta Q_t) + L_t(x + \delta Q_t) + \mathbb{E}[V_{t+1}(x + \delta Q_t - d_t, \boldsymbol{q}_{t+1})]\}, \quad t < T, \tag{5}$$

where δ is a binary variable that represents the reordering decision in period t when the initial inventory level is x; finally,

$$V_T(x, \boldsymbol{q}_T) \triangleq \min_{\delta \in \{0,1\}} \{c(\delta Q_T) + L_T(x + \delta Q_T)\} \tag{6}$$

is the boundary condition. Therefore, considering all combinations, the optimal expected total cost when the inventory level at the beginning of the planning horizon is x can be defined as

$$V_0(x) \triangleq \min_{\boldsymbol{q}_1 \in \mathcal{Q}_1} \{V_1(x, \boldsymbol{q}_1)\}. \tag{7}$$

Let the optimal order quantity be represented by the vector $\boldsymbol{q}_t^* \triangleq \langle Q_t^*, \ldots, Q_T^* \rangle$.

Next we show that the policy found by the formulation in Section 3.1 is of an (s_t, Q_t) form. The following discussion is inspired by the work of Gallego and Toktay (2004) on all-or-nothing ordering policies under a capacity constraint. For any opening inventory level x and a vector of order quantities $\boldsymbol{q}_t$, let $J_t(x, \boldsymbol{q}_t)$ and $\hat{J}_t(x, \boldsymbol{q}_t)$ denote the expected total cost when the decision in period t is not to order $(\delta = 0)$ and to order $(\delta = 1)$ respectively, it follows that

$$J_t(x, \boldsymbol{q}_t) \triangleq L_t(x) + \mathbb{E}[V_{t+1}(x - d_t, \boldsymbol{q}_{t+1})] \tag{8}$$

and

$$\hat{J}_t(x, \boldsymbol{q}_t) \triangleq c(Q_t) + L_t(x + Q_t) + \mathbb{E}[V_{t+1}(x + Q_t - d_t, \boldsymbol{q}_{t+1})]. \tag{9}$$

Recall that Eq. (5) optimises the system over the reorder decision $\delta \in \{0,1\}$ and is equivalent to

$$\begin{aligned} V_t(x, \boldsymbol{q}_t) &= \min\{\hat{J}_t(x, \boldsymbol{q}_t),\ J_t(x, \boldsymbol{q}_t)\} \\ &= \min\{K + zQ_t + L_t(x + Q_t) + \mathbb{E}[V_{t+1}(x + Q_t - d_t, \boldsymbol{q}_{t+1})], \\ &\qquad\qquad L_t(x) + \mathbb{E}[V_{t+1}(x - d_t, \boldsymbol{q}_{t+1})]\} \\ &= \min\{K + zQ_t + J_t(x + Q_t, \boldsymbol{q}_t),\ J_t(x, \boldsymbol{q}_t)\} \\ &= J_t(x, \boldsymbol{q}_t) + \min\{K + zQ_t - \Delta J_t(x, \boldsymbol{q}_t),\ 0\}, \end{aligned} \tag{10}$$

where we define

$$\Delta J_t(x, \boldsymbol{q}_t) \triangleq J_t(x, \boldsymbol{q}_t) - J_t(x + Q_t, \boldsymbol{q}_t). \tag{11}$$

From Eq. (10), it is optimal to reorder in period t with opening inventory x when $\Delta J_t(x, \boldsymbol{q}_t) > K + zQ_t$ and not to reorder otherwise. If we choose not to reorder when $\Delta J_t(x, \boldsymbol{q}_t) = K + zQ_t$, then the range of opening inventory level x for which it is optimal to reorder can be expressed as

$$\{x : \Delta J_t(x, \boldsymbol{q}_t) > K + zQ_t\}. \tag{12}$$

If $\Delta J_t(x, \boldsymbol{q}_t)$ is non-increasing in x for given order quantities $\boldsymbol{q}_t^*$, then either there exits an s_t such that it is optimal to order in period t when $x < s_t$ and not otherwise, or it is never optimal to order in period t; and it hence leads to the (s_t, Q_t) policy. In the following, for any given $\boldsymbol{q}_t$, we show the monotonicity of $\Delta J_t(x, \boldsymbol{q}_t)$ in x.

Lemma 1. $L_t(y) - L_t(y + a)$ is non-increasing in y for any $a > 0$ and $t = 1, \ldots, T$.

Proof. Since $L_t(y)$ is convex, its derivative $L_t^{'}(y)$ is non-decreasing by the definition of convexity. For any $a > 0$ and any $t = 1, \ldots, T$, $[L_t(y) - L_t(y + a)]^{'} = L_t^{'}(y) - L_t^{'}(y + a) \leq 0$; therefore, $L_t(y) - L_t(y + a)$ is non-increasing in y. □

Lemma 2. For a given $\boldsymbol{q}_t$, the function $\Delta J_t(x, \boldsymbol{q}_t)$ is non-increasing with respect to the opening inventory level x for any $t = 1, \ldots, T$.

Proof. We prove this by induction. For period T,

$$\Delta J_T(x, \boldsymbol{q}_T) = J_T(x, \boldsymbol{q}_T) - J_T(x + Q_T, \boldsymbol{q}_T) = L_T(x) - L_T(x + Q_T)$$

is non-increasing by Lemma 1. Assuming that $\Delta J_t(x, \boldsymbol{q}_t)$ is non-increasing in x, we want to show that $\Delta J_{t-1}(x, \boldsymbol{q}_{t-1})$ is non-increasing in x. We find that

$$\begin{aligned}
& K + zQ_t + V_t(x + Q_t, \boldsymbol{q}_t) - V_t(x, \boldsymbol{q}_t) \\
= \; & K + zQ_t + J_t(x + Q_t, \boldsymbol{q}_t) - J_t(x, \boldsymbol{q}_t) + \min\{0, K + zQ_t - \Delta J_t(x + Q_t, \boldsymbol{q}_t)\} \\
& \qquad - \min\{0, K + zQ_t - \Delta J_t(x, \boldsymbol{q}_t)\} \\
= \; & K + zQ_t - \Delta J_t(x, \boldsymbol{q}_t) + \min\{0, K + zQ_t - \Delta J_t(x + Q_t, \boldsymbol{q}_t)\} - \min\{0, K + zQ_t - \Delta J_t(x, \boldsymbol{q}_t)\} \\
= \; & \max\{0, K + zQ_t - \Delta J_t(x, \boldsymbol{q}_t)\} + \min\{0, K + zQ_t - \Delta J_t(x + Q_t, \boldsymbol{q}_t)\}
\end{aligned}$$

is the sum of two non-decreasing functions because $\Delta J_t(x, \boldsymbol{q}_t)$ is assumed to be non-increasing. It follows that $V_t(x, \boldsymbol{q}_t) - V_t(x + Q_t, \boldsymbol{q}_t)$ is non-increasing. Consequently, since $L_{t-1}(x) - L_{t-1}(x + Q_{t-1})$ is non-increasing in x,

$$\begin{aligned}
\Delta J_{t-1}(x, \boldsymbol{q}_t) &= J_{t-1}(x, \boldsymbol{q}_{t-1}) - J_{t-1}(x + Q_{t-1}, \boldsymbol{q}_{t-1}) \\
&= L_{t-1}(x) - L_{t-1}(x + Q_{t-1}) + \mathbb{E}[V_t(x - d_{t-1}, \boldsymbol{q}_t) - V_t(x + Q_t - d_{t-1}, \boldsymbol{q}_t)]
\end{aligned}$$

is the sum of two non-increasing functions; therefore, $\Delta J_{t-1}(x, \boldsymbol{q}_t)$ is non-increasing in x. This completes the proof by induction. □

For a given $\boldsymbol{q}_t$, the monotonicity of $\Delta J_t(x, \boldsymbol{q}_t)$ in x assures the unique existence of the reorder point s_t, which defines the region of opening inventory $x < s_t$ for which it is optimal to reorder, where s_t can be denoted as

$$s_t = \inf\{x : \Delta J_t(x, \boldsymbol{q}_t) < K + zQ_t\}; \tag{13}$$

if the inventory levels are discrete, then s_t is the minimum value of x such that $\Delta J_t(x, \boldsymbol{q}_t) < K + zQ_t$, where Q_t is the first argument of the order quantities $\boldsymbol{q}_t$. The reorder points associated with the optimal order quantities $\boldsymbol{q}_t^*$ hence can be denoted as $\boldsymbol{s}_t^* \triangleq \langle s_t^*, \ldots, s_T^* \rangle$.

Example 2. Consider a 4–period stochastic lot-sizing problem under Poisson-distributed demand with rates $d_t = \langle 2, 1, 5, 3 \rangle$. The cost parameters are $K = 5$, $z = 0$, $h = 1$ and $b = 3$. The maximum order quantity is set to 9. After exhaustive enumeration of all order quantity vectors, we obtain $\boldsymbol{q}_1^* = \langle 3, 3, 8, 5 \rangle$ and the associated reorder points $\boldsymbol{s}_1^* = \langle 1, 0, 4, 1 \rangle$. The expected total cost of the optimal (s_t, Q_t) policy is 22.5 when the initial inventory is 0. Under discrete inventory levels with Poisson demand, Fig. 2 and Fig. 3 illustrate determining s_1^* by scatter plots. In Fig. 2, $s_1^* = 1$ is selected as the minimum value such that $\Delta J_1(I_0, \boldsymbol{q}_1^*) < K$, which is equivalent to $J_1(I_0, \boldsymbol{q}^*) > \hat{J}_1(I_0, \boldsymbol{q}^*)$ when $I_0 \leq 0$, suggesting it is optimal to order; and $J_1(I_0, \boldsymbol{q}^*) < \hat{J}_1(I_0, \boldsymbol{q}^*)$ when $I_0 \geq 1$, suggesting it is optimal not to order, as Fig. 3 shows.

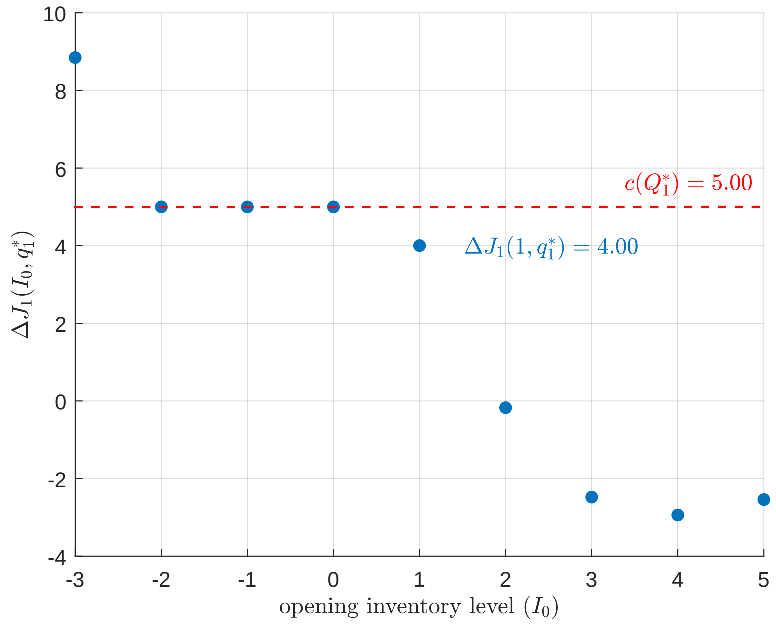

Figure 2: $s_1^* = 1$ determined by comparing $\Delta J_1(I_0, \boldsymbol{q}_1^*)$ and $c(Q_1^*)$.

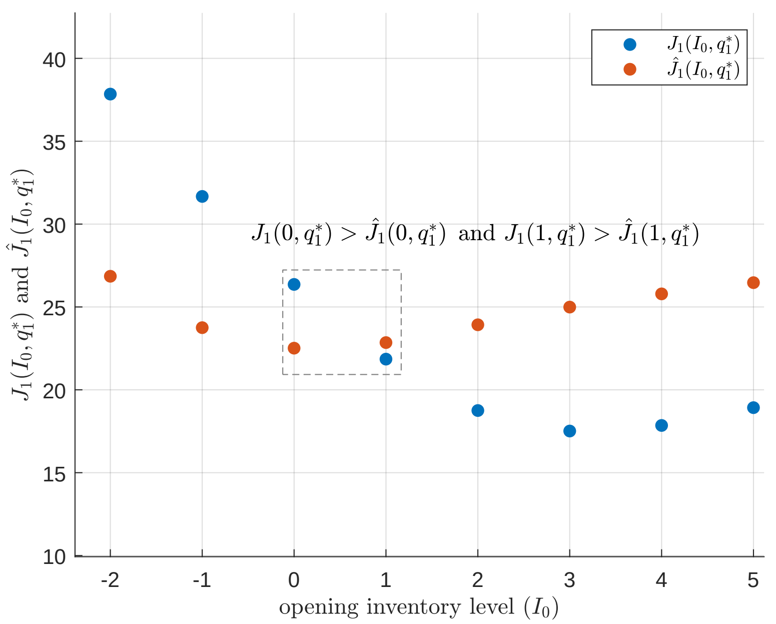

Figure 3: $s_1^* = 1$ determined by comparing $J_1(I_0, \boldsymbol{q}_1^*)$ and $\hat{J}_1(I_0, \boldsymbol{q}_1^*)$.

3.2 A stochastic dynamic program for the (s_t, Q) policy

An (s_t, Q) policy places a replenishment order of size Q if the inventory level falls below the reorder point s_t and does not place an order otherwise. It is therefore a special case of (s_t, Q_t) in which all Q_t's are equal. We modify the vector space $\mathcal{Q}_t$ introduced in section 3.1 to explore the (s_t, Q) policy.

Let $\dot{\boldsymbol{q}}_t \triangleq \langle Q, \ldots, Q \rangle$ be a $(T-t+1)$-dimensional vector of reorder quantities for the (s_t, Q) policy and $\dot{\mathcal{Q}}_t$ be the vector space containing all combinations of order quantities $\dot{\boldsymbol{q}}_t$. It follows that $\dot{\mathcal{Q}}_t$ is a subspace of $\mathcal{Q}_t$. For a given $\dot{\boldsymbol{q}}_t \in \dot{\mathcal{Q}}_t$, the expected total cost over periods t to T when the opening inventory level is x is

$$V_t(x, \dot{\boldsymbol{q}}_t) = \min_{\delta \in \{0,1\}} \{c(\delta Q) + L_t(x + \delta Q) + \mathbb{E}[V_{t+1}(x + \delta Q - d_t, \dot{\boldsymbol{q}}_{t+1})]\}, \quad t < T, \tag{14}$$

and

$$V_T(x, \dot{\boldsymbol{q}}_T) = \min_{\delta \in \{0,1\}} \{c(\delta Q) + L_T(x + \delta Q)\} \tag{15}$$

is the boundary condition. The optimal expected total cost under the (s_t, Q) policy with opening inventory level x can be defined as

$$V_0(x) = \min_{\dot{\boldsymbol{q}}_1 \in \dot{\mathcal{Q}}_1} \{V_1(x, \dot{\boldsymbol{q}}_1)\}. \tag{16}$$

We let the optimal order quantity vector be $\dot{\boldsymbol{q}}_t^* \triangleq \langle Q^*, \ldots, Q^* \rangle$. Since $\dot{\mathcal{Q}}_t$ is a subspace of $\mathcal{Q}_t$, Lemma 2 holds for any $\dot{\boldsymbol{q}}_t \in \dot{\mathcal{Q}}_t$. The determination of reorder points under the (s_t, Q) policy follows in the same fashion as for the (s_t, Q_t) policy by Eq. (13). We denote the reorder points associated with $\dot{\boldsymbol{q}}_t^*$ as $\dot{\boldsymbol{s}}_t^* \triangleq \langle \dot{s}_t^*, \ldots, \dot{s}_T^* \rangle$.

Example 1 (**Continued**). Recall the 4–period stochastic lot-sizing problem under Poisson-distributed demand with rates $d_t = \langle 20, 40, 60, 40 \rangle$. Under the (s_t, Q) policy, the optimal order quantity is $Q^* = 83$ as illustrated by Fig. 4. The reorder points associated with $\dot{\boldsymbol{q}}_1^*$ are determined as $\dot{\boldsymbol{s}}_1^* = \langle 13, 33, 54, 24 \rangle$. Fig. 5 and Fig. 6 illustrate determining $\dot{s}_1^* = 13$. Note that we apply curves to show the trend of expected costs, while the system is in fact discrete. In Fig. 6, a unique sign change in $[\Delta J_1(I_0, \dot{\boldsymbol{q}}_1^*) - c(Q^*)]$ is detected between $I_0 = 12$ and 13 and so, by Eq. (13), $I_0 = 13$ is chosen as $\dot{s}_1^*$.

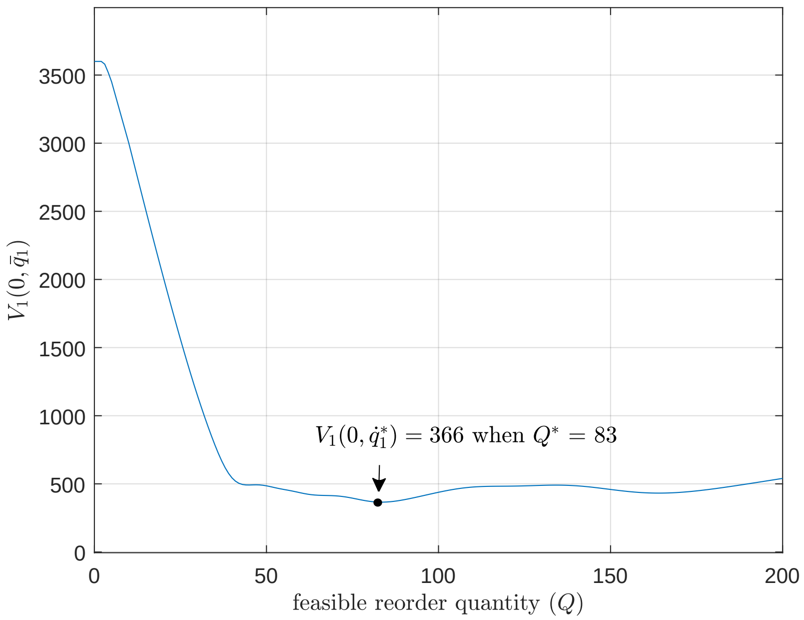

Figure 4: $Q^* = 83$ under (s_t, Q) policy for Example 1.

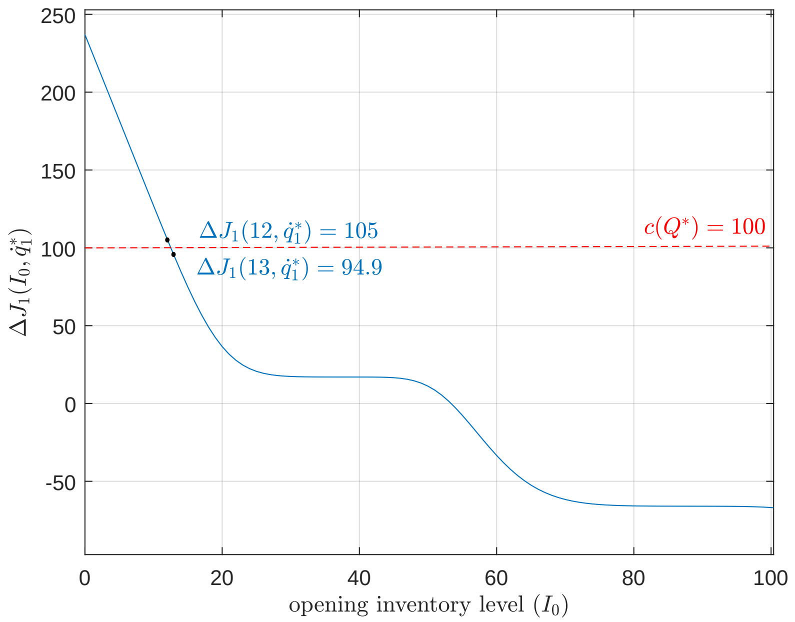

Figure 5: $\dot{s}_1^* = 13$ determined by comparing $\Delta J_1(I_0, \dot{\boldsymbol{q}}_1^*)$ and $c(Q^*)$.

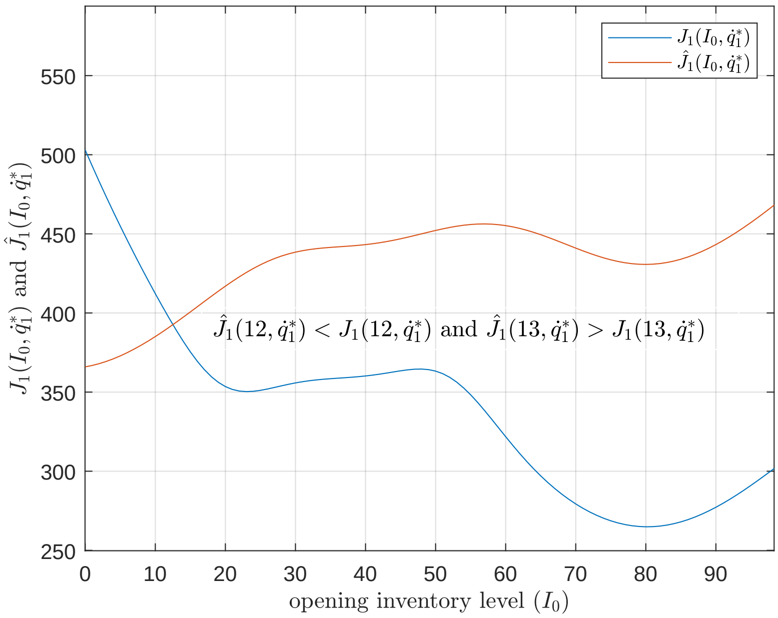

Figure 6: $\dot{s}_1^* = 13$ determined by comparing $J_1(I_0, \dot{\boldsymbol{q}}_1^*)$ and $\hat{J}_1(I_0, \dot{\boldsymbol{q}}_1^*)$.

4 A MINLP-based heuristic for the (s_t, Q_t) policy

Optimal (s_t, Q_t) and (s_t, Q) policies can be obtained by enumerating all possible order quantities and using the stochastic dynamic programming formulations presented in Section 3. However, as the length of planning horizon increases, the enumeration increases exponentially and it becomes impractical to use this method. In this section, we therefore introduce an effective heuristic to compute near-optimal (s_t, Q_t) and (s_t, Q) policy parameters in reasonable time. Our heuristic leverages a MINLP approximation of $V_t(\cdot)$ and, similarly to Bookbinder and Tan (1988), it comprises two steps: in the first step, we determine a set of near-optimal order quantities; in the second step, we compute the associated reorder points.

4.1 Step I: Order quantity Q_t of the (s_t, Q_t) policy

We first aim to derive a vector of near-optimal order quantities $\hat{\boldsymbol{q}}_t \triangleq \langle \hat{Q}_1 \ldots, \hat{Q}_T \rangle$ for our heuristic (s_t, Q_t) policy. The reader should note that we seek a policy that is near-optimal in terms of expected total cost, not in terms of how close the policy parameters obtained are to the true optimal ones. Therefore, our approximated order quantities and reorder points do not need to be close to the true optimal ones for the (s_t, Q_t) policy, as long as the expected total cost they provide is close enough to the expected total cost of an optimal policy.

Note that if an order is placed in period t under the (s_t, S_t) policy, the order quantity is at least $S_t - s_t$; in fact, if the opening inventory level $I_{t-1} < s_t$ in period t, a further $s_t - I_{t-1}$ items will be ordered to ensure the order-up-to level is reached. In our heuristic (s_t, Q_t) policy, we define $\hat{Q}_t \triangleq S_t - s_t$ to be our approximate order quantity in period t; and we will denote the vector of approximate order quantities as $\hat{\boldsymbol{q}}_t \triangleq \langle \hat{Q}_t, \ldots, \hat{Q}_T \rangle$. While these $\hat{Q}_t$'s may not be optimal, we will compensate for this in Section 4.2, by computing suitable reorder points that are tailored for these approximate order quantities.

Of course, to compute $\hat{Q}_t$, we need optimal or near-optimal values of parameters s_t and S_t of the (s_t, S_t) policy. We use the approach introduced by Xiang et al. (2018) to compute near-optimal s_t and S_t values. For completeness, the model used is presented in Appendix B.

4.2 Step II: Reorder point s_t of the (s_t, Q_t) policy

Since approximate order quantities $\hat{\boldsymbol{q}}_t$ are a lower bound for order quantities observed under an (s_t, S_t) policy, we cannot directly use the reorder points from the optimal (s_t, S_t) policy as the reorder points for a heuristic (s_t, Q_t) policy. To compensate for the under-estimation in the order quantities, we need higher reorder points.

For a given vector $\hat{\boldsymbol{q}}_t$ of approximate order quantities, we may compute the associated optimal reorder points by using an SDP formulation. This would be relatively straightforward for Poisson demand, but would require a discretisation step for continuous demand distributions. In order to provide a framework that can be applied to Poisson, normal, and possibly other continuous demand distributions, we modify the model in Xiang et al. (2018) to capture the characteristics of an (s_t, Q_t) and provide an approximation $\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t)$ of $J_t(x, \hat{\boldsymbol{q}}_t)$ that can be used in Eq. (13) to compute near-optimal reorder points. Let $\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t)$ be our approximation of $J_t(x, \hat{\boldsymbol{q}}_t)$ for the set of near-optimal order quantities $\hat{\boldsymbol{q}}_t$ computed in Section 4.1.

$$\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t) = \min \quad h\tilde{H}_t + b\tilde{B}_t + \sum_{k=t+1}^{T} [h\tilde{H}_k + b\tilde{B}_k + c(\delta_k \hat{Q}_k)], \tag{17}$$

$$\text{s.t.} \quad \delta_t = 0, \tag{18}$$

$$\tilde{I}_t + \tilde{d}_t = \tilde{I}_{t-1}, \tag{19}$$

$$\delta_k = 0 \to \tilde{I}_k + \tilde{d}_k - \tilde{I}_{k-1} = 0, \qquad k = t+1, \ldots, T, \tag{20}$$

$$\delta_k = 1 \to \tilde{I}_k + \tilde{d}_k - \tilde{I}_{k-1} = \hat{Q}_k, \qquad k = t+1, \ldots, T, \tag{21}$$

$$P_{jk} \geq \delta_j - \sum_{r=j+1}^{k} \delta_r, \qquad k = t, \cdots, T \text{ and } j = t, \ldots, k, \tag{22}$$

$$\sum\nolimits_{j=t}^{k} P_{jk} = 1, \qquad k = t+1, \ldots, T, \tag{23}$$

$$P_{jk} = 1 \to \tilde{H}_k = \hat{\mathcal{L}}(\tilde{I}_k + \tilde{d}_{jk}, d_{jk}), \qquad k = t, \ldots, T \text{ and } j = t, \ldots, k, \tag{24}$$

$$P_{jk} = 1 \to \tilde{B}_k = \mathcal{L}(\tilde{I}_k + \tilde{d}_{jk}, d_{jk}), \qquad k = t, \ldots, T \text{ and } j = t, \ldots, k, \tag{25}$$

$$\tilde{H}_k, \tilde{B}_k \geq 0, P_{jk}, \delta_k \in \{0, 1\}, \qquad k = t, \ldots, T \text{ and } j = t, \ldots, k. \tag{26}$$

Let $\tilde{H}_k$ and $\tilde{B}_k$ denote the expected positive inventory and back-order levels at the end of period k, respectively; their values are computed by following the piecewise-linear approximation strategy in Rossi et al. (2015), which is based on the first-order loss function $\mathcal{L}$ and its complement $\hat{\mathcal{L}}$. We discuss in detail the loss function and its piecewise-linear approximation under non-stationary Poisson demand in Appendix C.

In line with (Tarim and Kingsman, 2006), we introduce the binary decision variable δ_k that takes value 1 if and only if an order is placed in period k. In the model above, the objective function $\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t)$ approximates the expected total cost over horizon (t, T) with no order in period t. Constraint (18) indicates that no order is placed in period t and leads to the flow balance equation for period t as constraint (19). Constraints (20) and (21) are indicator constraints[1] representing the flow balance equations and reorder conditions under the (s_t, Q_t) policy that applies order quantities $\hat{\boldsymbol{q}}_{t+1}$ over horizon $(t+1, T)$.

We introduce a binary variable P_{jk} to properly account for demand variance while computing the first-order loss function. Let P_{jk} ($j \leq k$) take value of 1 if the last order before period k (including period k itself) is placed at the beginning of period j. Note that the combination of constraints (22) and (23) ensures that demand variance is properly accounted even when no

[1] An indicator constraint (denoted by $\to$), see e.g. (Belotti et al., 2016), is a way to express relationships among variables by identifying a binary variable to control whether or not a specified constraint is active. Indicator constraints are standard constructs that are nowadays implemented in most off-the-shelf MILP solvers.

order takes place within the horizon (j, k). Constraints (24) and (25) are indicator constraints modelling end of period k expected excess inventory and back-orders by means of the first order loss function (Xiang et al., 2018).

Since $J_t(x, \boldsymbol{q}_t)$ is approximated as $\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t)$, the near-optimal reorder point $\hat{s}_t$ can be determined, following Eq. (13), as

$$\hat{s}_t = \inf\{x : \Delta\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t) < K + z\hat{Q}_t\}, \tag{27}$$

or as the minimum value of x such that

$$\Delta\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t) < K + z\hat{Q}_t \tag{28}$$

for discrete inventory levels, where $\Delta\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t) \triangleq \mathcal{J}_t(x, \hat{\boldsymbol{q}}_t) - \mathcal{J}_t(x + \hat{Q}_t, \hat{\boldsymbol{q}}_t)$. Note that $\Delta\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t)$ is not necessarily monotonic in x, since the piecewise linearisation produces errors; our model applies the optimal partitioning strategy to maintain a minimum error (Rossi et al., 2014, Thm. 11). We denote the vector of near-optimal reorder points associated with $\hat{\boldsymbol{q}}_t$ as $\hat{\boldsymbol{s}}_t \triangleq \langle \hat{s}_t, \ldots, \hat{s}_T \rangle$.

4.3 A binary search approach to approximate the reorder points s_t

A line search for $\hat{s}_t$ following Eq. (27) may be too time-consuming for large-scale instances. This subsection introduces a heuristic algorithm to approximate $\hat{s}_t$ and reduce computational complexity.

The algorithm applies a binary search on $\Delta\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t)$ with $\hat{\boldsymbol{q}}_t$ known as an input. For any period t, the opening inventory level x_0 and given step-size w ($w > 0$) define an interval of inventory level $[x_0,\, x_0 + w]$, which maps to $[\Delta\mathcal{J}_t(x_0 + w, \hat{\boldsymbol{q}}_t),\, \Delta\mathcal{J}_t(x_0, \hat{\boldsymbol{q}}_t)]$. The binary search halves the length of the interval in each iteration until $\hat{s}_t$ is detected according to Eq. (27). If the initial interval does not span the point at which the sign of $\Delta\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t) - K - z\hat{Q}_t$ changes, we renew $[x_0,\, x_0 + w]$ by panning it w units to the left if $\Delta\mathcal{J}_t(x_0 + w, \hat{\boldsymbol{q}}_t) < K + z\hat{Q}_t$ or to the right, otherwise; and then proceed with the binary search.

We present the following algorithm for integer inventory levels. One can extend it to discrete systems with any interval between two adjacent inventory levels. For integer inventory levels, the algorithm terminates if a pair of inventory levels x and $x + 1$ are found such that $\Delta\mathcal{J}_t(x, \hat{\boldsymbol{q}}_t) \leq K + z\hat{Q}_t \leq \Delta\mathcal{J}_t(x + 1, \hat{\boldsymbol{q}}_t)$, and then $\hat{s}_t = x + 1$. The procedure in detail is as follows.

Algorithm 1 Computing the reorder points $\hat{s}_t$ associated with $\hat{\boldsymbol{q}}_t$.

1: **Input:** demand rates $\tilde{d}_t$; cost parameters (K, z, h, b); the step-size w; an opening inventory x_0; order quantities $\hat{\boldsymbol{q}}_t$.
2: **Output:** reorder point $\hat{s}_t$ associated with $\hat{\boldsymbol{q}}_t$.
3: **for** $t = 1 \to T$ **do**
4:     Compute the ordering cost of placing an order $\mathcal{J}_0 = K + z\hat{Q}_t$;
5:     $x_l = x_0$ and $x_r = x_0 + w$;
6:     compute $\mathcal{J}_l = \Delta\mathcal{J}_t(x_l, \hat{\boldsymbol{q}}_t)$ and $\mathcal{J}_r = \Delta\mathcal{J}_t(x_r, \hat{\boldsymbol{q}}_t)$ with $\hat{Q}_t$;
7:     **if** $\mathcal{J}_l > \mathcal{J}_0 > \mathcal{J}_r$ **then**
8:         $x_m = \lfloor \frac{x_l + x_r}{2} \rfloor$ and $\mathcal{J}_m = \Delta\mathcal{J}_t(x_m, \hat{\boldsymbol{q}}_t)$;
9:         **if** $\mathcal{J}_m > \mathcal{J}_0$ **then**
10:             **if** $\Delta\mathcal{J}_t(x_m + 1, \hat{\boldsymbol{q}}_t) < \mathcal{J}_0$ **then**
11:                 output $\hat{s}_t = x_m$;
12:             **else** $x_l = x_m$, $x_r = x_r$, and repeat lines 6 – 20;
13:             **end if**
14:         **else**
15:             **if** $\Delta\mathcal{J}_t(x_m - 1, \hat{\boldsymbol{q}}_t) > \mathcal{J}_0$ **then**
16:                 output $\hat{s}_t = x_m - 1$;
17:             **else** $x_l = x_l$, $x_r = x_m$, and repeat lines 6 – 20;
18:             **end if**
19:         **end if**
20:     **end if**
21: **end for**

Example 2 (Continued). Recall the 4–period stochastic lot-sizing problem under Poisson-distributed demand with rates $d_t = \langle 2, 1, 5, 3 \rangle$. Applying 20 partitions in the piecewise linearisation approximation, $\hat{\boldsymbol{q}}_1 = \langle 3, 4, 9, 5 \rangle$ approximates $J_1(I_0, \boldsymbol{q}_1^*)$ as shown in Fig. 7 for $I_0 \in [-4, 14]$. The curves are plotted to demonstrate the difference between $J_1(I_0, \boldsymbol{q}_1^*)$ and $\mathcal{J}_t(I_0)$, while the system is in fact discrete. Table 1 compares the (s_t, Q_t) policy parameters obtained by the SDP in Section 3.1 and the heuristic for a zero initial inventory level.

Table 1: Parameters of the (s_t, Q_t) policy for Example 2 computed by SDP and the heuristic.

	$\hat{Q}_t$				$\hat{s}_t$			
t	1	2	3	4	1	2	3	4
SDP	3	3	8	5	1	0	4	1
Heuristic	3	4	9	5	1	-2	4	0

Taking $G_1(0) = 21.8$ as a benchmark, the optimality gaps of the (s_t, Q_t) policy determined by SDP and our heuristic, relative to the (s_t, S_t) policy are shown in Table 2. We note that the (s_t, Q_t) policy produces large optimality gaps in Example 2, where values of the expected total

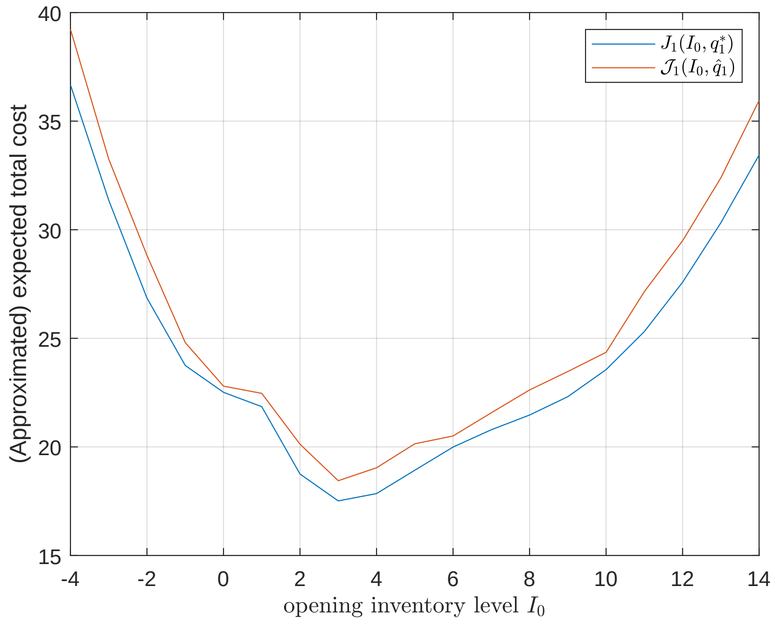

Figure 7: Plot of $J_1(I_0, \boldsymbol{q}_1^*)$ and $\mathcal{J}_1(I_0, \hat{\boldsymbol{q}}_1)$.

cost are small, while the approximation accuracy of the heuristic $(23.1 - 22.5)/22.5 \times 100\% = 2.67\%$ is acceptable. We will extend our computational study in Section 5 to investigate how the (s_t, Q_t) policy performs on several problem instances.

Table 2: Expected total cost (ETC) and optimality gap (OG) of SDP and the heuristic for Example 2 under the (s_t, Q_t) policy.

	ETC	OG(%)
SDP	22.5	3.33
Heuristic	23.1	5.93

4.4 Approximation of the (s_t,Q) policy parameters

For the (s_t, Q) policy, a direct way to approximate the order quantity is to simplify the model in Appendix B by replacing Q_t with Q and then follow the steps in Sections 4.1 and 4.2; however, this is found to produce large optimality gaps in terms of the expected total cost.

Following the line of reasoning illustrated in Section 4.1 for (s_t, Q_t), one can derive a single order quantity in period 1 as $S_1 - I_0$ for a known opening inventory I_0. However, a high value for I_0 may result in a low order quantity imposed over a long period. In our heuristic (s_t, Q) policy, we define $\hat{Q} \triangleq S_1$ to be our approximate order quantity for horizon $(1, T)$; and we denote the vector of approximate order quantities as $\hat{\boldsymbol{q}} \triangleq \langle \hat{Q}, \ldots, \hat{Q} \rangle$. The reorder points are adjusted to compensate for the over-estimation for cases with high opening inventory levels.

The computation of reorder points $\hat{\boldsymbol{s}}_t$ associated with order quantity $\hat{Q}$ follows the same procedure proposed in Section 4.2 for (s_t, Q_t). We apply Model 4.2 with $\hat{Q}$ to obtain the approximated expected cost over horizon (t, T) when no order is placed in period t, denoted as $\mathcal{J}_t(x, \hat{\boldsymbol{q}})$, and we apply our previously introduced heuristic algorithm on the function $\Delta \mathcal{J}_t(x, \hat{\boldsymbol{q}})$ to determine $\hat{s}_t$.

Example 1 (Continued). Applying 20 partitions in the piecewise linearisation approximation, Fig. 8 approximates $J_1(I_0, \dot{\boldsymbol{q}}_1^*)$ by $\mathcal{J}_1(I_0, \hat{\boldsymbol{q}})$. While the inventory system is discrete, we plot interpolated curves for the sake of clarity. For a zero initial inventory level, Table 3 compares the

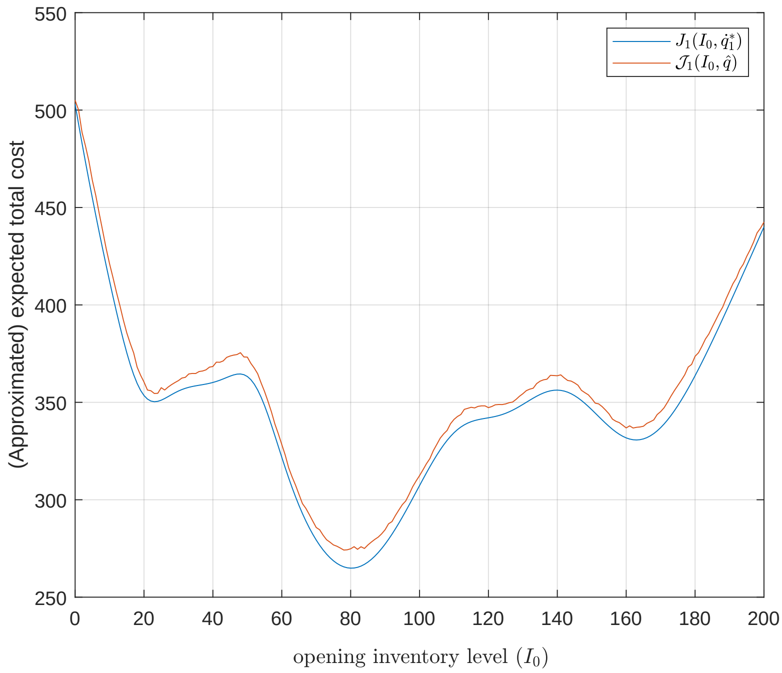

Figure 8: Plot of $J_1(I_0, \dot{\boldsymbol{q}}_1^*)$ and $\mathcal{J}_1(I_0, \hat{\boldsymbol{q}})$.

parameters of the (s_t, Q) policy computed by SDP and the approximation. Taking $G_1(0) = 481$ as the benchmark, Table 4 summarises the optimality gaps of the (s_t, Q) policy computed by SDP and the heuristic. The approximation accuracy $(505 - 503)/503 \times 100\% = 0.398\%$ is high.

Table 3: Parameters of the (s_t, Q) policy for Example 1 computed by SDP and the heuristic.

	$\hat{Q}$	$\hat{s}_t$			
t	–	1	2	3	4
SDP	83	13	33	54	24
Heuristic	84	14	34	55	24

We discuss the performance of the (s_t, Q) policy in detail in the next section.

Table 4: Expected total cost (ETC) and optimality gap (OG) of SDP and the heuristic for Example 1 under the (s_t, Q) policy

	ETC	OG(%)
SDP	503	4.57
Heuristic	505	4.99

5 Computational analysis

This section presents a computational analysis to evaluate (s, Q)-type policies under non-stationary stochastic demand. The analysis considers both the stochastic dynamic programming formulations and our heuristics for the (s_t, Q_t) and (s_t, Q) policies. In Section 5.1, we consider a test set comprising small problem instances with 6 periods; we investigate the performance of optimal (s, Q)-type policies against that of optimal non-stationary (s, S) policies, and we evaluate the difference between optimal (s, Q) and heuristic (s, Q) policies. In Section 5.2, we consider a test set comprising large problem instances with 25 periods; we investigate the performance of (s, Q)-type heuristics versus that of the optimal non-stationary (s, S) policy; we also compare the performance between our (s, Q)-type heuristics and another existing static-dynamic uncertainty heuristic, namely the (R_t, S_t) policy discussed in (Rossi et al., 2015).

We name the optimal policy for the stochastic lot-sizing problem, which takes an (s, S) form, (s_t, S_t)-SDP. In our experiment we consider two variants of the (s, Q) policy: the (s_t, Q_t) policy, and the (s_t, Q) policy; presented in Section 3.1 and Section 3.2, respectively. For each variant, we discuss results for the optimal SDP formulation, named (s_t, Q_t)-SDP and (s_t, Q)-SDP, respectively; and results for our MINLP heuristics formulations presented in Section 4, named (s_t, Q_t)-Heuristic and (s_t, Q)-Heuristic, respectively. We apply 10 partitions in the piecewise approximation for both heuristics. We simulate each test instance with the policy parameters obtained from the heuristics and derive the average total cost of 500,000 simulation runs.

For each approach, we always use the optimal (s, S) policy as a benchmark. Approaches are compared in terms of their expected total cost (ETC) using the percent optimality gap computed as $100 \times (\mathrm{ETC}_2 - \mathrm{ETC}_1)/\mathrm{ETC}_1$, where ETC_1 is the expected total cost of the optimal non-stationary (s, S) policy, and ETC_2 is the expected total cost of the other approach benchmarked. We set a zero initial inventory for all test instances and test the robustness of heuristics for (s, Q)-type policies.

In our numerical study, we consider ten expected demand patterns: two life cycle patterns, one moves from the launch stage to maturity via a growth (LCY1) and the other moves from the growth stage through maturity and into decline (LCY2); two sinusoidal patterns, one with stronger (SIN1) and the other with weaker (SIN2) oscillations; a stationary demand pattern (STAT); a random demand pattern (RAND); and lastly, 4 empirical patterns derived according to (Strijbosch et al., 2011).

All computations are performed by a 4.0 (1.90+2.11) gigahertz Intel(R) Core(TM) i7-8650U CPU with 16.0 gigabytes of RAM in JAVA 1.8.0_201.

5.1 A test set with 6-period Poisson-distributed demand

The first test set involves 60 instances over a 6-period planning horizon in which the demand follows a non-stationary Poisson distribution. Our aim is twofold: first, we aim to investigate the performances of optimal (s, Q)-type policies obtained via SDP against the optimal non-stationary (s, S) policy; second we aim to evaluate the difference between the optimal and heuristic (s, Q) policies.

We assume the maximum order quantity is 9, which allows us to enumerate all combinations of order quantities for the (s_t, Q_t) policy by stochastic dynamic programming. The problems in this test set are designed with very small mean demands λ_t, as illustrated in Fig. 9. The values of λ_t are set to be between 1 and 7 in all cases which allows variation in the optimal values of Q_t and ensures that the optimal order quantity is never as high as 9 in any period. The problem coefficients are considered over $z \in \{0, 1\}$ and the three sets of K and b shown in Table 5 with different ratios of K to b. Holding cost is set as $h = 1$ for all instances.

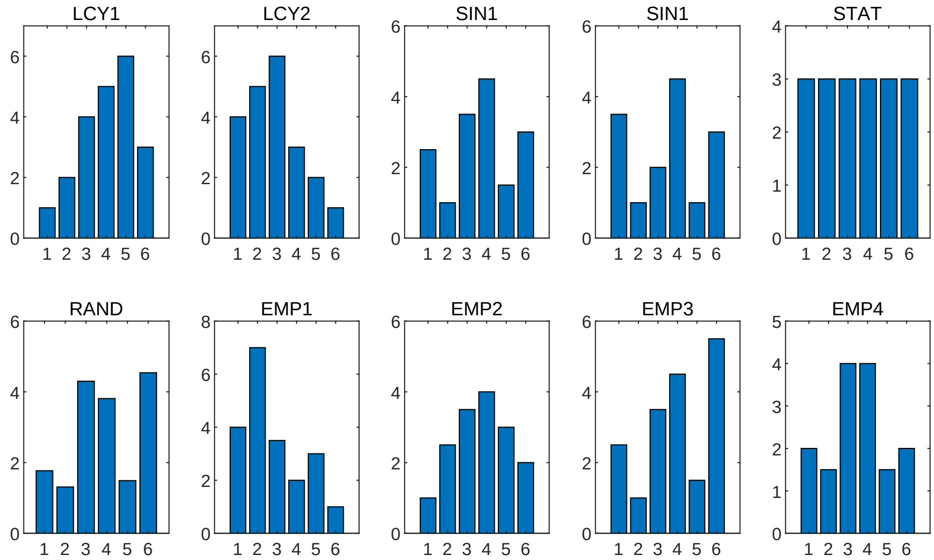

Figure 9: Demand patterns of 6-period instances.

Table 5: parameter groups of fixed ordering cost (K) and penalty cost (p)

set	K	b	ratio
1	5	3	1.67
2	10	3	2.00
3	10	7	1.43

For each approach considered, Table 6 reports the optimality gaps observed relative to the optimal (s_t, S_t) policy. The results for (s_t, Q_t)-SDP and (s_t, Q)-SDP give the exact optimality gaps for these policies against optimal (s_t, S_t) policy, which are on average 1.91% and 3.61% respectively. In detail, (s_t, Q_t)-SDP performs better than (s_t, Q)-SDP in every individual demand pattern; and (s_t, Q)-SDP is dominated by (s_t, Q_t)-SDP even in the case of a stationary demand pattern. In view of cost parameters, there is no obvious relation between optimality gaps and the variation in demand patterns or in the ratio of K to b. Optimality gaps also decrease when the unit cost increases. On the other hand, the increase in penalty cost results in a small increase in the optimality gap for both (s_t, Q_t)-SDP (1.43% to 1.49%) and (s_t, Q)-SDP (2.92% to 3.22%).

For (s_t, Q_t)-Heuristic and (s_t, Q)-Heuristic we found that the optimality gaps increase by an average of 0.85% and 1.05%, respectively. The largest average increases arise under demand pattern EMP3 (1.04%) for (s_t, Q_t)-Heuristic and RAND (1.75%) for (s_t, Q)-Heuristic. We conclude that the difference between SDP and the heuristic approach is generally low.

Table 6: Average percent optimality gap over our 6-period test set under different demand patterns and pivoting parameters.

Problem Settings	(s_t, Q_t)-SDP	(s_t, Q_t)-Heuristic	(s_t, Q)-SDP	(s_t, Q)-Heuristic
demand pattern				
LCY1	1.96	2.60	2.55	3.30
LCY2	2.70	3.60	5.37	6.11
SIN1	1.95	2.89	3.96	4.80
SIN2	2.13	3.04	3.18	4.75
STAT	1.54	2.41	2.45	4.00
RAND	1.17	2.02	3.12	4.86
EMP1	1.98	2.87	3.98	5.33
EMP2	2.32	2.94	3.56	4.44
EMP3	1.13	2.17	3.11	3.66
EMP4	2.21	3.11	4.80	5.39
unit cost				
0	2.03	2.93	3.83	5.15
1	1.79	2.59	3.38	4.18
set				
1	2.81	3.76	4.67	5.76
2	1.43	2.29	2.92	3.96
3	1.49	2.23	3.22	4.28
Average	1.91	2.76	3.61	4.66

5.2 A test set with 25-period Normally-distributed demand

We extend the planning horizon to 25 periods. The purpose of implementing this test set is twofold. First we aim to investigate the performance of (s, Q)-type heuristics versus that of the optimal non-stationary (s, S) policy for larger instances; second, we aim to compare the performances of (s, Q)-type heuristics and the non-stationary (R, S) policy introduced in (Rossi et al., 2015), which we name (R_t, S_t)-Heuristic.

Since the computation of piecewise linearisation parameters consumes a large amount of computation time for large non-stationary demand following a Poisson distribution, in what follows we will focus on normally distributed demand patterns, for which Rossi et al. (2014) present precomputed optimal partitioning coefficients.

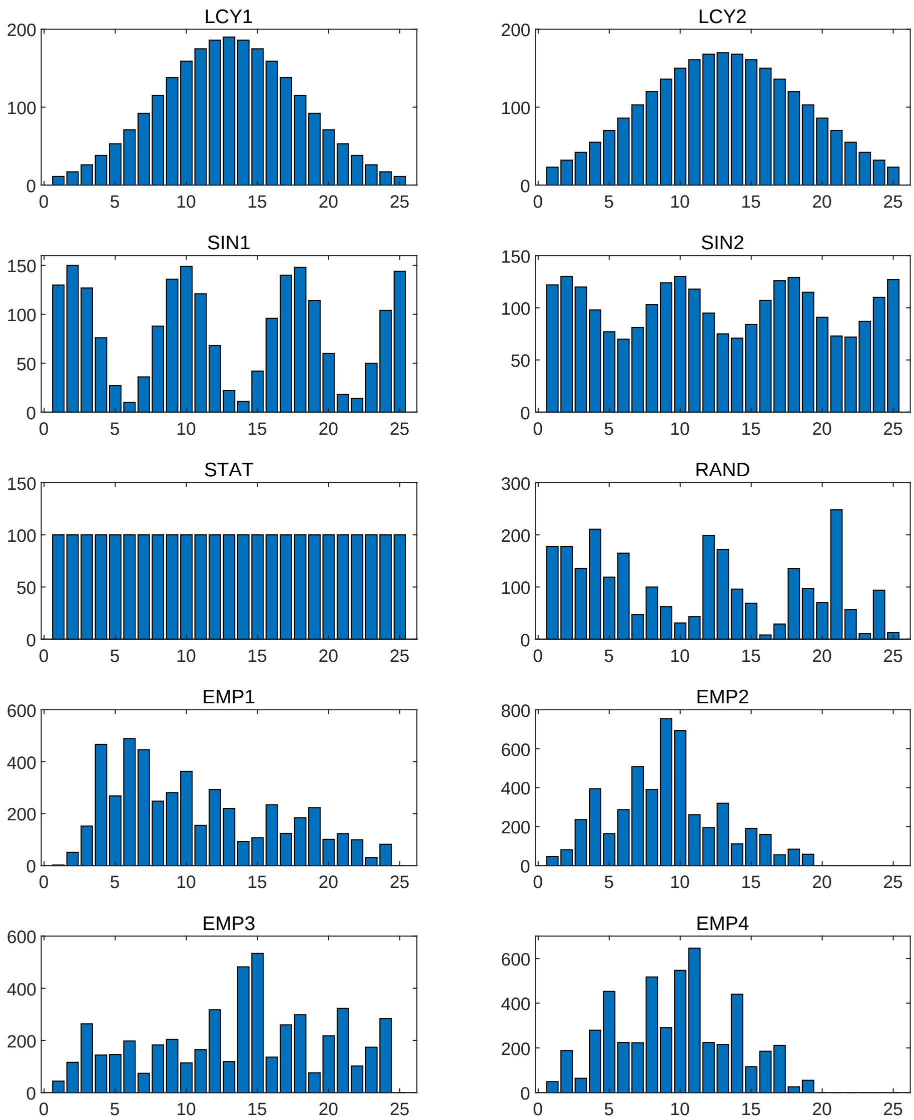

Figure 10: Demand patterns of 25-period instances.

We refer to the 25-period instances in Xiang et al. (2018). The demand d_t in each period t is assumed to be a normally distributed random variable with known mean $\tilde{d}_t$ and standard deviation $\sigma_t = \rho \cdot \tilde{d}_t$, where ρ denotes the coefficient of variation of the demand, which remains fixed over time as prescribed in Bollapragada and Morton (1999); demands are assumed to be independent of each other. We allow the standard deviation parameter ρ to vary over $\rho \in$

$\{0.1, 0.2, 0.3\}$. Demand patterns are illustrated in Figure 10. Other problem parameters are $K \in \{500, 1000, 1500\}$; $b \in \{5, 10, 20\}$; $z \in \{0, 1\}$; and $h = 1$.

The reader should note that, since stochastic dynamic programming is pseudo-polynomial, an increase in the average value of the demand or of its standard deviation will lead to a dramatic increase in the state space and hence of computational times (Dural-Selcuk et al., 2020). The (s_t, Q)-SDP can be implemented by bounding the inventory level, while it is no longer possible to compute (s_t, Q_t)-SDP within a reasonable time for normal demand or large planning horizons such as 25.

Table 7 reports average optimality gaps for our 25-period instances. For the (s_t, Q_t)-Heuristic, the average optimality gap is 2.31%, which is similar to the result obtained for the 6-period test problems. The optimality gap also exhibits similar trends with the penalty cost and the unit cost, while the gap increases with penalty cost and decreases when the unit cost is increased. For the normal distribution, the increase of the standard deviation parameter ρ reduces the optimality gap, which suggests the (s_t, Q_t) policy performs slightly better when the standard deviation of demand is higher.

For (s_t, Q)-Heuristic, once more, as with 6-period test set, we observe that the (s_t, Q)-Heuristic is not satisfactory. The average optimality gap now increases up to 11.5%; and for an individual demand pattern, the optimality gap reaches 25.9%. We also cross-validated results against optimal (s_t, Q) parameters obtained via SDP, to ensure the accuracy of the result, but found that the optimality gap remained as large as 10.5% on average. This confirms that it is not just the approximation, but the policy itself that performs poorly. We believe that under non-stationary demand, when the length of planning horizon increases, the single order quantity Q in the (s_t, Q) policy cannot properly hedge against demand, and thus it produces substantially higher expected cost than other policies that provide more flexibility. It should be noted that the maximum optimality gaps observed for (s_t, Q)-SDP (24.9% and 23.8%) concern empirical demand patterns with a series of 0 demand. A single order quantity for all periods causes either a large amount of holding cost for 0-demand periods or penalty cost for large-demand periods. Despite the unsatisfactory performance of the (s_t, Q) policy, it is worth noting that the results show the same trends with respect to ρ, b and z as the (s_t, Q_t) policy.

The optimality gaps for the (R_t, S_t)-Heuristic are 2.90% on average, which is larger than the optimality gap observed for the (s_t, Q_t)-Heuristic over all demand patterns and pivoting parameters. As a result, we conclude that in the context of our test set the (s_t, Q_t) is better than (R_t, S_t) policy in terms of expected cost.

Table 7: Average percent optimality gap over our 25-period test set under different demand patterns and pivoting parameters.

Problem Settings	(s_t, Q_t)-Heuristic	(R, S)-Heuristic	(s_t, Q)-SDP	(s_t, Q)-Heuristic
demand pattern				
LCY1	2.38	2.50	9.56	10.5
LCY2	2.20	2.20	7.06	7.60
SIN1	2.52	2.87	6.25	8.06
SIN2	2.00	2.03	3.29	3.79
STA	1.45	1.50	1.91	2.25
RAND	2.58	2.99	7.24	8.98
EMP1	2.62	3.19	12.5	13.3
EMP2	2.50	4.22	24.9	25.9
EMP3	2.19	2.79	8.73	9.49
EMP4	2.70	4.71	23.8	25.3
std parameter				
0.1	2.52	2.68	10.3	11.4
0.2	2.48	2.50	11.0	11.9
0.3	1.94	3.53	10.3	11.3
fixed ordering cost				
500	2.71	3.36	13.8	14.7
1000	1.86	2.61	9.97	10.8
1500	2.35	2.69	7.70	8.90
penalty cost				
5	2.15	2.37	8.79	9.93
10	2.17	2.97	10.8	11.6
20	2.62	3.37	12.0	13.0
unit cost				
0	2.53	2.47	11.7	12.7
1	2.10	3.33	9.32	10.3
Average	2.31	2.90	10.5	11.5

6 Conclusion

This paper investigated (s, Q)-type policies for the non-stationary stochastic lot-sizing problem. By adopting a variant of the Bookbinder and Tan (1988) static-dynamic uncertainty strategy in which order quantities are fixed once and for all at the beginning of the planning horizon, we derived a stochastic dynamic formulation for the problem and proved that the associated optimal policy must take the (s, Q) form.

To compute optimal policy parameters, we enumerated all possible order quantity configurations to determine an optimal one, and then used a dynamic programming recursion to determine associated reorder points. Since this brute force approach is not scalable, we introduce MINLP-based heuristics to tackle large-size problems under (s, Q)-type policies. Our heuristics leverage the MINLP approaches introduced in Xiang et al. (2018) for the non-stationary (s, S) policy, in which the non-linearity is dealt with via a piecewise linearisation of the cost function.

We carried out extensive computational experiments on a test set of small problems with short (6-period) planning horizons and a test set of large problems with long (25-period) planning horizons. Both test sets include 10 demand patterns and various coefficient settings. In the numerical study on small problems, our results show that the average optimality gaps for the (s_t, Q_t) policy and the (s_t, Q) policy versus the optimal (s_t, S_t)-SDP are 1.91% and 3.61%, respectively; and the optimality gaps associated with (s_t, Q_t)-Heuristic and (s_t, Q)-Heuristic (2.76% and 4.66%, respectively) are close to those of the corresponding SDP.

In the numerical study on large problems, we found that the average optimality gaps of the (s_t, Q_t)-Heuristic remained small (2.31%); while the optimality gap of the (s_t, Q)-Heuristic remained unsatisfactory (11.5%). Our comparison against the (R_t, S_t)-Heuristic showed that the optimality gap of the (s_t, Q_t)-Heuristic was slightly better than that of the (R_t, S_t)-Heuristic (2.90%).

Our investigation demonstrates the effectiveness of (s, Q)-type policies for the non-stationary stochastic lot-sizing problem. The (s_t, Q_t) policy can be well approximated by a heuristic that provides satisfactory results in reasonable time. (Near-)optimal parameters for the $(s_t,\ Q)$ policy can be found in a reasonable time using either SDP or a heuristic, but it produces larger optimality gaps than the (s_t, Q_t) policy.

A Notations

Table A1: Notations of important functions

Functions	Explaination
$c(Q)$	cost of an order of size Q
$L_t(y)$	expected immediate holding and penalty cost when the inventory level after replenishment is y at period t
$C_t(x)$	expected total cost of the optimal policy over periods t to T when the opening inventory level is x
$G_t(y)$	expected total cost over periods t to T when the opening inventory level is y and no order is placed in period t
$V_t(x, \boldsymbol{q}_t)$	expected total cost with a combination of reorder quantities $\boldsymbol{q}_t \in \mathcal{Q}_t$ when the opening inventory level is x
$V_0(x)$	minimum expected total cost over $\mathcal{Q}$, the set of possible order quantities, when opening inventory level is x
$J_t(x, \boldsymbol{q}_t)$	expected total cost with a combination of reorder quantities $\boldsymbol{q}_t \in \mathcal{Q}_t$ when no order is placed for opening inventory level x in period t
$\hat{J}_t(x, \boldsymbol{q}_t)$	expected total cost with a combination of reorder quantities $\boldsymbol{q}_t \in \mathcal{Q}_t$ when an order is placed for opening inventory level x in period t
$\Delta J_t(x, \boldsymbol{q}_t)$	$= J_t(x, \boldsymbol{q}_t) - J_t(x + Q_t, \boldsymbol{q}_t)$, the difference between expected total costs with opening inventory levels x and $x + Q_t$
$\mathcal{J}_t(x, \hat{\boldsymbol{q}})$	an approximation of $J_t(x, \boldsymbol{q}_t^*)$ by MINLP

B MINLP model to compute S_t

This appendix section presents the MINLP model introduced in (Xiang et al., 2018) to compute the order-up-to level S_t of the (s_t, S_t) policy. To properly account for the proportional ordering cost z, we modify the objective function in line with Tarim and Kingsman (2006). We apply a superscript 'S' to distinguish decision variables from other formulations.

$$\min \quad z(\tilde{I}_T^S + \tilde{d}_{tT}) + \sum_{k=t}^{T}(K\delta_k^S + Q_k^S + h \cdot \tilde{H}_k + b \cdot \tilde{B}_k),$$

$$\text{s.t.} \quad \delta_t^S = 1, \tag{B1}$$

$$\tilde{I}_t^S + \tilde{d}_t = S_t, \tag{B2}$$

$$\delta_k^S = 0 \to \tilde{I}_k^S + \tilde{d}_k = \tilde{I}_{k-1}^S, \qquad k = t+1, \ldots, T, \tag{B3}$$

$$\delta_k^S = 1 \to \tilde{I}_k^S + \tilde{d}_k = \tilde{I}_{k-1}^S + Q_k^S, \qquad k = t+1, \ldots, T, \tag{B4}$$

$$\sum\nolimits_{j=t}^{k} P_{jk}^S = 1, \qquad k = t+1, \ldots, T, \tag{B5}$$

$$P_{jk}^S \geq \delta_j^S - \sum_{r=j+1}^{k} \delta_r^S, \qquad k = t, \ldots, T \text{ and } j = t, \ldots, k, \tag{B6}$$

$$P_{jk}^S = 1 \to \tilde{H}_k = \hat{\mathcal{L}}(\tilde{I}_k^S + \tilde{d}_{jk}, d_{jk}), \qquad k = t, \ldots, T \text{ and } j = t, \ldots, k, \tag{B7}$$

$$P_{jk}^S = 1 \to \tilde{B}_k = \mathcal{L}(\tilde{I}_k^S + \tilde{d}_{jk}, d_{jk}), \qquad k = t, \cdots, T \text{ and } j = t, \ldots, k, \tag{B8}$$

$$Q_k^S, \tilde{H}_k, \tilde{B}_k \geq 0, \qquad k = t, \ldots, T, \tag{B9}$$

$$P_{jk}^S, \delta_k^S \in \{0, 1\}, \qquad k = t, \ldots, T \text{ and } j = t, \ldots, k. \tag{B10}$$

We add constraints (B1) and (B2) to force the system to place an order in the first period of the horizon (t, T) in order to approximate S_t. The other constraints remain as in Xiang et al. (2018). Constraints (B3) and (B4) capture the inventory flow balance equations and reorder conditions. Constraint (B6) forces $P^S_{jk} = 1$ if the most recent replenishment before period k in horizon (t, k) is placed in period j; constraint (B5) ensures $P^S_{jk} = 0$ otherwise. Constraints (B7) and (B8) model the expected inventory and back-order levels at the end of period k through first order loss functions.

C Piecewise approximation with non-stationary Poisson demand

Consider a random variable ω and a scalar variable x, the first order loss function is defined as $\mathcal{L}(x, \omega) = \mathbb{E}[\max(\omega - x, 0)]$ and its complement as $\hat{\mathcal{L}}(x, \omega) = \mathbb{E}[\max(x - \omega, 0)]$. Decision variables $\tilde{H}_t \geq 0$ and $\tilde{B}_t \geq 0$ denote the expected inventory and back-order levels at the end of period t.

Rossi et al. (2014) presented the approach with bounding techniques to generate piecewise linear lower and upper bounds and discussed the implementation on the standard normal distribution. Instances in this paper involve non-stationary Poisson demand to enable the computation analysis on problems with small means of demand. Therefore, we extend the results of Rossi et al. to the Poisson distribution.

To minimise the expected inventory and back-order levels at the end of each period with a lower bounding piecewise linear approximation, $\tilde{H}_t$ is constrained by

$$\tilde{H}_t \geq (\tilde{I}_t + \sum_{j=1}^{t} \tilde{d}_{jt} P_{jt}) \sum_{k=1}^{i} p_k + \sum_{j=1}^{t} (\sum_{k=1}^{i} p_k \mathbb{E}[d_{jt}|\Omega_{jt}]) P_{jt}, \tag{C1}$$

and $\tilde{B}_t$ by

$$\tilde{B}_t \geq -\tilde{I}_t + (\tilde{I}_t + \sum_{j=1}^{t} \tilde{d}_{jt} P_{jt}) \sum_{k=1}^{i} p_k + \sum_{j=1}^{t} (\sum_{k=1}^{i} p_k \mathbb{E}[d_{jt}|\Omega_{jt}]) P_{jt}. \tag{C2}$$

where d_{jt} follows the notation in section 4.1 denoting the convolution of d_j to d_t, demand d_t is a random variable that is of a Poisson distribution with mean λ_t, and its domain $\mathbb{R}^+$ is partitioned into N disjoint adjacent subregions $\Omega_1, \Omega_2, \cdots, \Omega_N$.

According to the technique in (Rossi et al., 2014), $\Omega_1 = [0, a_1]$, $\Omega_i = [a_{i-1}, a_i]$ for $i = 2, \cdots, N-1$ and $\Omega_N = [a_{N-1}, \infty]$. Let the probability density function of d_t be $g_{\lambda_t}(k) = e^k/\lambda_t!$ and $g^{-1}_{\lambda_t}(p)$ be its inverse function, which returns the value of k satisfying $g_{\lambda_t}(k) = p$, then

$$a_i = g^{-1}_{\lambda_t}(\frac{i}{N}),$$

and the probability p_i that a realisation of the Poisson random variable d_t (i.e. a value of demand d_t) locates within the subregion i is

$$p_i = \Pr\{d_t \in \Omega_i\} = \int_{\Omega_i} g_{\lambda_t}(u) \, \mathrm{d}u, \tag{C3}$$

and

$$\mathbb{E}[d_t|\Omega_i] = \frac{N}{i} \int_{\Omega_i} u g_{\lambda_t}(u) \, \mathrm{d}u, \tag{C4}$$

where $i = 1, 2, \cdots, N$.

References

Askin, R. G. (1981). A procedure for production lot sizing with probabilistic dynamic demand. *AIIE Transactions*, 13(2):132–137.

Bellman, R. (1957). *Dynamic Programming.* Princeton University Press, Princeton, NJ, USA.

Belotti, P., Bonami, P., Fischetti, M., Lodi, A., Monaci, M., Nogales-Gómez, A., and Salvagnin, D. (2016). On handling indicator constraints in mixed integer programming. *Computational Optimization and Applications*, 65(3):545–566.

Bollapragada, S. and Morton, T. E. (1999). A simple heuristic for computing nonstationary (s, S) policies. *Operations Research*, 47(4):576–584.

Bookbinder, J. H. and Tan, J.-Y. (1988). Strategies for the probabilistic lot-sizing problem with service-level constraints. *Management Science*, 34(9):1096–1108.

Dural-Selcuk, G., Rossi, R., Kilic, O. A., and Tarim, S. A. (2020). The benefit of receding horizon control: Near-optimal policies for stochastic inventory control. *Omega*, 97: 102091 .

Gallego, G. and Toktay, L. B. (2004). All-or-nothing ordering under a capacity constraint. *Operations Research*, 52(6):1001–1002.

Özen, U., Doğru, M. K., and Tarim, S. A. (2012). Static-dynamic uncertainty strategy for a single-item stochastic inventory control problem. *Omega*, 40(3):348–357.

Rossi, R., Kilic, O. A., and Tarim, S. A. (2015). Piecewise linear approximations for the static–dynamic uncertainty strategy in stochastic lot-sizing. *Omega*, 50:126–140.

Rossi, R., Tarim, S. A., Prestwich, S., and Hnich, B. (2014). Piecewise linear lower and upper bounds for the standard normal first order loss function. *Applied Mathematics and Computation*, 231:489–502.

Scarf, H. E. (1960). Optimality of (s, S) policies in the dynamic inventory problem. In Arrow, K. J., Karlin, S., and Suppes, P., editors, *Mathematical Methods in the Social Sciences*, pages 196–202. Stanford University Press, Stanford, CA.

Silver, E. (1978). Inventory control under a probabilistic time-varying, demand pattern. *AIIE Transactions*, 10(4):371–379.

Silver, E., Pyke, D., and Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling*. Wiley, New York, 3 edition edition.

Silver, E. A. and Meal, H. C. (1973). A heuristic for selecting lot size quantities for the case of a deterministic time-varying demand rate and discrete opportunities for replenishment. *Production and Inventory Management*, 14(2):64–74.

Sox, C. R. (1997). Dynamic lot sizing with random demand and non-stationary costs. *Operations Research Letters*, 20(4):155–164.

Strijbosch, L. W., Syntetos, A. A., Boylan, J. E., and Janssen, E. (2011). On the interaction between forecasting and stock control: The case of non-stationary demand. *International Journal of Production Economics*, 133(1):470–480.

Tarim, S. A., Dogru, M. K., Özen, U., and Rossi, R. (2011). An efficient computational method for a stochastic dynamic lot-sizing problem under service-level constraints. *European Journal of Operational Research*, 215(3):563–571.

Tarim, S. A. and Kingsman, B. G. (2004). The stochastic dynamic production/inventory lot-sizing problem with service-level constraints. *International Journal of Production Economics*, 88(1):105–119.

Tarim, S. A. and Kingsman, B. G. (2006). Modelling and computing (R^n, S^n) policies for inventory systems with non-stationary stochastic demand. *European Journal of Operational Research*, 174(1):581–599.

Tunc, H., Kilic, O. A., Tarim, S. A., and Rossi, R. (2018). An extended mixed-integer programming formulation and dynamic cut generation approach for the stochastic lot-sizing problem. *INFORMS Journal on Computing*, 30(3):492–506.

Vargas, V. (2009). An optimal solution for the stochastic version of the Wagner–Whitin dynamic lot-size model. *European Journal of Operational Research*, 198(2):447–451.

Visentin, A., Prestwich, S., Rossi, R., and Tarim, S. A. (2021). Computing optimal (R, s, S) policy parameters by a hybrid of branch-and-bound and stochastic dynamic programming. *European Journal of Operational Research*.

Wagner, H. M. and Whitin, T. M. (1958). Dynamic version of the economic lot size model. *Management Science*, 5(1):89–96.

Xiang, M., Rossi, R., Martin-Barragan, B., and Tarim, S. A. (2018). Computing non-stationary (s, S) policies using mixed integer linear programming. *European Journal of Operational Research*, 271(2):490–500.

Zheng, Y. and Federgruen, A. (1991). Finding optimal (s, S) policies is about as simple as evaluating a single policy. *Operations Research*, 39(4):654–665.